Doctoral theses at NTNU, 2021:201

Davit Gigilashvili

On the Appearance of Translucent Objects: Perception and Assessment by Human Observers

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of Science and Technology

Davit Gigilashvili

On the Appearance of Translucent Objects: Perception and Assessment by Human Observers

Thesis for the Degree of Philosophiae Doctor

Gjøvik, June 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Computer Science

ISBN 978-82-326-5547-2 (printed ver.)
ISBN 978-82-326-5981-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:201

Printed by NTNU Grafisk senter

Abstract

Appearance characterizes visual features of objects and materials. It is a multiplex psychovisual phenomenon that is usually broken into several appearance attributes for simplification of its measurement and communication, and for studying its nature. Color, texture, gloss, and translucency are considered the major appearance attributes. Significant research work has been done in metrology for accurate instrumental measurement of optical properties of materials, and considerable advances have been made in computer graphics, permitting the generation of highly photorealistic visual stimuli. Nevertheless, the knowledge remains limited on how humans perceive appearance, how we behave to assess appearance, what factors impact our perception, how different attributes interact with each other, and all in all how optical properties relate with their perceptual counterparts.

In this thesis, we explore various aspects of appearance perception with a focus on the appearance of translucent objects. For this purpose, we conducted a series of social and psychophysical experiments with real and synthetic visual stimuli. Elucidating appearance perception of translucent objects has implications for industrial, academic and artistic applications alike.

In the initial stage of the study, we organized a social experiment in order to collect qualitative observations on the process of appearance assessment, construct a qualitative model of material appearance and generate relevant research hypotheses. The hypotheses have been analyzed in context of the state-of-the-art.

Afterwards, we tested the most interesting hypotheses quantitatively, in order to assess their generalization prospects. The experimental results have provided indications in support of the hypotheses. We have observed that translucency of an object impacts perception of glossiness, while detection of translucency difference depends on geometric thickness of the objects and optical thickness of the materials they are made of. Additionally, we examined a potential role of several cues in translucency perception that are present in the image detected by either a camera or a human observer. We found that blurriness of the image and the presence of caustics can impact apparent translucency.

Finally, we conducted a comprehensive survey on translucency perception, advancing the state-of-the-art with our findings, and outlining unanswered questions for future research.

Sammendrag

Utseende karakteriserer visuelle egenskaper ved gjenstander og materialer. Det er et mangfoldig psykovisuelt fenomen som vanligvis blir brutt ned til flere utseendeattributter, for å forenkle dets måling og kommunikasjon, og studering av dets natur. Farge, tekstur, glans og gjennomskinnelighet anses som de viktigste utseendeattributtene. Det er gjort betydelig forskningsarbeid innen metrologi for nøyaktig instrumentell måling av materialers optiske egenskaper, og betydelige fremskritt innen datagrafikk som tillater generering av meget fotorealistiske visuelle stimuli. Likevel er kunnskapen fortsatt begrenset om hvordan mennesker oppfatter utseende, hvordan vi oppfører oss for å vurdere utseende, hvilke faktorer som påvirker vår oppfatning, hvordan forskjellige attributter innvirker på hverandre, og alt i alt hvordan optiske egenskaper relateres til deres perseptuelle motstykker.

I denne avhandlingen utforsker vi ulike persepsjonsaspekter med fokus på utseendet til gjennomskinnelige objekter. For dette formålet gjennomførte vi en serie sosiale og psykofysiske eksperimenter med ekte og syntetiske visuelle stimuli. Kunnskap om uteseende til gjennomskinnelige gjenstander har implikasjoner for både industrielle, akademiske og kunstneriske anvendelser.

I den innledende fasen av studien gjennomførte vi et sosialt eksperiment for å samle kvalitative observasjoner om prosessen med utseendevurdering, konstruere en kvalitativ modell for materialutseende og frembringe relevante forskningshypoteser. Hypotesene er analysert i sammenheng med kunnskapsfronten.

Etterpå testet vi de mest interessante hypotesene kvantitativt, for å vurdere deres muligheter for generalisering. De eksperimentelle resultatene har gitt indikasjoner til støtte for hypotesene. Vi har observert at et objekts gjennomskinnelighet påvirker oppfatningen av glans, mens deteksjon av gjennomskinnelighetsforskjeller avhenger av gjenstandenes geometriske tykkelse og materialene de er laget av sin optiske tetthet. I tillegg har vi undersøkt rollen til flere potensielle perseptuelle indikatorer for gjennomskinnelighet, som kan finnes i bilder som er registrert enten av et kamera eller av en menneskelig observator. Vi har funnet at bildeuskarphet og kaustikk kan påvirke oppfattelsen av gjennomskinnelighet.

Til slutt gjennomførte vi en omfattende undersøkelse om perseptuell gjennomskinnelighet, oppdaterte kunnskapsfronten med våre funn, og skisserte ubesvarte spørsmål for fremtidig forskning.

Dedication

Dedicated to the memory of my father Alexander & my grandma Lamara. Thank you for all that I am today!..

Acknowledgements

The work I have carried out could have been infeasible without the support and contributions from the people around me.

First and foremost, I would like to express my gratitude to my supervisors Prof. Jon Yngve Hardeberg, Prof. Marius Pedersen and Assoc. Prof. Jean-Baptiste Thomas for their invaluable feedback that always helped me keep on the right track and orient myself towards the right directions.

Secondly, I want to express special gratitude to Prof. Holly Rushmeier for being the opponent on my pre-defense session and for her guidance during my stay at Yale University. I also want to thank all my co-authors Philipp Urban, Midori Tanaka, Weiqi "Justin" Shi, Zeyu "Zach" Wang, Lucas Dubouchet, Fereshteh Mirjalili and others for fruitful collaborations and for bringing interesting perspectives into the project. I want to show my appreciation to the anonymous reviewers for their comments that significantly increased the scientific value of my publications and to all observers who dedicated their valuable time to the participation in psychophysical experiments and thus, altruistically contributed to science. Likewise, I thank Dali Khomeriki, an educational advisor of mine, whose competent consultations made my overall academic journey a lot smoother.

Thirdly, the project would not have happened without the support of the Research Council of Norway. The research has been funded by the Measuring and Understanding Visual Appearance - MUVApp (#250293) and Material Appearance Network for Education and Research - MANER (#288187) projects. I want to thank the Kingdom of Norway for promoting research.

And last but definitely not least, the personal support from my loved ones has been of vital importance. I want to wholeheartedly thank my beloved wife Ana for the infinite inspiration and care, as well as for her invaluable advise on linguistic matters. I also want to express my gratitude to my uncles Vazha and David, and aunt Mzia for encouraging the constant pursuit of new knowledge, for motivation, guidance, moral and material support at different stages of my education.

Contents

Chapter 1

Introduction

1.1 Motivation

Vision is one of the fundamental senses human beings rely on for interpreting their surrounding. Appearance is a visual sensation attributing particular properties to surrounding objects and materials. Based on how they look, we can tell whether food is fresh or spoiled, whether a sidewalk is slippery or not, or whether a cup is made of soft and elastic plastic or rigid and fragile glass. We are surprisingly good at assessing appearance and deducing material properties from it. The sensation of appearance impacts a broad range of our behaviors, from performing simple daily routines to making choices between lavish consumer products. Therefore, understanding how to acquire, reproduce and communicate appearance has considerable implications for academia, industry and arts alike.

Appearance is a result of light interacting with different objects and materials in a scene. While instrumental measurement (hard metrology) (Pointer (2003) and Choudhury (2014)) and digital modeling of optical material properties (Dorsey et al. (2010)) have advanced considerably, the physical material properties remain poor predictors of what humans perceive, as our understanding of how our visual system perceives appearance remains limited. This gave rise to the development of soft metrology – an attempt of finding a correlation between objective measures and subjective human responses, where the paramount goal is to come up with a measurement scale which will predict subjective response based on objectively measurable quantities (Pointer (2003), Eugène (2008), and Leloup et al. (2014)).

Appearance is a complex psychovisual phenomenon. In order to simplify quantification and studying its nature, appearance is usually broken into distinct appearance attributes, color, gloss, translucency and texture being usually the most significant and prevalent ones (CIE (2006) and Eugène (2008)). Color is undeniably the most salient, as well as the most studied appearance attribute. Color science has a long history and the mechanisms of color perception are relatively well understood. However, the same cannot be said about other appearance attributes. Appearance research has emerged from and can be considered an extension of color science (Sole et al. (2019)). Translucency is among the most understudied albeit significant attributes of appearance (Anderson (2011)). We interact with translucent objects and materials on a daily basis, which in addition to food, beverages, countless plastic, glass, wax and paper products, also includes our own skin. Translucency helps us distinguish fresh juicy food from dry spoiled ones (Di Cicco et al. (2020b)), metals from glass, or human skin from plastic dummies. Proper reproduction of the appearance of translucent objects is critical in many fields, such as 3D printing (Brunton et al. (2018) and Urban et al. (2019)), cultural heritage (Kaltenbach (2012) and Barry (2011)), architecture (Murray (2013) and Kaltenbach (2012)) (see Figure 1.1), computer graphics (Frisvad et al. (2020) and Nunes et al. (2019)), cosmetology (Giancola and Schlossman (2015) and Emmert (1996)), aesthetic dentistry (Liu et al. (2010) and Lopes Filho et al. (2012)), food industry (Hutchings (1977) and Hutchings (2011)) and visual arts (Wijntjes et al. (2020), Di Cicco et al. (2020a), and Di Cicco et al. (2020b)) – making research on translucent objects and materials largely interdisciplinary. The standards for measuring particular optical properties, such as the extinction coefficient, clarity or haze, might differ among industries (Pointer (2003), Dorsey et al. (2010), and Frisvad et al. (2020)), but they all suffer from the common problem – physical measurements are poor predictors of what humans perceive. Furthermore, measurements are conducted for small sets of materials, objects and illumination conditions, and little is known how appearance varies in the complex and dynamic environment we usually interact with the objects and materials in. The research on translucency perception will help us identify these links between the physical and the perceptual properties, which is relevant for all above-mentioned fields. In the industries, where the visual appearance of the products has enormous significance, such as the industries of food, fashion, cosmetics, electronics and other accessories, understanding how the appearance of translucent objects is perceived by the customers will enable the manufacturers predict, produce and replicate the desired appealing looks. In arts and cultural heritage, understanding perception will not only facilitate designing, but also the conservation, restoration, archiving and cross-media reproduction processes. The development of the perception-aware material mixing or rendering algorithms in the rapidly emerging fields of 3D printing and computer graphics, respectively, will make it possible to generate more realistic visual effects in more cost-effective ways. Understanding visual perception of translucent materials in the dynamic and varying environment will be especially important in the extended reality applications - e.g. for achieving the realistic telepresence.

Translucency implies that light penetrates the material, propagates through it and emerges from a different part of it. Therefore, image structure detected at the human retina can result from an infinite number of combinations between surface reflection and subsurface transport of light. While disentangling these contributions and understanding the complex process of light and matter interaction is an ill-posed problem, the human visual system (HVS) manages to deduce the properties of translucent objects in a surprisingly consistent and robust manner (Anderson (2011) and Fleming and Bülthoff (2005)). The exact mechanisms of this ability are yet to be unearthed.

The fact that material appearance research is in an early phase of its development, with yet ample unknowns, motivated us to observe the process of material appearance assessment by humans with an objective to generate relevant research hypotheses and to pave the way for future research. Afterwards, we aimed our attention at a particular subset of visual stimuli – translucent materials and objects made of them. We explored not only translucency proper as an appearance attribute, but also the perception of glossiness on translucent objects. We want to highlight the following: although translucency as an optical phenomenon is a property of materials, we usually view and interact with different objects that are made of those materials. In addition to optical properties, geometric properties of an object, such as shape, roughness and size, also impact what we perceive. Therefore, in the rest of this thesis, we discuss perceiving the translucency of particular objects, not that of materials as abstract entities.

Finally, while computer graphics enables us to manipulate material and object's properties in an easy, cheap and systematic manner, manufacturing physical objects that cover a broad range of materials is a substantially harder task. On the other hand, computer graphics which suffers from a lower dynamic range and lacks interactivity, tactile information and binocular vision, does not fully emulate the natural experience we usually have in our daily lives. The need for an inevitable trade-off prompted us to conduct our study both on real and digital stimuli, which itself can reveal intriguing differences between the media.

This fundamentally interdisciplinary work, which incorporates components from computer science, social science, vision science and experimental psychology, has implications for a broad range of fields, such as 3D printing, computer graphics and even visual arts (Hodgson (2020)).

1.2 Research Objectives

The preeminent goal of this work is to unveil the visual mechanisms of material appearance and to find the correlation between physical and perceptual properties, with particular emphasis on, but not limited to, translucent materials and objects made of them. Considering the complex nature of the problem, we believe the goal should be reached incrementally, by generating interesting hypotheses, followed either by their falsification or inability thereof. Consequently, we divided the project into distinct parts according to four major objectives:

First of all, we aimed for constructing a qualitative model of material appearance and generating relevant research hypotheses, which if supported by the state-of-the-art and validated quantitatively, would enable us to generalize our observations incrementally. Although translucent objects remain the focus of this thesis, the objective at this stage has been to observe the process of assessing material appearance in general, to provide a bigger picture and to propose hypotheses both on translucent and non-translucent objects.

Figure 1.1: Beinecke Rare Book and Manuscript Library is located on Yale University campus, in New Haven, Connecticut. It was designed by Gordon Bunshaft and the construction was completed in 1963. The library is built with translucent marble panels. This is a vivid example of using translucent building materials in modern architecture and respective visual appearance generated with that. While the panels look opaque most of the time (the left wall in the image), they start to transluce and glow (the right wall) as soon as direct sunlight hits them. The visual effect is achieved with a phenomenon that objects look more translucent when they are back-lit. [Photo by Davit Gigilashvili]

Secondly, we tested the interesting hypotheses about interactions between translucency and other appearance properties, such as geometric shape and perceived glossiness.

Afterwards, we attempted to identify how information about material appearance (namely, translucency and glossiness) is encapsulated in the image structure.

Finally, we concentrated on translucency as an appearance attribute. The objective at this stage has been to analyze the findings, use them to advance the state-of-the-art about translucency perception and to outline future steps needed for reaching the preeminent goal.

It is worth noting that the objective of this thesis is limited neither to translucency perception, nor the appearance of translucent objects. Translucency coexists with other appearance attributes, being a piece of a puzzle in a picture of total appearance. We started from a general topic and narrowed our focus as the work progressed and more data was being obtained. This is summarized in Table 1.1.

1.3 Research Questions

The details regarding the generated hypotheses and research questions are summarized in Chapter 3. Below we enlist the pivotal research questions for this work. How these research questions serve to the four objectives discussed above is shown in Figure 1.2.

1. How do people behave when assessing appearance, and which factors facilitate this process?

Table 1.1: We started collecting experimental observations on material appearance assessment in general. Gradually narrowing the focus, we tested the hypotheses quantitatively and eventually surveyed the updated state-of-the-art on a particular topic of translucency perception.

	Objective	**Appearance Attributes Addressed**	**Level of Generality**
1	Hypotheses generation	Virtually any	Appearance in general
2	Interaction of translucency, gloss and shape	Translucency, gloss	Two attributes
3	Impact of image structure on apparent translucency and gloss	Translucency, gloss	Two attributes
4	Knowledge status in translucency perception	Translucency	Focus on a single attribute

2. Does the human visual system manifest constancy in translucency perception similarly to color constancy, and to what extent?
3. Does translucency contribute to glossiness perception?
4. Does the shape of the object impact the perceived magnitude of translucency?
5. Does the shape of the object impact detection of translucency differences?
6. Does the magnitude of subsurface scattering impact our ability to detect translucency differences?
7. Does appearance assessment differ between physical objects and displayed images, and how vital is the direct interaction with the objects when judging their appearance?
8. Does the presence of caustics impact the perceived magnitude of translucency?
9. Does image blur impact the perceived magnitude of translucency?
10. Can the luminance statistics be used for prediction of apparent gloss and translucency?
11. What are the major obstacles to advancing translucency perception research?
12. What is the knowledge status on translucency perception and where should we go next?

1.4 Research Methodology

1.4.1 Methods used in the project

The initial stage of the project was dedicated to qualitative research using an inductive research method. We started the project with a qualitative research

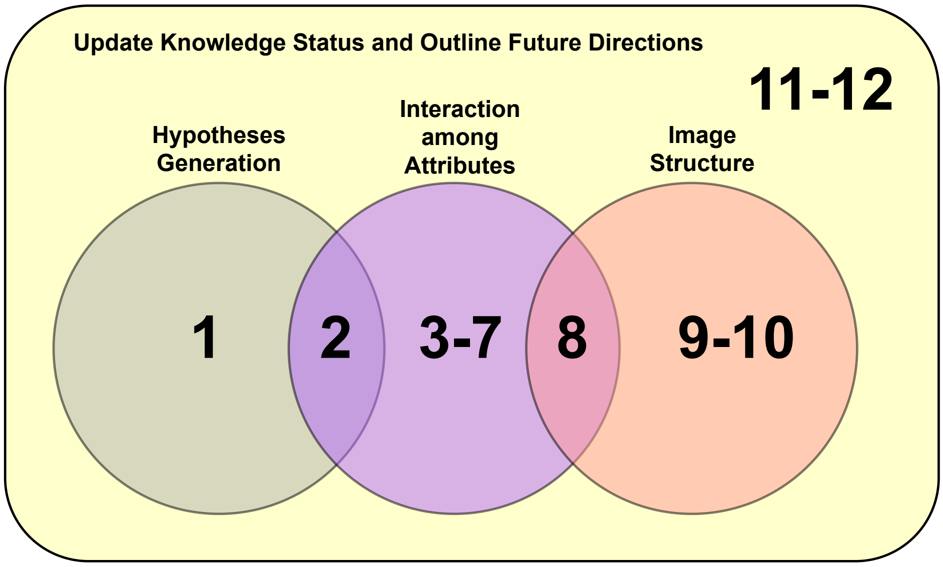

Figure 1.2: The figure summarizes how the research questions relate to the objectives of the project. The numbers correspond to the respective research questions. For instance, research question 2 about translucency constancy helps us generate research hypotheses and also understand how translucency interacts with other attributes. All research questions, including 11-12, serve the objective to update the knowledge status on translucency perception and identify the avenues worth taking in the future.

methodology with an intention of:

- Building a qualitative model of material appearance that is rooted in the experimental data. While qualitative models are usually based on subjective interpretation by the authors and their philosophical rationales, to the best of our knowledge, no model exists that is fully rooted in experimental data.
- Generating relevant research hypotheses for future deductive studies, which, if validated with quantitative research methods, will help the generalization of the model.

We hypothesize that appearance is a social interaction, either a human subject interacting with the object in a scene, or two subjects communicating the appearance. Therefore, we approached the problem with a methodology from social science and conducted a social experiment to observe this interaction. As well-noted by Anderson (2011), the experimental scenes are usually oversimplified, creating a risk that the experimenters remove information essential to the visual system and "*those experiments may provide little insight into the normal functioning of the visual system*". In order to see a broader picture of the appearance assessment process and make the interaction as close as possible to natural everyday behavior, unrestricted interaction with the objects was permitted and the experimental conditions have not been fixed.

The process was videotaped and the transcripts have been analyzed with the Grounded Theory Analysis (GTA) (Paillé (1994)). The GTA is an inductive research method derived from the Grounded Theory Approach (Glaser et al. (1968) and Corbin and Strauss (2015)) in social science. The method consists of six stages of analysis:

1. Coding – assigning labels to all experimental observations.
2. Categorization – grouping conceptually similar observations into categories.
3. Co-linking – identifying how different categories relate to each other.
4. Integration – putting the categories into a single system and reinforcing the original links with additional data which could be either quantitative frequency analysis or the overview of the state-of-the-art.
5. Modelling – creating a model that describes the underlying structure of the data.
6. Theorization – creating a provisional theory, which is far from a general theory, but is conceptually and structurally more advanced than a mere description of observations.

The examples of using this methodology for addressing numerous social aspects can be found in the works by Jacob and Holmes (2011), Gaucher and Payot (2011), and Rippon et al. (2020). In parallel to qualitative analysis, quantitative frequency analysis was also conducted to augment and strengthen the qualitative observations - more specifically, to identify the most common observations and to formulate research hypotheses based on them. Afterwards, the literature has been reviewed and the observations have been scrutinized in the context of the state-of-the-art.

At the second stage of the project, the most relevant hypotheses were tested quantitatively. We conducted psychometric scaling experiments (Engeldrum (2000)) and tried to correlate physical material properties with the perception of particular attributes among the human observers, as well as to measure the statistical significance of these correlations. Several experimental setups were used in different studies, including pair-wise comparisons (***Articles D*** and ***G***), rank order (***Articles A*** and ***C***), category judgment (***Article F***), and the method of constant stimuli (***Article E***). The visual stimuli have been presented: as physical objects (***Articles A, B*** and ***C***), computer-generated imagery (***Articles D, E*** and ***F***) or RGB images (***Article G***). Additionally, image statistics of the RGB photographs, particularly the first four moments of luminance histogram, were also analyzed (***Article H***) to understand how changes in optical properties and visual appearance are reflected in the image structure.

Finally, an exhaustive literature review was produced that advanced the state-of-the-art with our findings obtained in the previous steps. Figure 1.3 illustrates how these fundamentally different methods fit together in the loop of generating new knowledge.

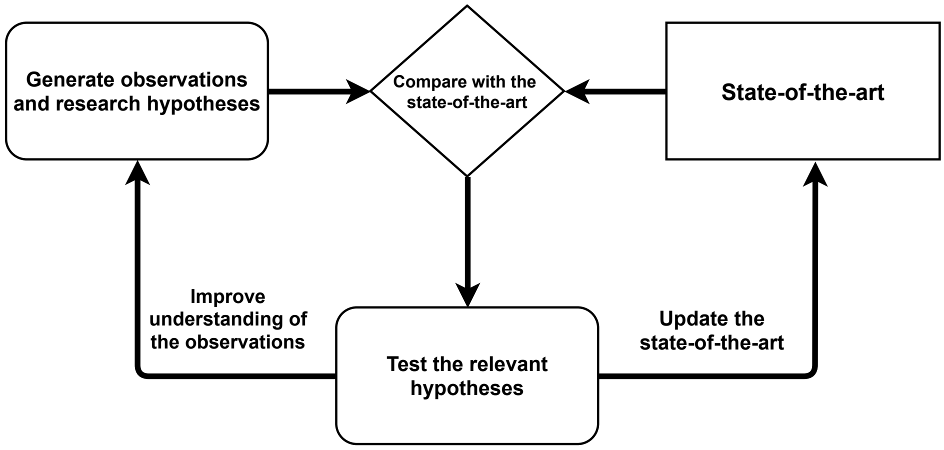

Figure 1.3: After generating new hypotheses and observations, they are compared with the state-of-the-art. The ones considered most relevant are tested quantitatively. Validation or falsification of the hypotheses helps us not only to update the knowledge status, but also improve our understanding of the original observations.

1.4.2 Rationale for using an inductive research method

We are aware that the current academic community is dominated by "hypothetico-deductive" research and the scepticism towards the methods based on Grounded Theory (GT) is not unheard in the scientific community (Luckerhoff and Guillemette (2011)). However, considering the interdisciplinarity and complexity of the problem, the research methodology has been chosen with full awareness of the latter fact. Below we will explain the rationale for using the inductive research method derived from the GT.

Luckerhoff and Guillemette (2011) have analyzed methodological peculiarities of the GT that are oftentimes reason for rejection of the GT-based research proposals by evaluation committees. However, we believe that this project took great advantage of these very features that are specific to this inductive method.

Typical quantitative studies test the research hypotheses by fixing particular optical properties of the materials while systematically varying others – trying to measure their impact on observer responses (Anderson (2011)). However, material appearance research is still in its infancy and little remains known about the complex process of behavioral and psychovisual mechanisms of material appearance assessment. This creates a fundamental problem that even before raising the question of how particular research hypotheses should be tested, first of all, we need to identify what those hypotheses are. When Glaser and Strauss (1965) introduced the GT method, they argued that some sociologists "*over-emphasize rigorous testing of hypotheses, and de-emphasize the discovering of what concepts and hypotheses are relevant for the substantive area being researched*". While a colossal area in the field of material appearance remains to be explored, generating research hypotheses and bringing new concepts to light is a valuable contribution in itself.

In traditional deductive studies a literature review is conducted before setting up an experiment, while in GT-based approaches reference to the literature is postponed in order to avoid prejudices and ensure a higher degree of openness among the experimenters. The observations are compared with the state-of-the-art once they are collected and a researcher is open to whatever emerges from the data rather than *"forcing the data to comply with existing theories"*. (Luckerhoff and Guillemette (2011))

Furthermore, while traditional research methodologies are linear by nature (proposing a hypothesis, setting up an experiment, testing the hypothesis, drawing the conclusions), the GT is characterized with circularity - as a constant refinement loop is allowed by the GTA, where every new piece of the data can be used to return to the original observations and improve their understanding.

We believe these peculiarities of the inductive research method are especially important for generating new unbiased ideas and guiding future research, which can be crucial for such an understudied field as material appearance. This is well summarized by Starrin et al. (1997): "*Usually you collect the data, then analyze them. When collecting theoretical puzzle pieces, you have no idea ahead of time what you will collect. Above all, you do not know where they will lead you. By discovering codes and trying to saturate them by seeking comparable groups, you get a growing feeling of where you should look for more data.*"

Finally, we want to mention that our research objectives could be to some limited extent reached with structure discovery techniques, such as multidimensional scaling (MDS). However, those techniques could not fully substitute the benefits of using GTA for the following reasons: first, GTA is a qualitative inductive research method, while MDS is a quantitative method of a deductive nature. When using structure discovery techniques, some hypotheses about the structure are assumed – for instance, in MDS we assume dimensionality. However, we refused to accept any pre-existing hypotheses due to above-discussed reasons. Secondly, structure discovery methods, such as MDS, deal with scale, numerical and ordinal data (e.g. similarity of the objects by glossiness). However, unlike GTA, they cannot measure and capture the complex socio-behavioral aspects of the interaction. Thirdly, quantitative structure discovery methods (such as MDS) require a high number of visual stimuli, which would have been impractical with physical objects. Using computer generated imagery as an alternative would have considerably limited the naturalness of the behavior due to a simpler environment and the lack of the interaction. Indeed, it is not to deny that the methods such as MDS are powerful tools for building reliable quantitative visual models, but on the other hand, the methods such as GTA, are more suitable for observing a broad range of the behavioral and social processes involved in appearance assessment. It is important to highlight that we neither consider these methods mutually exclusive alternatives, nor have we abstained from using the MDS – instead, we postponed it in time (MDS was later used in ***Article D***). We see GTA and quantitative methods as the methods suitable for different stages of the recursive process. We first generate observations and hypotheses free from assumptions and state-of-the-art bias, and only afterwards we validate them with the quantitative methods.

1.5 List of Articles

The thesis is based on 10 articles, out of which 9 have been either published or accepted for publication in the peer-reviewed publication channels, while the remaining 1 is awaiting the peer-review at a scientific journal. The publications are listed with alphabet-based enumeration, based on their occurrence in the thesis narrative. The articles come in four types: qualitative research, quantitative research, review and position paper. The experiments and/or visual demonstrations in the articles are based on three different types of visual stimuli: physical tangible objects the observers have been able to interact with, synthetic images generated with computer graphics and displayed on a monitor, and RGB photographs displayed on a monitor. The types of articles and the relation among them are illustrated in Figure 1.4. The content of the articles is summarized in Chapter 3.

The following articles are included in the thesis. Journal articles are shown in boldface, while conference articles are shown in regular typeface:

Article A Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2018). "Behavioral investigation of visual appearance assessment." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 294–299 DOI: `https://doi.org/10.2352/ISSN.2169-2629.2018.26.294`

Article B Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). "On the appearance of objects and materials: Qualitative analysis of experimental observations." In: *Accepted for publication in the Journal of the International Colour Association (JAIC)*, 33 pages

Article C Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (2019). "Perceived Glossiness: Beyond Surface Properties." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 37–42 DOI: `https://doi.org/10.2352/issn.2169-2629.2019.27.8`

Article D Davit Gigilashvili, Weiqi Shi, Zeyu Wang, Marius Pedersen, Jon Yngve Hardeberg, and Holly Rushmeier (2021). "The Role of Subsurface Scattering in Glossiness Perception." In: *ACM Transaction on Applied Perception* 18.3, 10:1–10:26 DOI: `https://doi.org/10.1145/3458438`

Article E Davit Gigilashvili, Philipp Urban, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). "The Impact of Optical and Geometrical Thickness on Perceived Translucency Differences." In: *Under review in a journal*, 13 pages

Article F Davit Gigilashvili, Lucas Dubouchet, Marius Pedersen, and Jon Yngve Hardeberg (2020). "Caustics and Translucency Perception." In: *Material Appearance 2020, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 033:1–033:6 DOI: `https://doi.org/10.2352/ISSN.2470-1173.2020.5.MAAP-033`

Article G Davit Gigilashvili, Marius Pedersen, and Jon Yngve Hardeberg (2018). "Blurring impairs translucency perception." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 377–382 DOI: `https://doi.org/10.2352/ISSN.2169-2629.2018.26.377`

Article H Davit Gigilashvili, Midori Tanaka, Marius Pedersen, and Jon Yngve Hardeberg (2020). "Image Statistics as Glossiness and Translucency Predictor in Photographs of Real-world Objects." In: *10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*. Vol. 2688. CEUR Workshop Proceedings, pp. 1–15

Article I Davit Gigilashvili, Jean Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2020). "On the Nature of Perceptual Translucency." In: *8th Annual Workshop on Material Appearance Modeling (MAM2020). Eurographics Digital Library*, pp. 17–20 DOI: `https://doi.org/10.2312/mam.20201141`

Article J Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (n.d.). "Translucency perception: A review." In: *Accepted for publication in the Journal of Vision*, 45 pages

Two of the above-mentioned works won the accolades. Namely, ***Article C*** has received the 2019 Best Student Paper Award at the 27th Color and Imaging Conference. ***Article F*** received the Best Paper Award at Material Appearance 2020 conference, IS&T International Symposium on Electronic Imaging.

1.6 Supporting Articles

In addition to 10 above-mentioned articles, 5 additional articles have been published within the course of the PhD program. Although those articles are not included as a part of the thesis, they play a supporting role. They have facilitated progress through the overall project and provided additional insight into the data. Therefore, we list them below, as we believe that some readers might find them interesting:

Article K Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (2019). "Material appearance: ordering and clustering." In: *Material Appearance 2019, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 202:1–202:6 DOI: `https://doi.org/10.2352/ISSN.2470-1173.2019.6.MAAP-202`

Article L Davit Gigilashvili, Philipp Urban, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2019). "Impact of Shape on Apparent Translucency Differences." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 132–137 DOI: `https://doi.org/10.2352/issn.2169-2629.2019.27.25`

Article M Davit Gigilashvili, Fereshteh Mirjalili, and Jon Yngve Hardeberg (2019). "Illuminance Impacts Opacity Perception of Textile Materials." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 126–131 DOI: https://doi.org/10.2352/issn.2169-2629.2019.27.24

Article N Aditya Sole, Davit Gigilashvili, Helene Midtfjord, Dar'ya Guarnera, Giuseppe Claudio Guarnera, Jean-Baptiste Thomas, and Jon Yngve Hardeberg (2019). "On the acquisition and reproduction of material appearance." In: *International Workshop on Computational Color Imaging*. Springer, pp. 26–38 DOI: https://doi.org/10.1007/978-3-030-13940-7_3

Article O Ana Amirkhanashvili and Davit Gigilashvili (2020). "Color Naming and Communication of Color Appearance: Is it Different for Native Georgian Speakers?" In: *10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*. Vol. 2688. CEUR Workshop Proceedings, pp. 1–15

Article K is based on the same experiment as ***Articles A*** and ***B***, providing analysis of appearance-based clustering and initiating the discussion on potential appearance ordering systems. It is a preliminary work and the observations collected from ***Article K*** have been used for generating research hypotheses and strengthening the conclusions of ***Article B***. However, it is not included as a part of the thesis as it neither tests any particular hypothesis, nor provides a comprehensive report of qualitative observations. The content of ***Article L*** is to a large extent covered in ***Article E***. ***Article M*** tested the hypothesis proposed in ***Article B*** that opacity does not imply the complete absence of transmission. However, the specific type of visual stimuli (textiles) and their context put the work out of the scope of this thesis. ***Article N*** revisits ***Articles A*** and ***G*** and puts them in context of the general problem of material appearance acquisition and reproduction. ***Article O*** has explored communication of appearance - namely, how native Georgian speakers communicate color appearance in comparison with English speakers.

1.7 Ethical Considerations

Conducting psychophysical and social experiments imply collection of personal data, which must be processed in an ethical and responsible manner. The study was conducted with full adherence to research ethics, as well as national and international legal requirements. Participation was voluntary and all participants provided a priori written consent. Demographic information (age, gender, professional background etc.) has been collected and treated anonymously and has not been used for any purpose other than scientific research. The work reported in ***Articles A*** and ***B*** implied collection of sensitive personal information (videotapes of face and voice). Therefore, the study was reported to and approved by the NSD - Norwegian Centre for Research Data (approved project number 59754). The data is to be fully anonymized as soon as ***Article B*** clears the peer review.

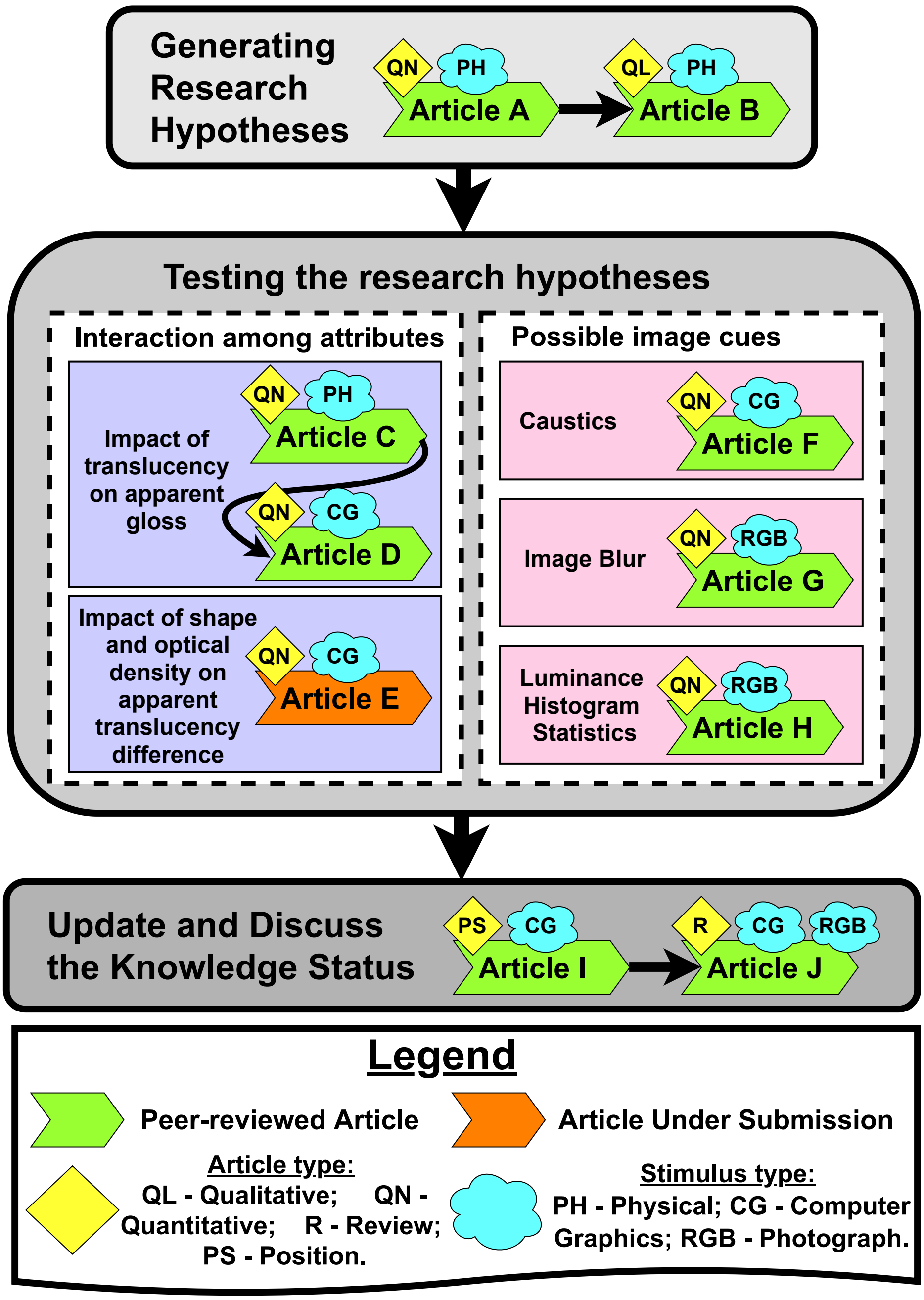

Figure 1.4: The figure explains how the articles are related to each other and where they fit in the narrative of the thesis.

1.8 Thesis Organization

The thesis consists of two parts. Part I consists of the umbrella chapters with the general overview of the work carried out, while the 10 articles mentioned above are appended in Part II. The Introduction chapter covers the motivation of the work, research objectives to be reached, research questions to be answered, and methodologies applied to answer these questions.

The background chapter provides definitions of appearance and its attributes, a brief discussion of qualitative appearance models which are based on philosophical rationale rather than experimental data. As we primarily focus on translucency and gloss, the background chapter also summarizes the optical aspects of translucency and gloss, followed by the state-of-the-art in translucency and gloss perception research.

The third chapter is the summary of the contributions, where the major takeaways from each of the ten articles are summed up. In Chapter 4 we discuss the results, answer the research questions raised in Section 1.3 and analyze how the findings could refine and strengthen the qualitative model proposed in ***Article B***. In the same chapter, we also analyze the limitations of the work and the shortcomings of the articles that have been revealed in the course of the doctoral project. In Chapter 5, we draw conclusions, which is followed by the outline of the future work and the overview of the short and long term perspectives in Chapter 6.

Chapter 2

Background

2.1 Definition of Appearance and its Attributes

2.1.1 Appearance and Total Appearance

According to the ASTM - Standard Terminology of Appearance (ASTM E284-17 (2017)), appearance of an object is "*the collected visual aspects of an object or a scene*", while perceived appearance is defined as "*the visual perception of an object, including size, shape, color, texture, gloss, transparency, opacity, etc., separately or integrated.*" Appearance is a complex phenomenon that is far from being comprehensively understood. Considering its complex nature, appearance is usually broken down into various attributes which entail just particular dimensions of appearance. The CIE[1] defines color, gloss, translucency and texture as four major appearance attributes (Eugène (2008) and CIE (2006)). Pointer (2003) argues that while appearance might imply description of color information only, *total appearance* requires "*a description of the shape, size, texture, gloss and **any other apparent quality***". Appearance has long been a point of scholarly interest, Hunter and Harold (1987) providing the first significant summary of appearance measurement techniques extending Hunter's momentous contributions to understanding different appearance attributes (Hunter (1937)). Although the title "*The Measurement of Appearance*" implies some extent of total appearance measurement, Hunter and Harold primarily focus on individual attributes, with color being the major focus of the textbook. Discussion of total appearance is based on a very constrained qualitative analysis. According to the authors, the objective of appearance measurement is "*to obtain numbers that are representative of the way objects and materials look.*" (Hunter and Harold (1987)) However, they consider that comprehensive analyses of the total appearance is impossible and impractical and argue that "*measurements of specific attributes of appearance can be exceedingly useful and economically important*". This work is not only far from modeling total appearance, but also provides little guidance on the correlation between metro-

[1]Commission internationale de l'éclairage, The International Commission on Illumination - an international organization dealing with color and illumination-related aspects.

logical and perceptual aspects of it.

It is very unlikely that the four attributes of appearance are independent. We have observed that appearance attributes impact each other and the same has been previously proposed by Eugène (2008) as well. There has been an extensive amount of work on appearance in computer graphics, vision, and metrology, the vast majority of them focusing on very narrow specific cases and providing quantitative analysis of particular appearance attributes (Hunter (1937), Motoyoshi (2010), Motoyoshi et al. (2007), Nicodemus (1965), Nishida and Shinya (1998), Xiao et al. (2014), Chowdhury et al. (2017), and Fleming and Bülthoff (2005)), and to the best of our knowledge, there is no comprehensive model and terminology standard for total appearance. However, there have been some attempts to debunk the concept of *total appearance*.

Some aspects of total appearance have been discussed by Hutchings (1995a). His work is "*an attempt to emphasize the continuity of science and art, helping practitioners of these traditionally disparate disciplines work together to achieve a greater understanding and control of the visual images we create and manage in our crowded world.*" He thinks that appearance communication "*can be based on a quantitative understanding of the basic perceptions of form, colour, translucency, gloss, and movement.*" He describes a structure of the factors affecting total appearance (Hutchings (1995a)):

- Appearance Images (e.g. gestalt principles, recognition, emotional and sensory responses)
- Immediate environment factors (e.g. geographical, social, medical)
- Inherited and learned responses (e.g. culture, memory, fashion)
- Receptor mechanisms (color vision, aging effects, adaptation, other senses)
- Design (e.g. aesthetics of paintings, performing arts)
- Object's properties (e.g. optical properties, like spectral reflectance; shape and size; movement and temporal aspects)
- Light source properties (e.g. illumination spectrum and direction)

Hutchings (1999) takes the total appearance concept up to the level of a scene understanding and defines it as follows: "*total appearance combines a description of the appearance of each element of a scene... with a personal interpretation of the total scene in term of its recognition and expectation.*" However, Eugène (2008) also highlights that CIE recommends the following definition: "*the total appearance points out the visual aspects of objects and scenes*". He considers appearance measurement challenging, because it involves subjective judgment and argues that "*a goal of making measurements that ensures appropriate quality control in the manufacturing process is probably achievable, but the measurement process will be multidimensional, product specific and probably application specific*".

Choudhury (2014) has also reviewed total appearance as a concept and described a four-step flow of total appearance from molecular composition of an object to the high level cognitive interpretation of appearance by an observer.

Despite these qualitative attempts to put total appearance perception into some system, all above-mentioned works are theoretical reasoning without being based on particular experimental observations and the behavior of humans.

Translucency and gloss are appearance attributes that play a significant role in total appearance. As they remain relatively understudied unlike color, we decided to investigate the perception of these two attributes. These terms can have different meanings to different people and in different industries (Pointer (2003)). Thus, in the two following subsections we present and discuss the definitions of the terms *translucency* and *gloss*, which should be used for interpreting this work. Afterwards, in the subsequent sections we provide a brief state-of-the-art summary on translucency and gloss perception, respectively.

2.1.2 Definition of Translucency

Translucency appearance is a result of stimuli emitted by an object possessing some degree of subsurface light transport. Translucency relates to spatial variation of color, which takes place "*due to the relationship between the light transmitted, the light reflected, and the light scattered by the body of the object*" (Pointer (2003)).

According to Eugène (2008), "*translucency occurs between the extremes of complete transparency and complete opacity... If it is possible to see only a "blurred" image through the material (due to some diffusion effect), then it has a certain degree of transparency and we can speak about translucency*". Gerbino et al. (1990) make a more clear distinction between transparency and translucency, postulating that "*transparent substances, unlike translucent ones, transmit light without diffusing it.*" ASTM - Standard Terminology of Appearance (ASTM E284-17 (2017)) defines translucency as "*the property of a specimen by which it transmits light diffusely without permitting a clear view of objects beyond the specimen and not in contact with it.*". While technical definitions usually connote subsurface scattering and resulting blur of the see-through image, *translucent* as an adjective in everyday use can be also used to describe transparent and lucid media (*Merriam-Webster Dictionary* (n.d.)). The CIE (2006) highlights that "*translucency is a subjective term that relates to a scale of values going from total opacity to total transparency.*" We have observed a high degree of subjectivity in the interpretation of the term (***Articles A, B*** and ***E***), and discussed potential challenges related to this in ***Article I***.

2.1.3 Definition of Gloss

Gloss is usually associated with surface shininess and is perceived separately from color (Pointer (2003)); According to CIE, gloss is "*the mode of appearance by which reflected highlights of objects are perceived as superimposed on the surface due to the directionally selective properties of that surface*" (CIE (1987) cited in Eugène (2008)) and "*gloss perception is particularly depending on the way that light is reflected from the surface of the object at and near the specular direction.*" (Eugène (2008)) ASTM Standard Terminology of Appearance (ASTM E284-17 (2017)) defines gloss as "*angular selectivity of reflectance, involving surface-reflected light, responsible for the degree to which reflected highlights or images of objects may be seen* *as superimposed on a surface.*" In his classic work, Hunter (1937) postulated six different types of gloss:

1. **Specular gloss** - "brilliance of specularly reflected light, shininess"; (Figure 2.1(a))
2. **Sheen** - "shininess at grazing angles"; (Figure 2.1(b))
3. **Contrast gloss** - "contrast between specularly reflecting areas and other areas"; (Figure 2.1(c))
4. **Absence-of-bloom gloss** - "absence of smear or excess semi-specular reflection adjacent to reflected highlights and images"; (Figure 2.1(d))
5. **Distinctness-of-reflected-image gloss** - "distinctness and sharpness of reflected images"; (Figure 2.1(e))
6. **Absence-of-surface-texture gloss** - "surface evenness, absence of texture, indicated by difficulty of recognizing presence of surface." (Figure 2.1(e))

He proposed that glossiness might be correlated with surface specular reflectance and concluded that reflectance distribution functions "*offer the only means by which the reflectance properties of surfaces responsible for their glossiness may be completely specified.*" This traditional definition that gloss is surface-specific quality is challenged in ***Articles C*** and ***D***.

2.2 Measurement, Modeling and Simulation of Appearance

When discussing the measurement of appearance, it is important to make a distinction between *soft metrology* and *hard metrology*. Soft metrology implies using human response to determine an objective property of the target (Pointer (2003)). In order to study the correlation between physical properties and perception, proper generation of visual stimuli based on these properties is of the utmost importance. The physical accuracy of the rendering in computer graphics is constrained by the accuracy of the input physical material properties, dubbed as "*the input problem*" by Rushmeier (1995). This makes accurate instrumental measurement of these optical properties (hard metrology) important. The most comprehensive and up-to-date survey regarding the acquisition of the optical properties of translucent materials is done by Frisvad et al. (2020).

A pivotal contribution to modeling light and matter interaction has been made by Nicodemus et al. (1977) who proposed bidirectional distribution functions characterizing macro-level interaction between light and materials, and that come in form of BSDF (Bidirectional Scattering Distribution Function) and BSSRDF (Bidirectional Subsurface Scattering Distribution Function). The fundamental difference between the two is that the BSDF is a local approximation of BSSRDF, which assumes that incidence and emergence points are the same, while BSSRDF considers light globally, i.e. light can be incident at one point and emerge from another point. BSDF is a combination of BRDF (Bidirectional Reflectance Distribution Function) and BTDF (Bidirectional Transmittance Distribution Func-

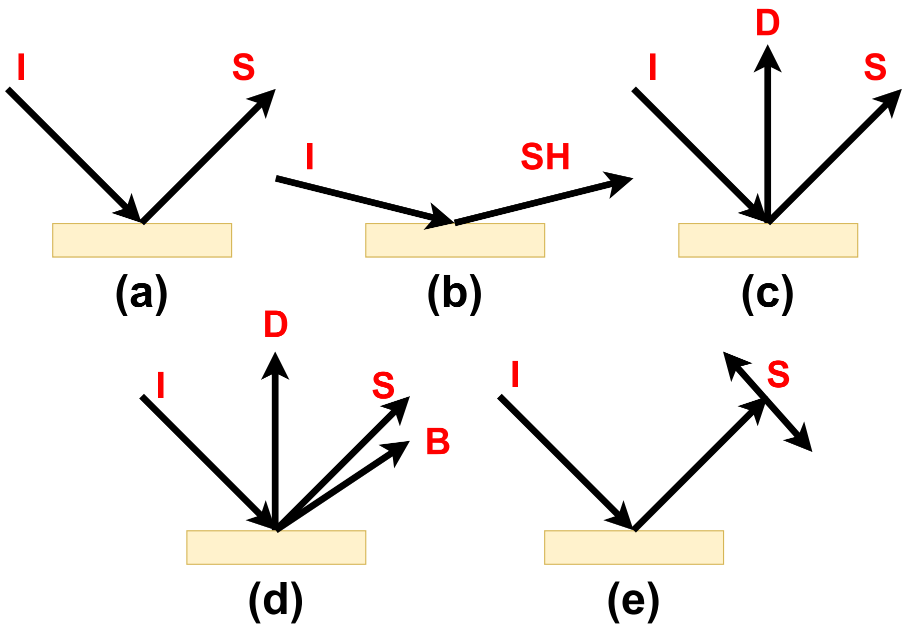

Figure 2.1: Hunter identified six types of gloss. (a) **specular gloss** - shininess due to the mirror reflection (i.e. incident (I) and reflected (S) rays form identical angles with the surface normal); (b) **sheen** (SH) - shininess on different gazing angles (other than specular); (c) **contrast gloss** - the contrast between specular (S) and other areas (D); (d) **absence-of-bloom gloss** - absence of haze or smear (B) in the areas adjacent to specular highlights (S); (e) **distinctness-of-reflected-image gloss** - distinctness and sharpness of reflected image; **absence-of-surface-texture gloss** - inability to detect surface irregularities in the reflected image (surface appears perfectly smooth).

tion). BRDF characterizes the light that is reflected at the point of incidence, i.e. re-emerges towards the same hemisphere it has arrived from, while BTDF characterizes the light that re-emerges on the opposite side. BSDF is usually enough to approximate the light and matter interaction when subsurface scattering is negligible. However, unlike BSSRDF, it cannot account for scattering inside the volume. BSSRDF is eight-dimensional (four spatial and four angular) and it provides the relation between incident radiant flux at a given point x_i from direction $\overrightarrow{\omega_i}$ and outgoing radiance at another point x_j towards direction $\overrightarrow{\omega_j}$. A simplified representation of these functions can be found in Figure 2.2.

Instrumental measurement of BSDF is conceptually more straightforward than that of BSSRDF. Frisvad et al. (2020) discuss goniometric techniques as per ASTM Standard (ASTM E2387-05 (2011)). However, image-based techniques have also been demonstrated (summarized in Dorsey et al. (2010)). The principle in goniometric measurement is the following (ASTM E2387-05 (2011)): a sample object is illuminated from a given direction, while the detector moves and measures how emerging light intensity varies from angle to angle. Afterwards, the illumination

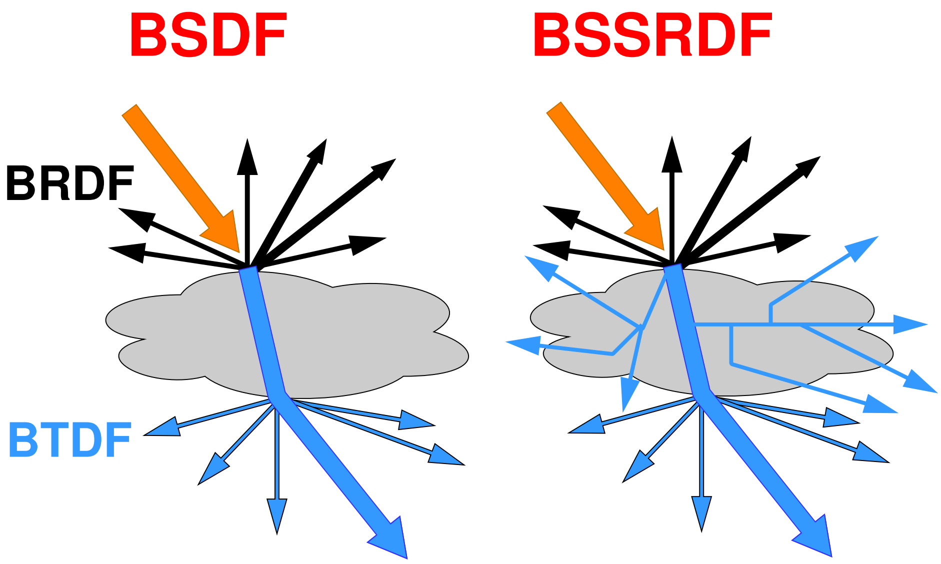

Figure 2.2: Representation of BSDF and BSSRDF. Orange arrow corresponds to incident light, black arrows signify surface scattering, while blue arrows correspond to subsurface light transport. Surface scattering is characterized by BRDF, while BTDF describes transmission, when scattering inside the medium is negligible. The BRDF and BTDF constitute BSDF which is an approximation of more complex BSSRDF. In addition to light and matter interaction characterized by BSDF, BSSRDF also accounts for multiple scattering events taking place inside the material. In BSSRDF, light incident at one point of a surface can emerge from a different point on any side of the object. If the penetration depth is negligibly small due to high absorption and scattering, light that re-emerges back from non-specular areas is in some scenarios approximated as "*diffuse reflectance*". BRDF is usually thought to be descriptive of glossiness. However, we challenge this opinion in ***Articles C*** and ***D***

.

angle is changed by moving either the light source or the object. BRDF is measured in reflection setup (detector and illuminant are in the same hemisphere), while BTDF is measured in transmission setup (detector and illuminant are in different hemispheres). The process is sketched in Figure 2.3. A detailed review of the techniques and instruments for the BRDF acquisition can be found in the work by Leloup et al. (2008).

On the other hand, the high dimensional nature of BSSRDF makes it virtually infeasible to apply the same principle to it. Therefore, according to Frisvad et al. (2020), neither a standardized sampling of directions, nor respective equipment exists. BSSRDFs are usually measured using camera-based techniques, as proposed by Jensen et al. (2001) or such as proposed by Gkioulekas et al. (2013). Piadyk et al. (2020) proposed a light field imaging system for BSSRDF acquisition and built a low-cost prototype setup.

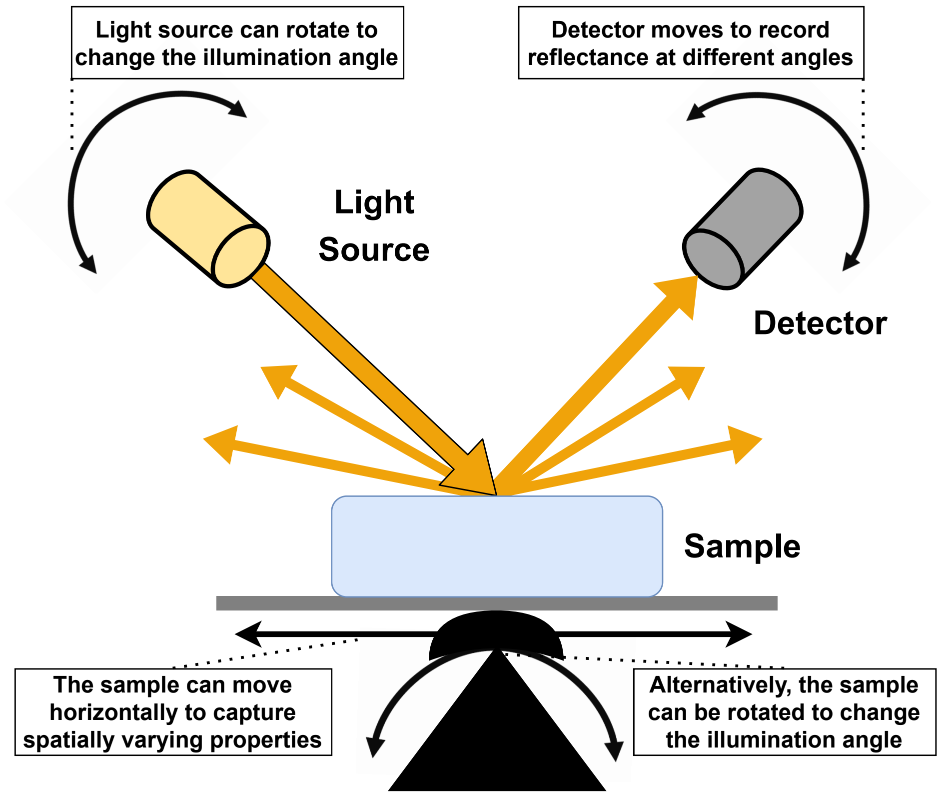

Figure 2.3: A schematic representation of goniometric measurement of material properties in reflectance geometry. For a fixed illumination geometry, the detector moves and quantifies reflected energy at different angles. Afterwards, illumination geometry is changed by rotation of either a light source, or a sample. While this method measures material property at a given point, the sample can additionally be displaced horizontally as well, in order to capture spatially varying properties (properties across different points on the surface).

It is worth noting that as BSSRDF includes a spatial component, it is a function of object's shape and geometry. Therefore, in addition to intrinsic optical properties, the acquisition of object's geometry is also of vital importance. However, capturing the shape of translucent materials to date remains a challenging task for 3D scanners, and various invasive techniques have been proposed as workarounds, such as covering with a layer of diffuse opaque dust in order to "*turn off subsurface scattering*". (Goesele et al. (2004))

The seminal work by Jensen et al. (2001) pioneered using BSSRDF in computer graphics, which remarkably advanced translucency rendering as well as translucency perception studies. The authors simplified the problem by assuming that when the light propagates through a homogeneous translucent medium and scatters multiple times, diffusion theory can be applied. Instead of addressing all scattering events individually, they use diffusion equation and approximate the subsurface scattering with a single scattering term. In other words, when a photon gets scattered many times, its directionality becomes somewhat random (however, this might not work well for thin objects as observed by Gkioulekas et al. (2013) and Xiao et al. (2014)). The model is called the classical dipole model (Jakob (2010)). The parameters of the BSSRDF are the index of refraction, scattering and absorption coefficients, and the scattering phase function which defines the directionality of the scattered light. Nowadays, a broad range of techniques exists to avoid dipole-type of approximations (Jensen et al. (2001)) and to model and simulate light and matter interaction in an accurate manner, which by Frisvad et al. (2020) is divided into roughly two categories: radiometric models and field models. The latter is applied when a rigorous description of the electromagnetic field and e.g. solving Maxwell's equations are needed. This could be the case when replication of the wave optics phenomena (e.g. interference and diffraction) is desired. On the other hand, the problem can be approached in the radiometric domain and light and matter interaction can be modeled as a variation of radiant energy due to absorption and scattering phenomena. The process entails modeling coefficients of absorption and scattering, as well as scattering phase function, and solving the *radiative transfer equation* (Chandrasekhar (1960)). One of the most popular methods for solving the *radiative transfer equation* is the Monte Carlo method. The Monte Carlo method is a probabilistic approach. For instance, Monte Carlo ray tracing entails following light rays through the scene. Whether the ray is absorbed or scattered inside the medium, or whether it is reflected or refracted at the boundary, is decided stochastically. Monte Carlo methods have been broadly used in appearance perception research to generate translucent visual stimuli for psychophysical experiments (e.g. Urban et al. (2019), Gigilashvili et al. (2019), Xiao et al. (2014), and Gkioulekas et al. (2015)).

While the above-mentioned techniques attempt to acquire and model physical material properties of translucent objects, no technique has been proposed to date for measuring overall perceptual qualities instrumentally. There are multiple application-specific instruments on the market for measuring distinct visual attributes related to transmission-related properties and appearance (*BYK Gardner GmbH. Haze-gard Transparency Transmission Haze Meter* (n.d.)). Two most common attributes studied in relation to translucency are **clarity** - "*clarity, defined in terms of the ability to perceive the fine detail of images through the material*", and **haze** - "*defined as a property of the material whereby objects viewed through it appear to be reduced in contrast*" (Pointer (2003)). Haze is usually associated with wide angle scattering (when the angle between incident illumination and transmitted light is more than 2.5 degrees, according to the ASTM standard (ASTM D 1003 (2003)) of light that causes blur and loss of contrast of the see-through image, while clarity usually results from narrow angle (less than 2.5 degrees) scattering. However, it is important to highlight that no clear link between translucency as an appearance attribute, on the one hand, and clarity and haze, on the other hand, has been established. Pointer (2003) argues that "*the concept of translucency can perhaps be regarded as a descriptor of the combined effects defined above as clarity* *and haze. This implies that it is a more general term and, perhaps, should be limited to use as a subjective term, keeping clarity and haze as descriptors of objective, or measurable, correlates.*" This subjectivity can raise the question, whether translucency is the right attribute to study and to be measured at all. However, we need to highlight two factors: first of all, neither clarity, nor haze alone can fully characterize the complexity of subsurface light transport properties, and it is translucency that encapsulates the effects of both combined; secondly, the definitions of clarity and haze to some extent imply the visibility of the background image through the object. However, oftentimes it is not possible to see the background through the objects and materials, and the luminance variation on the object's body is the sole indicator of subsurface light transport. Therefore, the appearance characteristics of the objects made of, for example, marble or wax, are better conveyed by translucency.

2.3 The Gap between Physics and Perception

Pointer (2003) asserts that appearance consists of three aspects: *physical* - spatial and spectral distribution of the light emerging from an object, which depends on its optical properties; *physiological* - the stimulation of the HVS by this light, the sensory response; and *psychological* - the ability to interpret the sensed stimuli "*thanks to long training*". Despite the advance in the acquisition (Frisvad et al. (2020)) and modeling (Dorsey et al. (2010)) of the optical properties of materials, it remains largely unknown how these objective physical properties from the scene relate to what people perceive. Advance in computer graphics makes virtual prototyping and creating digital twins possible. However, the knowledge gap between physical properties and perception limits our ability to generate desired visual effects from scratch, and to predict and replicate appearance across different objects, scales and observation conditions. This motivates the attempts of *soft metrology*, particularly relying on psychophysics - "*the study of the functions relating the physical measurements of stimuli and the sensations and perceptions the stimuli evoke.*" (ASTM E284-17 (2017)). This explains the ever increasing interest in gloss and translucency perception research in modern vision science and the broad range of industries, such as 3D printing.

2.4 Translucency Perception

We want to make it clear that in this thesis we are addressing *translucency* and not *transparency*. As already mentioned above, it is usually accepted that *"transparent substances, unlike translucent ones, transmit light without diffusing it."* (Gerbino et al. (1990)) Unlike *translucency* perception, mechanisms of *transparency* perception are relatively well-explored and understood. However, *transparency* and *translucency* are not mutually exclusive and a given visual stimulus might to some extent evoke perception of the both attributes.

Up until recently, studies on light transmittance have been limited to transparency perception, modeling an object as a thin filter. The classical studies in this direction have been done by Metelli who proposed the episcotister model (Metelli (1974)) which models transparency as a linear color fusion between opaque part of the rotating circle and background seen through its cut-out sector (refer to Figure 2.4). Further studies paid attention that the spatial arrangement of the luminance intensities at the filter-background boundaries, called X-junctions (Beck and Ivry (1988) and Gerbino et al. (1990)), could assist the HVS in inferring material transmittance properties (refer to Figure 2.5). While Metelli's model implies an *additive color mixture* to model transparency, filter models have also been proposed as an alternative approach (Beck et al. (1984), Faul and Ekroll (2002), and Faul and Ekroll (2011)). In these models the transparent overlay is presented as a filter that reflects part of the light at its surface (additive component), while the rest gets refracted and continues propagation through the filter, where it can get absorbed depending on the filter's thickness and absorbance (*subtractive color mixture*). Faul and Ekroll (2011) demonstrated that under the diffuse illumination the surface reflection of the filter might evoke the perception of translucency even without subsurface scattering, as the background contrast is decreased and it serves as a cue to translucency. Singh and Anderson (2002) studied thin transparent filters that scatter light that propagates through them. They proposed that blur and apparent contrast of the background image are the cues to translucency. However, if the magnitude of scattering is large enough, the background is not visible through the object and transparency cues, such as X-junctions, are absent. This means that transparency perception models simply cannot explain the perception of translucency in highly scattering media, which gave birth to translucency perception research as an independent direction.

Translucency depends on numerous intrinsic and extrinsic properties of an object and scene. The most extensive survey accounting for subsurface scattering in 3D objects has been carried out by Fleming and Bülthoff (2005). They studied the image cues affecting translucency perception and argue that the human visual system is not capable of *inverting optics*, but rather relies on simple image cues and statistics to judge translucency. They review a broad range of factors affecting the perceived translucency, like specular highlights, color, object's scale, image contrast and illumination direction. It is worth mentioning that their results are limited to a small number of rendered images with objects of simplified geometrical structure. The impact of illumination direction on perceived translucency was studied further by Xiao et al. (2014), who conclude that the perceived degree of translucency strongly depends on the illumination direction and most materials look more translucent when they are back-lit, rather than in case of front-side illumination. They introduce the concept of translucency constancy, i.e. an ability of human-beings *"to estimate translucency in a consistent way across different shapes and lighting conditions"* and make a counter-intuitive finding that the objects with complex geometric shapes demonstrate a higher degree of translucency constancy failure, even though complex objects provide more vi-

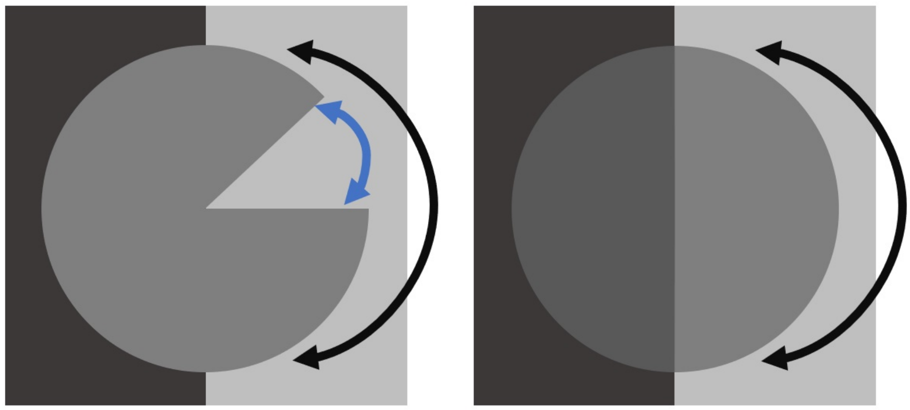

Figure 2.4: The linear algebraic modeling of transparency has been proposed by Metelli (1974). The episcotister is an opaque disc with a sector cut out (left image). When it rotates fast enough, the background and disc colors fuse in a linear fashion and the disc is perceived as a thin transparent film overlaid over the background (right image). The perceived color of the disc depends on the color of the opaque part, as well as the angle of the see-through sector (blue arrow).

sual cues (Fleming and Bülthoff (2005) and Xiao et al. (2014)) about translucency. Gkioulekas et al. (2013) have shown that translucency has at least two perceptual dimensions and they are impacted by the scattering phase function. The sharpness of the surface details is another factor that has been demonstrated to be impacting perceived translucency (Xiao et al. (2020) and Sawayama et al. (2019)). Motoyoshi (2010) has approached the question from the perspective of image structure. The author has shown that luminance contrasts within distinct spatial frequency bands of non-specular object regions carry relevant information on translucency appearance, and low luminance contrast in these regions is usually an indicator of translucency. Nagai et al. (2013) also identified that particular image regions are “hot spots” for translucency perception and the HVS relies on local luminance statistics in those regions. Interestingly, the authors reported that the informative regions observers have relied on are not universal, and they vary from person to person. Marlow et al. (2017) argue that the lack of co-variance between shape and shading might be the cue the HVS relies on for distinguishing translucent and opaque materials. They have even been able to evoke the illusory perception of translucency with an optically opaque material by manipulation of the diffuse light field which produced the shading non-covariant with surface geometry. This means that the perception of translucency might be inherently interconnected with shape perception. Marlow and Anderson (2021) have recently shown that both translucency and shape depend on the relations among the subsurface scattering, the specular reflections and the self-occluding contours, as all of these three factors are rooted in the same geometrical property

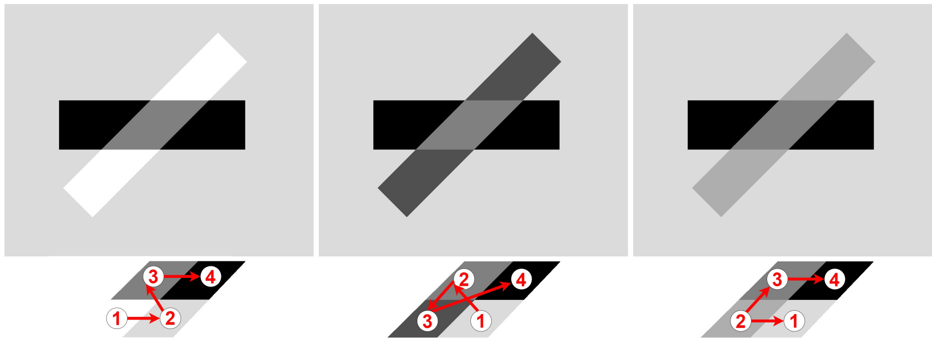

Figure 2.5: X-junctions, i.e. the relation between luminance intensities on the bounding regions, have been shown to be definitive for the perception of transparency. **S-shaped junction** (left image) creates a bi-stable image, where both figures can be perceived as being on top. This can be attributed to the fact that the overlapping polygon intensity is the mixture of the two, while on the background they both look solid opaque white and black and neither of them looks to be fused with the light gray background, which does not provide enough cue to understand which one is on the top and which one is on the bottom. **Crisscross junction** (the middle image) does not produce the impression of transparency, because the overlapping region is lighter than both of the figures, which makes it physically impossible to be the mixture of the two. **S-shaped junction** (right image) the diagonal figure looks overlaid over the horizontal black figure. In contrast with the left image (S-shaped junction), the rest of the diagonal figure is not pure white, it looks fused with the light gray background, which makes our visual system deduce that this is the transparent one and not the horizontal solid black object, which seemingly does not have a contribution from the background.

- the 3D surface curvature. This means that the HVS might be assessing the shape and translucency together from the combination of the same image cues. Finally, Chadwick et al. (2019) have recently demonstrated anatomical independence of translucency perception from that of color and texture. They showed that damage in cortical areas responsible for color and texture processing does not compromise the ability to perceive translucency. Despite those attempts, the exact mechanisms of translucency perception remain largely unidentified. A comprehensive review of the knowledge status in translucency perception research is given in ***Article J***.

2.5 Gloss Perception

The knowledge about gloss perception mechanisms also remains limited. Various image cues and statistics, such as skewness of luminance histogram (Motoyoshi et al. (2007) and Landy (2007)), contrast (Pellacini et al. (2000), Thomas et al. (2017), Marlow et al. (2012), and Marlow and Anderson (2013)), sharpness (Pellacini et al. (2000), Marlow et al. (2012), and Marlow and Anderson (2013)) and coverage area (Beck and Prazdny (1981), Marlow et al. (2012), Marlow and Anderson (2013), and Kerrigan and Adams (2013)) of the highlights have been proposed as potential glossiness cues. However, similar statistics might be found in the images of some non-glossy materials as well, meaning that those findings are subject to multiple photo-geometric constraints (Anderson and Kim (2009), Kim et al. (2011), and Marlow et al. (2011)). Pellacini et al. (2000) have identified two perceptual dimensions of gloss that are similar to contrast and distinctness-of-image. They concluded that "*darker objects look glossier than lighter ones*". Toscani et al. (2020) proposed that surface reflection has at least three perceptual dimensions: lightness, gloss, and metallicity.

Gloss perception is a complex psychophysical process that relies on the analysis and interpretation of several image cues and involves some degree of subjectivity. Wendt et al. (2010) have demonstrated that color, motion and disparity cues are used in the process, both separately and in combination. However, different observers prioritize different cues. Leloup et al. (2012) studied gloss perception using the physical objects and identified a very interesting dichotomy in the observers' approaches. They found that some observers prioritize the distinctness of the reflected image as a cue to glossiness, while others principally rely on the luminance contrast between the specular and diffuse areas.

Leloup et al. (2010) studied physical samples and reported that the perceived contrast is a better correlate of the perceived gloss than the instrumentally measured specular gloss, while the entire process is strongly impacted by the complexity and geometry of the illumination. The latter observation has been consistent with the claims by Fleming et al. (2003), who proposed that gloss depends on the illumination, and the matching accuracy of the surface reflectance properties, as well as the magnitude of the perceived gloss, is higher under a realistic complex illumination. Later, Leloup et al. (2011) introduced a perceptual metric that incorporates both surface and illumination characteristics, and predicts perceived glossiness based on the luminance measurements in the specular and non-specular areas. On the other hand, Obein et al. (2004) found that observers are able to compensate for the changes in the stimuli induced by the varying illumination geometry and hence, the HVS to some extent demonstrates gloss constancy, similarly to the color constancy. Recently, Faul (2019) also reported the high gloss constancy across the illumination conditions, and argued that the strong dependence of gloss on the illumination in the previous studies can be ascribed to the lack of the Fresnel effects in the visual stimuli and the simplistic shapes used in the experiments (spheres (Fleming et al. (2003)) and flat patches (Leloup et al. (2010) and Leloup et al. (2011)) had been used by the other authors). Gloss constancy is likely to be related to the ability of the HVS to identify and segment the specular and diffuse reflection components in the proximal stimulus. One of the instruments the HVS proposedly relies on for segmentation is the binocular vision – as the binocular disparity and the binocular depth cues facilitate isolation of the object body (diffuse component) from the specular highlights (Wendt et al. (2008)) and the mirror-reflection image of the immersing environment (Obein et al. (2004)). Gloss constancy is, however, limited and in addition to illumination (Fleming et al. (2003) and Olkkonen and Brainard (2011)), the perceived magnitude of glossiness can be also affected by object's shape (Vangorp et al. (2007), Olkkonen and Brainard (2011), and Marlow et al. (2012)), color (Nishida et al. (2008) and Wendt et al. (2010)) and motion (Wendt et al. (2010), Sakano and Ando (2010), and Doerschner et al. (2011)). Cheeseman et al. (2021) have recently studied sensitivity to the changes in the specular reflectance and found that the visual sensitivity to the gloss differences is lower when the magnitude of the specular reflection is high.

Finally, Ged et al. (2010) have noted that gloss contributes to material identification and discrimination. Additionally, they highlighted the importance of observing the materials from multiple angles, as the surface reflectance and the resulting intensities in the proximal stimulus are angle-dependent. A comprehensive review on gloss perception can be found in works by Chadwick and Kentridge (2015) and Leloup et al. (2014).

Chapter 3

Summary of Contributions

Ten manuscripts included as a part of the thesis are summarized in this chapter. We briefly summarize the objectives, methods and major takeaways of each work. For further details, refer to the respective manuscripts in Part II.

3.1 Article A: Behavioral investigation of visual appearance assessment

Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2018). "Behavioral investigation of visual appearance assessment." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 294–299

3.1.1 Objectives

While the psychophysical experiments studying material appearance are usually conducted in fixed and strictly controlled conditions, the process is far from what we experience in our daily lives and these kind of experiments might not reveal the actual behavioral patterns humans usually apply for assessing material appearance. Therefore, this study was conducted in uncontrolled conditions, permitting unrestricted interaction with the objects, as it is in a daily routine. Although this does not permit modeling the correlation between physics and perception, the objective of this study has been identification of interesting trends, generation of research hypotheses based on them and outlining the directions for future research.

3.1.2 Methods

We conducted series of social experiments in uncontrolled illumination and observation conditions. The observers were asked to describe the physical objects

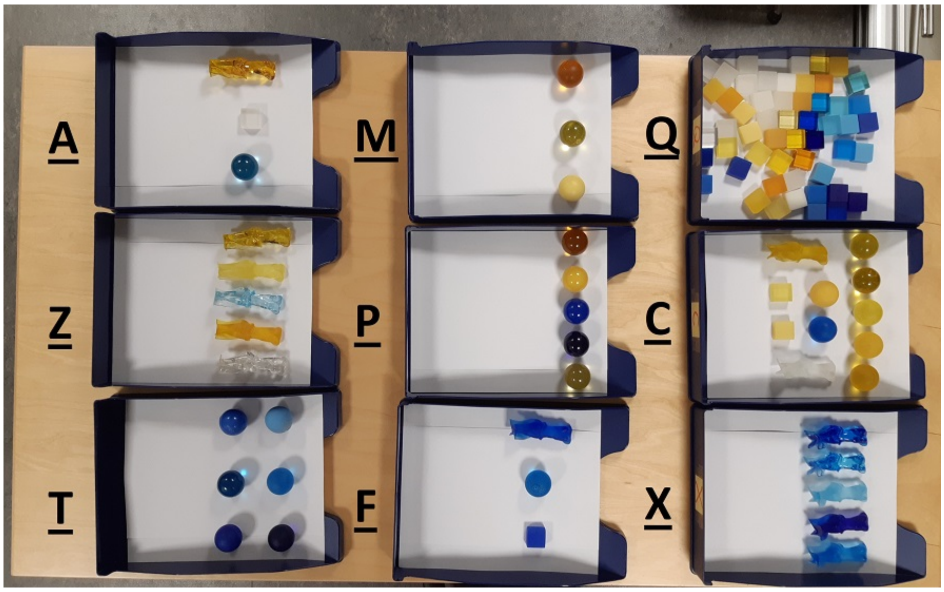

Figure 3.1: The objects used for the experiment. They have been distributed into 9 boxes and used for 11 different visual tasks. The letters are randomly assigned for reference purposes only.

and to perform eleven simple visual tasks that came in five different types: clustering objects by their appearance similarity, arranging objects in a space in any way observers consider "natural", ranking objects by glossiness, ranking objects by translucency and clustering objects into opaque and non-opaque categories. We used objects from the *Plastique* artwork collection (Thomas et al. (2018)). The objects are illustrated in Figure 3.1. We videotaped the experiment from two viewpoints to analyze the entire process subsequently. A sample frame from such video is shown in Figure 3.2. Frequency analysis of the task results was performed in order to identify interesting trends and hypotheses. Besides, the results of the frequency analysis have been used as an input for the qualitative analysis reported in ***Article B***.

3.1.3 Results

The frequency analysis of the task results has revealed several interesting trends that have been used to formulate research hypotheses and inspired future work. We have made several observations:

- Different tasks produced contradictory results on gloss perception. While on one occasion, an equal magnitude of apparent gloss was evoked with equal surface coarseness, in other cases, lightness and translucency also contributed to apparent gloss. Contradictory results have been obtained in the latter case as well: while some subjects considered lighter and translucent

Figure 3.2: A sample frame from the dual-view videotape of the experimental process.

objects glossier, because more light emerged from them, others opted for darker and opaque ones, which manifest a larger tonal range and contrast.

- Object's shape can have a significant impact on the magnitude of perceived translucency of a material, and the presence of thin parts can outweigh intrinsic material properties, such as density of the scattering particles.
- Opacity does not imply a complete absence of transmission. Classification of a material as *opaque* varies across illumination conditions. Besides, caustics can facilitate distinction between opaque and non-opaque materials.
- Definition of translucency has been found to be a challenging task. Interpretation of the instruction "*rank the objects by how light is going through them*" varied among observers and led to contrasting results.

The results of this article have inspired the rest of the work and have been used as an input for other articles in the following way:

- Whether translucency impacts perceived gloss was studied in ***Articles C*** and ***D***.
- Whether the presence of thin parts contributes to the detection of translucency, and particularly, translucency differences, was examined in ***Article E***.
- The role of caustics in translucency perception was investigated in ***Article F***. Besides, we also studied whether high illuminance backlight impacts discriminating opaque and non-opaque materials. Refer to ***supporting Article M*** for further details.
- Challenges related to the definition and interpretation of *translucency* as a term, made us question the current conceptual understanding of translucency and inspired our ***position Article I***.

3.2 Article B: On the appearance of objects and materials: Qualitative analysis of experimental observations

Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). "On the appearance of objects and materials: Qualitative analysis of experimental observations." In: *Accepted for publication in the Journal of the International Colour Association (JAIC)*, 33 pages

3.2.1 Objectives

The objective of the work is bi-fold: firstly, to build a qualitative model of material appearance assessment which is rooted in the experimental data; and secondly, to generate relevant research hypotheses. Validation or falsification of these hypotheses should not only strengthen the proposed model, but also advance general understanding of material appearance.

3.2.2 Methods

The work analyzes the experiment described in ***Article A***. However, instead of the quantitative analysis of the task results, this work analyzes the overall process of material appearance assessment in a qualitative way. The Grounded Theory Analysis (GTA) (Paillé (1994)) - an inductive research method has been used for this purpose. The method is described in Section 1.4.

3.2.3 Results

The resulted qualitative model is illustrated in Figure 3.3. The model consists of two sections: the essential Visual Part - which portrays the process from the introduction of the object to completion of the visual task on it; and auxiliary Decision-making Part which characterizes the factors that could impact a methodology selection for performing tasks. While decisions made on the methodology can affect the result of the task, the entire pipeline remains independent of the task and individual observer.

Object, with its absolute properties, such as size and shape, and conditions of observation, such as illumination geometry and spectral composition, produce the input stimulus for observers' visual system. When an observer is asked to perform a task based on the input visual stimulus, they need a relevant methodology for solving the task. Comparison with a reference has been a fundamental behavioral pattern all methodologies have relied on. The task instructions, social interaction with the experimenter and pre-existent expectations of the observers impact how they interpret the task and how they come up with a particular strategy to solve the given problem. When observers are asked to describe the object and communicate

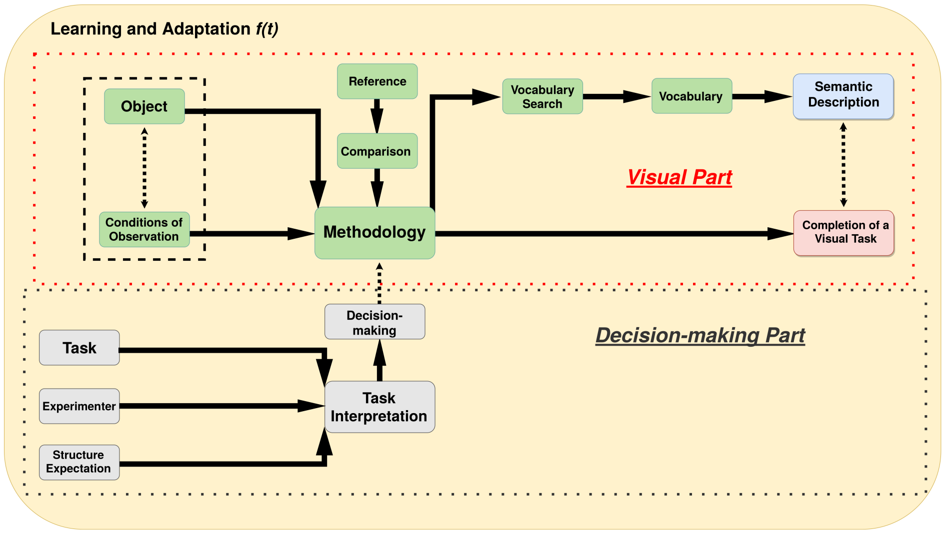

Figure 3.3: Qualitative model of material appearance assessment. The primary Visual Part of the model details the flow of the process from the introduction of an object in particular conditions to the semantic description of its appearance and completion of a visual task using this object. The auxiliary Decision-making part illustrates categories impacting methodology selection in the Visual Part, while Learning and Adaptation impacts the entire process as a function of time.

their appearance, they have to rely on respective vocabulary which itself is a result of an extensive vocabulary search process. Finally, the entire process changes over time due to the acquisition of new skills and information, as well as the change in the physiological state of the observer. An explicit example of how this model is rooted in the data can be found in *Appendix 2* of ***Article B***.

Considering qualitative analysis and quantitative results from ***Article A***, we formulated 20 research hypotheses and discussed their relevance in the light of the state-of-the-art. Below we list and discuss the most significant ones:

- **Translucency impacts perceived glossiness of an object.** This hypothesis has been addressed in ***Articles C*** and ***D***.
- **(a) A given material looks more translucent when an object made of it has thin parts; (b) Shape difference can dramatically impact appearance difference even for identical materials.** These two hypotheses inspired us to propose at the later stage that **thin parts facilitate detection of translucency differences**, which is explored in ***Article E***.
- **Presence of caustics is a cue for translucency assessment and may increase perceived degree of translucency.** This hypothesis is tested in ***Article F***.
- **(a) Multisensory information and interaction level impact the robustness of appearance constancy; (b) Motion facilitates gloss perception; (c) Back-lit is a preferred lighting geometry for translucency assess-**

ment. These three hypotheses made us study appearance in two different contexts: using physical tangible objects permitting interaction and using displayed stimuli with strictly controlled conditions and no interaction.

3.3 Article C: Perceived Glossiness: Beyond Surface Properties

Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (2019). "Perceived Glossiness: Beyond Surface Properties." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 37–42

3.3.1 Objectives

In this work, we attempted to challenge the established opinion that perceived gloss exclusively depends on surface-qualities of an object. As hypothesized in ***Articles A*** and ***B***, translucency impacts perceived glossiness. This is further supported by the notions that, subsurface light transport can modulate the image cues which are supposedly used for gloss perception (Motoyoshi et al. (2007) and Nishida and Shinya (1998)) and the HVS is poor at inverting optics (Motoyoshi et al. (2007) and Fleming and Bülthoff (2005)) and thus, is unlikely to fully separate transmission and reflection components.

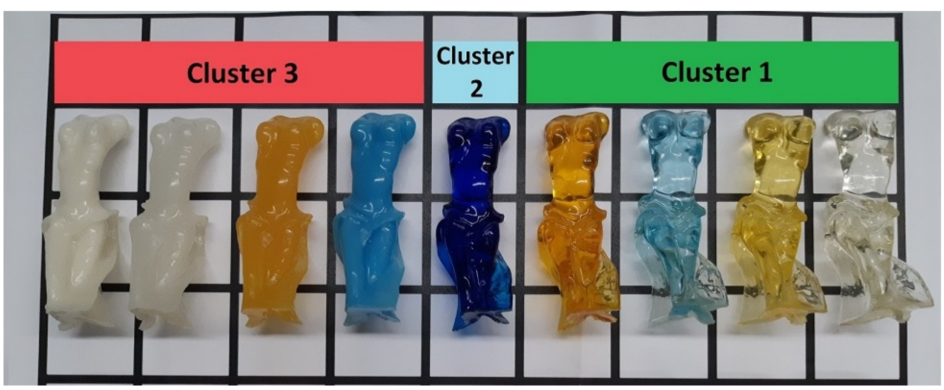

Figure 3.4: The female bust objects used for the experiment. Three clusters of objects emerged from the data: more transparent ones (cluster 1), which have been usually ranked glossiest by most observers; more opaque ones (cluster 3), which have been ranked glossiest by some observers, but were mostly considered least glossy; and the ranking of the dark-blue semi-transparent object (cluster 2) varied considerably. Although the latter was not as shiny as cluster 1 objects, its darker color and hence, higher contrast between specular and non-specular parts, landed it usually higher than cluster 3 objects in the ranking.

3.3.2 Methods

We conducted rank order psychophysical experiments under uncontrolled conditions. 107 observers ranked female bust plastic objects by their glossiness which had identical surface properties but differed in subsurface scattering. The tactile interaction with the objects (illustrated in Figure 3.4) has been unrestricted.

3.3.3 Results

The results varied considerably among observers and their approaches to the task can be categorized into four groups:

1. 10 people (9.35%) tied all objects considering them equally glossy.
2. 84 people (78.50%) considered more transparent objects glossier (cluster 1 objects in Figure 3.4).
3. 8 people (7.48%) considered more opaque objects glossier (cluster 3 objects in Figure 3.4).
4. 5 people (4.67%) used an approach that did not fit in any of the above-mentioned categories.

The results have led to three major observations:

- Identical surface properties do not necessarily yield the identical perception of gloss.
- Gloss perception function, or the interpretation of the concept, varies among individuals.
- Unlike spherical objects used in ***Article A***, fewer observers considered opacity to be positively correlated with gloss. This can be explained by the fact that the complex surface geometry of the female bust objects does not permit observation of mirror-reflection of the environment.

3.4 Article D: The Role of Subsurface Scattering in Glossiness Perception

Davit Gigilashvili, Weiqi Shi, Zeyu Wang, Marius Pedersen, Jon Yngve Hardeberg, and Holly Rushmeier (2021). "The Role of Subsurface Scattering in Glossiness Perception." In: *ACM Transaction on Applied Perception* 18.3, 10:1–10:26

3.4.1 Objectives

The study is inspired by ***Article C*** and aims to test the hypothesis that subsurface scattering impacts gloss perception. Additionally, the study investigates how this impact varies as a function of micro-scale surface roughness and macro-scale shape of the object.

3.4.2 Methods

The problem has been studied in the context of computer graphics applications. The physically-based rendering (Jakob (2010)) has been used to vary surface roughness, extinction coefficient and subsurface scattering albedo systematically, while isotropic scattering phase function has been used and the index of refraction has been fixed to 1.5. Two paired-comparison psychometric scaling experiments have been organized on the Amazon Mechanical Turk. While the initial experiment covered spherical objects only, four additional shapes have been introduced in the second experiment, which varied in surface curvature and thickness. The shapes are illustrated in Figure 3.5.

3.4.3 Results

The analysis of the two experiments made us conclude:

- We have not been able to falsify the null hypothesis that subsurface scattering properties do not contribute to perceived glossiness. There is ample evidence that subsurface scattering can impact apparent gloss. The examples are shown in Figure 3.6.
- The impact made by subsurface scattering differs among levels of micro-scale surface roughness and macro-scale shape of the object.
- For a spherical object the impact of subsurface scattering on gloss is stronger when the surface is smooth; conversely, for complex Stanford Lucy shape, surface roughness increases the role of subsurface scattering in gloss appearance; the impact remained limited for all cylindrical objects.
- For smooth spherical objects, apparent gloss is negatively correlated with albedo, but the correlation is positive for rough spherical objects. For Lucy, apparent gloss is negatively correlated with the extinction coefficient and positively correlated with albedo, regardless of roughness.
- The impact of subsurface scattering is relatively modest in comparison with the impact made by surface scattering. However, we have generated images with different roughness which equal in apparent gloss, because of the differences in subsurface scattering.
- Unlike ***Article C***, the inter-observer consistency has been higher, which makes us think that virtual stimuli demonstrate larger gloss constancy.

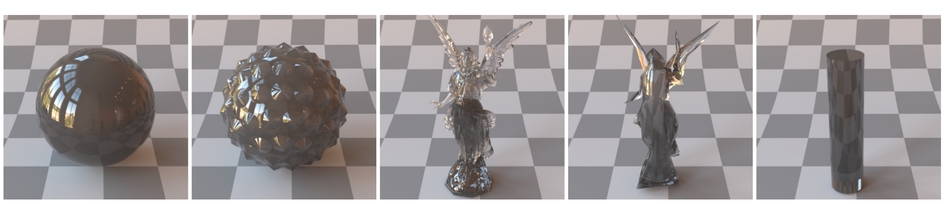

Figure 3.5: Five different shapes have been included in the experiments. Left to right: sphere, spiky sphere, Stanford Lucy, low-resolution Lucy and cylinder.

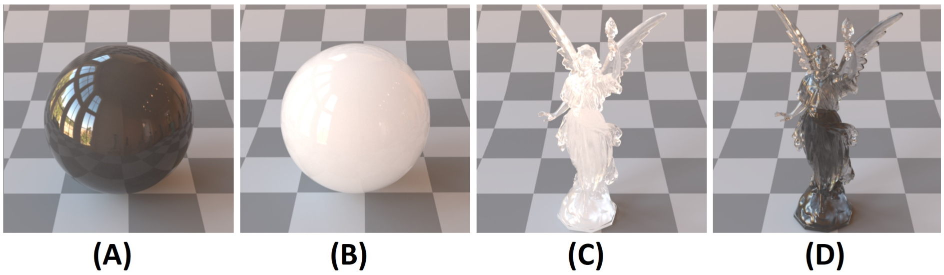

Figure 3.6: Although objects A and B have identical shapes, roughness and spectral reflectance, object A looks glossier, which can be attributed to its lower albedo. Low albedo generates higher contrast and permits observing mirror-like reflections. Contrarily, higher albedo Lucy (image C) has been considered glossier than the Lucy in image D, which only differs from it in subsurface scattering properties. In this case, high albedo generates more highlights. The complex shape of Lucy makes separation of reflectance and transmission components difficult and the highlights are mistaken for specular reflections.

3.5 Article E: The Impact of Optical and Geometrical Thickness on Perceived Translucency Differences

Davit Gigilashvili, Philipp Urban, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). "The Impact of Optical and Geometrical Thickness on Perceived Translucency Differences." In: *Under review in a journal*, 13 pages

3.5.1 Objectives

The primary objective of the study is to test the hypotheses that it is easier to detect changes in translucency when (a) the object has geometrically thin parts; (b) the object is made of an optically thin material. Additionally, the study aimed to produce further hypotheses for future translucency perception research.

3.5.2 Methods

We used a set of virtual materials that varied in absorption and scattering coefficients and were presented in a virtual viewing booth proposed by Urban et al. (2019). We conducted psychophysical experiments with a method of constant stimuli. The observers were shown two pairs of images and they had to select the one with a larger difference in translucency. One of the pairs has always been an anchor Buddha pair with suprathreshold translucency difference, identical to the one used in Urban et al. (2019). The second pair was composed of test images that came in five different shapes. A probit model was fitted to identify the difference in absorption and scattering coefficients necessary for yielding suprathresh-

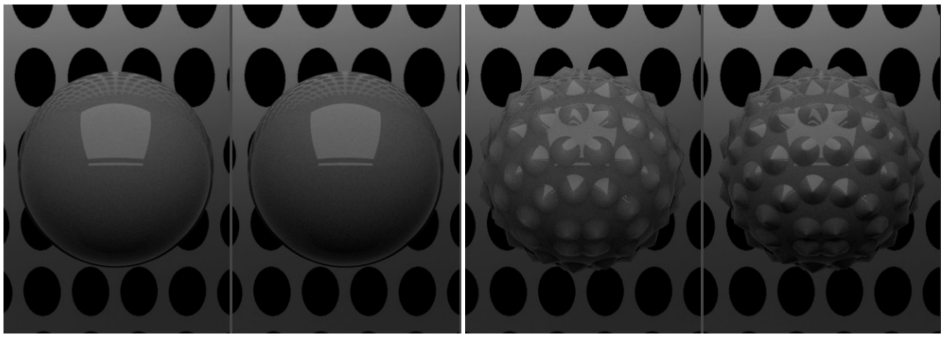

Figure 3.7: In both pairs, absorption and scattering coefficients equal to 77.5 in the left image and 1000 in the right one. Regardless of the considerable distance in absorption-scattering space, spheres look nearly identical. The difference becomes more apparent for bumpy objects, as the bumps produce sharper shadows when the material is more opaque, while the shadows are absent due to light transmission through the bumps when absorption and scattering are lower.

old translucency difference and studied how this varied across different shapes. The shapes were quantified with surface-to-medial-axis histograms. We repeated the experiment, but a transparent anchor-pair was substituted with an anchor pair that did not permit to see the background through the objects. In the second experiment, we studied how the detection of translucency differences varies between see-through and non-see-through materials.

3.5.3 Results

The experiments have produced ample evidence in support of both hypotheses. It is easier to spot suprathreshold translucency differences on spiky objects, which have thin parts than it is for compact spherical objects. This phenomenon is demonstrated in Figure 3.7. Despite this qualitative observation, we have not been able to find a quantitative model that would correlate this impact with an objective shape descriptor, such as a histogram of surface-to-medial-axis distances.

Besides, we also found that the HVS is more sensitive to changes in absorption and scattering when the material is optically thin and it permits seeing the blurred background through it. A larger difference in material properties is needed to spot the difference when no background can be seen through the object and the HVS relies solely on luminance distribution on the object's body. This makes us conclude that translucency and transparency involve interpretation of fundamentally different image cues and these phenomena should be studied separately.

3.6 Article F: Caustics and Translucency Perception

Davit Gigilashvili, Lucas Dubouchet, Marius Pedersen, and Jon Yngve Hardeberg (2020). “Caustics and Translucency Perception.” In: *Material Appearance 2020, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 033:1–033:6

3.6.1 Objectives

As shown in Figure 3.8, caustics might carry valuable information regarding the material, such as color and light transmission properties. We observed in the experiments reported in ***Articles A*** and ***B*** that human observers often use caustics as a cue for translucency assessment. This made us hypothesize that placing an object on a surface that does not permit observation of caustics will impact the magnitude of perceived translucency of a given object. The objective of this work is to test this hypothesis.

3.6.2 Methods

We generated a set of dielectric materials placed in a virtual Cornell box. The materials were presented in five different shapes. Each material was rendered twice - in the original Cornell box and in the Cornell box where the floor was made fully absorbing black (refer to Figure 3.9). Translucency was modulated with surface roughness, expressed as the root mean square (RMS) slope of microfacets (Jakob (2010)). We conducted category judgment psychophysical experiments on a QuickEval platform (Van Ngo et al. (2015)). The task of the observers was to assign a given material to one of the six categories, where 1 corresponds to *most translucent* and 6 corresponds to *least translucent*, which has been defined as "*closer to opacity*".

3.6.3 Results

The results of the experiment support our hypothesis. Objects were considered less translucent when placed on a black floor with caustics absent. The difference has been statistically significant for all shapes and all surface roughness levels, except for a perfectly smooth surface. We believe that sharp specular reflections present on smooth objects have assisted observers to identify transparent materials regardless of caustics and floor colors. On the other hand, it remains unclear whether the considerable difference in appearance can be attributed solely to the absence of caustics, or whether the overall luminance distribution that was certainly affected by a black floor also played a role.

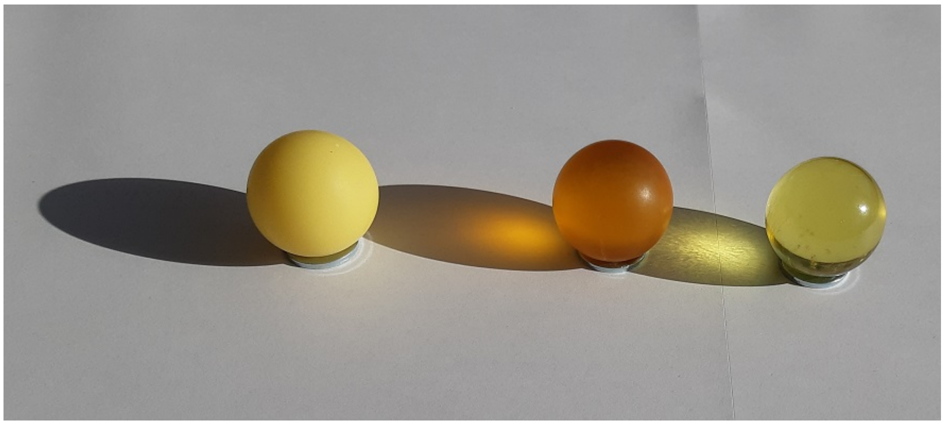

Figure 3.8: The caustics contain a lot of information regarding the properties of a material they are cast by.

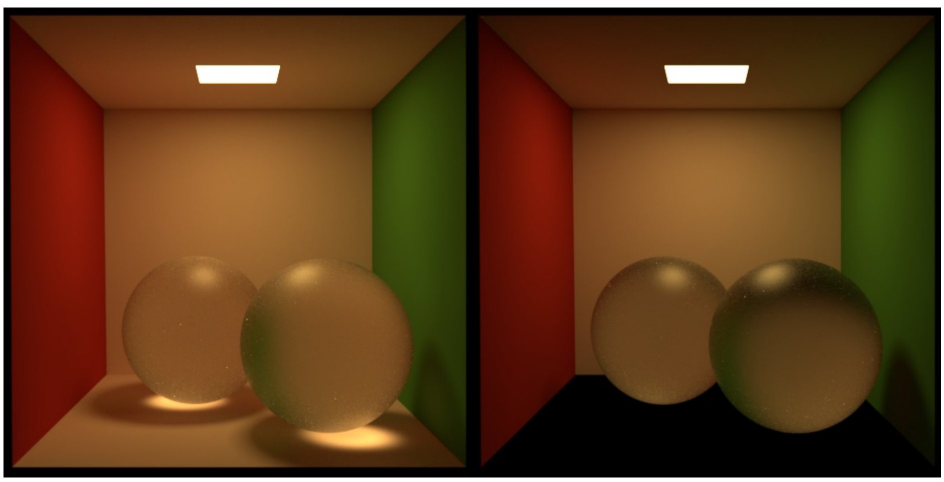

Figure 3.9: Although the material in both scenes is identical, the floor color affects its appearance.

3.7 Article G: Blurring Impairs Translucency Perception

Davit Gigilashvili, Marius Pedersen, and Jon Yngve Hardeberg (2018). "Blurring impairs translucency perception." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 377–382

3.7.1 Objectives

In this work, we have approached the question of translucency perception from the perspective of image quality and its impact on image structure. The HVS pro-

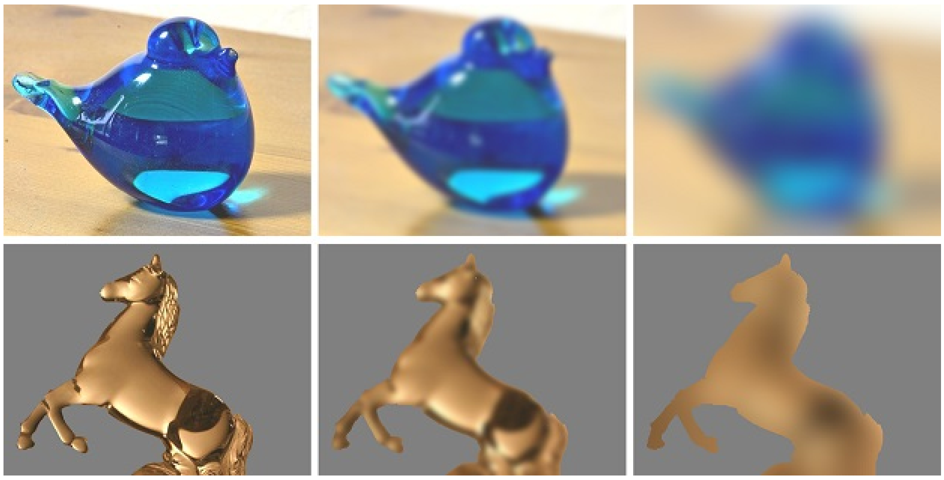

Figure 3.10: Two levels of Gaussian blur have been introduced in the images. The images were presented either with full scene context (top), or just the objects cropped and placed on a gray background.

posedly relies on luminance contrast information to perceive translucency (Fleming and Bülthoff (2005), Xiao et al. (2014), and Motoyoshi (2010)). The image blur has been shown to impair material categorization (Sharan et al. (2014)). We hypothesized that blurring the image removes necessary cues and decreases the perceived degree of translucency. The objective of this work has been testing this hypothesis.

3.7.2 Methods

We introduced different levels of Gaussian blur in RGB photographs of the glass objects from the Flickr Material Database (Sharan et al. (2014)). We hypothesized that if the blur was imposed on the entire scene, the HVS could to some extent discard its effect and keep perceived translucency relatively constant. To study this hypothesis, some blurred objects have been cropped and placed on a homogeneous neutral gray background, while others have been included in the experiment with the full scene context. The examples are illustrated in Figure 3.10. Paired-comparison experiments have been conducted on a calibrated display under controlled laboratory conditions. 20 observers participated in the experiment.

3.7.3 Results

The analysis of the experimental data has provided indications in support of our hypothesis. The higher the Gaussian blur, the weaker the apparent translucency. Contrary to our expectations, this effect has been stronger on full scene images than on the cropped ones.

3.8 Article H: Image Statistics as Glossiness and Translucency Predictor in Photographs of Real-world Objects

Davit Gigilashvili, Midori Tanaka, Marius Pedersen, and Jon Yngve Hardeberg (2020). "Image Statistics as Glossiness and Translucency Predictor in Photographs of Real-world Objects." In: *10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*. Vol. 2688. CEUR Workshop Proceedings, pp. 1–15

3.8.1 Objectives

Luminance statistics have been shown to co-vary with glossiness (Motoyoshi et al. (2007)) and translucency (Fleming and Bülthoff (2005)). In this work, we studied whether the first four moments of the luminance histogram and the area covered with specular highlights are correlated with gloss and translucency in real-world photographs. Unlike unnaturally perfect computer-generated imagery, we photographed the objects which have visible unintended artifacts, which also permits us to test the robustness of the aforementioned metrics.

3.8.2 Methods

We photographed spherical objects from the *Plastique* artwork collection (Thomas et al. (2018)) that came in three different levels of surface coarseness, three hues and different concentrations of scattering colorant particles inside the volume. Translucent objects have been photographed twice - on white and black backgrounds. The object was segmented from the background and the statistics of the CIE XYZ luminance channel (Y) have been analyzed. Furthermore, *k-means* clustering was conducted, in order to determine, whether the five statistical metrics (the first four moments of the histogram and the area covered with the highlights) are good predictors of the object's class. The example of the images is illustrated in Figure 3.11.

3.8.3 Results

The major takeaways of the work can be summarized as follows:

- As the surface becomes rougher, skewness and kurtosis of the luminance histogram decrease.
- Although specular highlights cover less than 1% of the total visible area of the sphere, they skew the luminance histogram and render a convincing glossy appearance.
- Mean luminance alone is not a good predictor of gloss. However, in particular cases mean luminance can provide information about contrast gloss (Hunter (1937)).

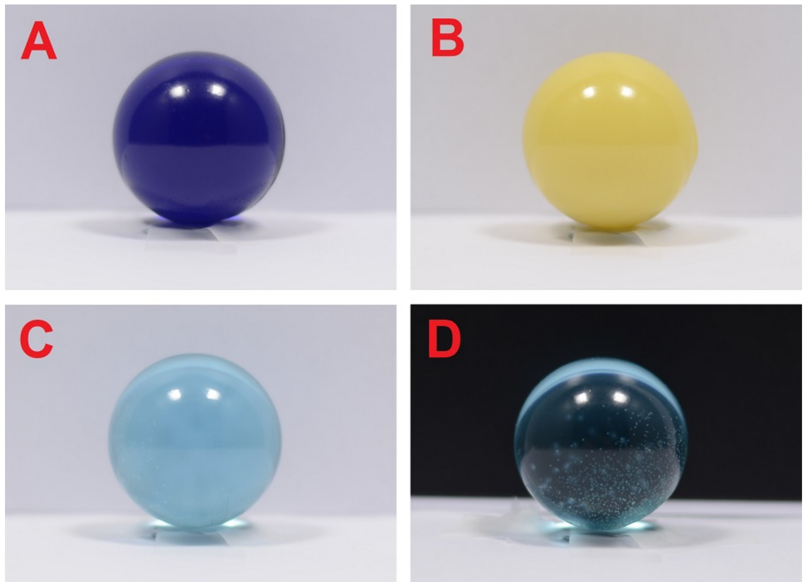

Figure 3.11: The statistics vary between blue and yellow opaque objects, the darker one (A) permitting visibility of a more clear reflection image of the environment than the lighter (B) one. The translucent object shown in illustrations C and D is the same, but its appearance differs considerably due to the change in the background color. Some artifacts and bubbles can be detected in image D.

- We did not identify any correlation between surface roughness and standard deviation, contradicting previous findings (Wiebel et al. (2015)).
- Change in variance and mean luminance across different backgrounds could potentially be predictors for translucency.
- The robustness of these findings can be compromised by dynamic and variable environment and might not be applicable to objects with low surface curvature.
- Image statistics alone are not enough for deducing glossiness and they are limited with photo-geometric constraints and semantic understanding of scene composition.

3.9 Article I: On the nature of perceptual translucency

Davit Gigilashvili, Jean Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2020). "On the Nature of Perceptual Translucency." In: *8th Annual Workshop on Material Appearance Modeling (MAM2020). Eurographics Digital Library*, pp. 17–20

3.9.1 Objectives

The concept of translucency as an appearance attribute is oftentimes more abstract than that of other attributes. We usually refer to the colors of the objects or to their textures, but rarely to their translucency. Hence the precise meaning of the term is not accepted universally. The objective of the position paper was to postulate ambiguities about perceptual translucency which make research on translucency perception difficult to conduct, communicate and interpret.

3.9.2 Summary

We have identified five issues observed throughout our experiments which we believe should be resolved in order to advance the translucency perception research:

- **Definition and conceptual understanding -** no single standard definition exists for translucency as an appearance attribute (Eugène (2008)), which leaves room for interpretation. This raises the question of how it should be defined to the participants of psychophysical experiments.
- **Perceptual dimensions of translucency -** some appearance attributes might be disentangled into distinct dimensions, such as hue, chromaticity and lightness, for color. It is unclear, whether translucency should be measured psychometrically as a whole, one-dimensional phenomenon, or it has distinct dimensions, where *haze* and *clarity* could be potential candidates.
- **Relation with transparency and opacity -** translucency exists between the extremes of transparency and opacity (CIE (2006)), but it remains unclear how it relates with them; is translucency orthogonal to transparency and opacity, can they co-exist to some degree, or are they mutually exclusive? We proposed that the magnitude of translucency might be conceptualized as a bell-shaped curve (illustrated in Figure 3.12) which is low near the extremes, gradually increases and peaks somewhere in between them.
- **How to quantify perceptual translucency -** it is not clear what *more translucent* implies; is it proximity to transparency, to opacity, or to some hypothetical maxima between the two? This uncertainty makes it difficult to apply magnitude estimation (Torgerson (1958)), or psychophysical scaling techniques (Engeldrum (2000)), such as rank order, for studying translucency.
- **Translucency constancy of objects and materials -** little is known how constant translucent appearance is across different conditions (Xiao et al. (2014)). Objects made of an identical material might differ considerably in terms of apparent translucency due to their shape and scale. We have noticed that observers found cross-shape translucency comparison difficult, because it was unclear what should be assessed - an appearance of a given object, or an absolute, shape-independent property of a material. We propose that perceptual translucency is a context-specific concept with limited constancy.

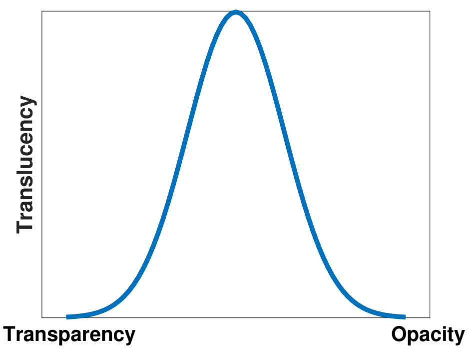

Figure 3.12: Translucency might be gradually increasing, reaching its peak and decreasing between transparency and opacity. However, transparency and opacity are unlikely to be discrete points - thus, translucency may co-exist with them.

3.10 Article J: Translucency perception: A review

Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (n.d.). "Translucency perception: A review." In: *Accepted for publication in the Journal of Vision*, 45 pages

3.10.1 Objectives

Translucency perception research is relatively novel and the findings are scattered around the literature. The objective of the review article is to put the current knowledge status about translucency together which includes the recent findings made in the course of this doctoral program. Additionally, the review aims to particularize existing knowledge gaps and outline the avenue for future translucency perception research.

3.10.2 Methods

We have conducted an exhaustive literature review about translucency perception, which covers a broad range of works in vision science, computer graphics and visual arts. First of all, we summarized and listed which factors and parameters affect apparent translucency and demonstrated these effects based on physically-based renderings and RGB photographs. Afterwards, we reviewed partial models of translucency perception proposed by other authors and manipulated computer-generated imagery to spotlight the limitations of those models.

3.10.3 Summary

The compilation of the literature led us to identify the factors that had been previously shown to be affecting translucency perception. These factors include: subsurface absorption and scattering coefficients, scattering phase function, index of refraction, object's scale and structural thickness, its surface roughness and geometry, illumination direction, illumination structure, object's color, its glossiness, caustics, motion and scene dynamics and high-level cognitive interpretation by the observers.

Analysis of the partial models of translucency perception made us conclude that a full perceptual model of translucency, which could simply take the scene and material properties as an input and provide an estimation of a perceptual correlate, remains beyond reach nowadays. It seems unlikely that modern vision science would solve this problem anytime in the foreseeable future. The knowledge status on translucency perception mechanisms can be conceptually summarized as follows:

1. It seems that neither luminance nor spatial information alone is enough for estimating apparent translucency. The HVS seemingly uses some sophisticated combination of the both.
2. Spatial regions where a photon can go through easily look brighter and contain rich information about material translucence. Examples of this kind of regions are edges, thin parts and sharp fine details of a surface geometry.
3. The regions that are usually shadowed in opaque objects are also informative about translucency, as they look brighter in translucent materials.
4. To summarize the two previous points: if in absence of subsurface light transport a considerably smaller amount of light could have reached a particular region, this region can be diagnostic for material translucence.
5. Understanding how much light could or could not have reached a particular region inherently involves understanding the surface geometry and global correlation among different local regions.
6. It is not known how the human visual system segments an image, how it identifies informative regions and how it calculates surface geometry.
7. These calculations are not unique and vary considerably across individuals. There can be multiple translucency cues in a proximal stimulus and different people can rely on different ones for yet unknown reasons.

Chapter 4

Discussion

This chapter is structured as follows. First of all, we answer the research questions presented in Section 1.3 and discuss what we have learned regarding them. This is followed by a *general discussion*, which addresses the topics that are not related to one particular research question. Finally, we analyze the limitations of our findings.

4.1 Research Questions

4.1.1 How do people behave when assessing appearance, and which factors facilitate this process?

In ***Articles A, B*** and ***C*** we allowed observers to interact freely with the objects while assessing their appearance. We noticed that they move objects, observe them from different viewpoints and under different illumination conditions, and in short, they rely on scene dynamics. This is consistent with the proposal by Fleming (2014) who argues that the HVS somehow identifies the salient features of materials, builds an internal generative model and characterizes systematic changes of these features to learn how materials and respective features behave under different conditions. It seems that humans are unconsciously aware of the rules which exist around us and which define the systematic changes in appearance. For example, objects look more translucent when they are lit from behind (Xiao et al. (2014)). We noticed that in order to assess translucency, people pick up objects and look through them towards the sun or an artificial light source. Apparently, they attempt to detect whether objects shine under back-light, which would be an indicator of subsurface light transport. It has also been shown earlier that specular reflections on a rotating object, unlike surface texture, remain static relative to the observer (Wendt et al. (2010) and Doerschner et al. (2011)). We observed that people move objects or their heads to assess glossiness - seemingly trying to separate specular reflections from surface texture. In general, humans use motion, whenever they are allowed to do so, and observe the change in appearance, which eventually helps them deduce material properties. This seems to be a result of prior training we undergo since birth. The motivation for motion is what we call *comparison with a reference* in our *qualitative model of material appearance* (***Article B***) and which, in our data, turned out to be of critical importance for assessing the appearance of the objects. Appearance is not assessed in isolation, it needs a reference which facilitates quantification of appearance. In the above-mentioned cases, humans use the appearance of a given object under a different condition as a reference and quantify the change relative to that when an object is moved to a new condition. For instance, comparing the appearance of a given object between front-lit and back-lit conditions can help observers deduce whether the object is translucent or opaque.

The importance of a proper reference has been further observed in subsequent works. In ***Article E***, we noticed how observer responses changed between different anchor (i.e. reference) pairs, while ***Article D*** demonstrated that the emphasis put on *subsurface* scattering differences depends on the *surface* scattering difference, i.e. on the compared materials. A reference also plays an important role in semantic communication of appearance - for instance, when a corpus of the visual stimuli is composed of objects with different shapes, the shape is more frequently mentioned in the description of appearance, rather than when the shape is identical in the entire corpus.

In addition to visual information, which encapsulates spatial and temporal aspects, humans tend to use multisensory information when it is available. We observed that visual information alone has oftentimes not been enough for identification of materials and accurate estimation of their mechanical properties (e.g. solid plastic objects were usually described as soft and elastic unless they had specular reflections). Observers use information from other senses to verify what they see, because "*the appearance of glass paired with a pepper sound is perceived as transparent plastic*", as noted by Fujisaki et al. (2014). People use tactile information to assess the surface of an object and auditory information to identify materials. Other senses complement the vision and integrated multisensory information is seemingly analyzed for the purpose of material identification, as well as for assessment of their appearance. Our study is limited to collecting these observations and we have not explored exactly why people need to identify materials or estimate their mechanical properties when assessing visual appearance. One explanation for this could be the existence of priors about familiar materials, i.e. they expect a particular appearance for particular materials (e.g. glass is usually glossy and transparent).

Material appearance assessment, both behavior and semantic communication are impacted by the individual observer's background and subjective traits, which is consistent with prior works (Hutchings (1995a) and Hutchings (1995b)). For example, observers with expertise in appearance studies inspect objects more scrupulously and tend to rely on literature definitions more often than artists. Semantic communication is often impacted by personal experience, as objects' descriptions involve comparisons with subjective references, such as materials for observer's childhood memories. On the other hand, communication of appearance inherently involves some degree of objectivity. Otherwise, it might have been impossible and unintelligible, which is not the case. For example, there is to some extent common understanding of what *green* and *jelly-like* mean.

To the best of our knowledge, **Article B** has been the first attempt to study the social and behavioral basis of appearance assessment. However, our observations come with particular limitations and need to be taken with care. Our assumption that the original social experiment was "*as natural and as close as possible to real-life situations*" does not fully hold. In order to motivate the social interaction, the process was driven by artificially imposed tasks, which themselves are unnatural and rarely performed in real life. We hardly ever rank objects by their glossiness and translucency. Therefore, this does not guarantee that the behavioral patterns we observed are identical to those applied when performing daily routines. Besides, the application of observed behavioral patterns was subject to the presence of particular illumination conditions. For instance, people put objects under high-illuminance backlight (looking towards the sun), but this might not be possible in diffuse ambient illumination.

4.1.2 Does the human visual system manifest constancy in translucency perception similarly to color constancy, and to what extent?

It has been demonstrated earlier that translucency constancy fails across different shapes and illumination conditions (Fleming and Bülthoff (2005) and Xiao et al. (2014)). Xiao et al. (2014) argue that the robustness of constancy depends on the scattering phase function and its location in the 2D perceptual space proposed by Gkioulekas et al. (2013). Our observations are consistent with the state-of-the-art and in ***Article J*** we have demonstrated that translucency constancy fails due to change in illumination direction (refer to Figure 4.1), as well as due to object's scale and shape (refer to Figure 9 in ***Article J***). Moreover, we have shown in ***Article F*** that translucency constancy is compromised by the environment color and subsequent removal of caustics. Color constancy relies on sensory adaptation as well as estimation and discounting of the illumination color. Sensory mechanisms of translucency are poorly understood. It is likely that its complex nature, which involves interpretation of luminance and spatial information, complicates sensory adaptation across different conditions. Moreover, discounting the effects of the illumination seems a challenging task, as translucency is a result of complex light and matter interaction, which for the HVS is difficult to understand and invert (Fleming and Bülthoff (2005)). For example, we observed in ***Article F*** that translucency constancy failed when the object involved complex surface scattering, while perceived translucency remained relatively constant, when smooth transparent objects were assessed, because it was easier to understand (and probably invert) the underlying optics.

In ***Article B***, we argue that constancy fails faster in real-life situations where interaction is possible. Therefore, this can be considered a limitation of our studies

Figure 4.1: The object is identical, however, the illumination geometry varies from back-lit (left) to side-lit (middle) and front-lit (right). Perceived translucency changes with the change of the illumination direction, as the back-lit object looks more translucent than a front-lit one.

conducted on displayed still images, as the absence of scene dynamics, interaction and multisensory information might have exaggerated the constancy of translucency appearance.

4.1.3 Does translucency contribute to glossiness perception?

Hunter (1937) argued that "*reflection distribution functions, though complex and cumbersome, offer the only means by which the reflectance properties of surfaces responsible for their glossiness may be completely specified.*" However, we believe that an apparent gloss model should also include subsurface scattering distribution functions. ***Articles A, C*** and ***D*** have provided ample evidence that even when the spectral reflectance is identical, subsurface scattering properties affect the perceived magnitude of gloss. However, our data does not permit understanding exactly how the subsurface scattering is contributing to gloss perception. Subsurface scattering parameters should be sampled more densely in order to quantitatively model the correlation between optical properties and perceived gloss. Moreover, it needs to be explored how subsurface scattering impacts image structure - i.e. exactly which image cues are affected by subsurface scattering and how those cues co-vary with subsurface scattering parameters.

For simple shapes, which permit observation of the reflection image of the environment, higher subsurface absorption generates higher contrast between specular and non-specular areas, also permitting to detect the reflection image more clearly. Interestingly, transparent objects which generate more caustics and permit to see-through, are also perceived glossy, either due to overall shininess or high-level cognitive factors. These trends are consistent between ***Articles A*** and ***C***. It was already proposed earlier that gloss cannot be characterized by specular reflectance only, and that diffusely reflecting areas contribute as well (Hunter (1937), Hunter and Harold (1987), Pellacini et al. (2000), and Thomas et al. (2017)). However, what is known as a diffuse reflection in simplified models, is actually light scattered backwards from the superficial layers of the subsurface. If the extinction coefficient is low enough, light can re-emerge far from the point of incidence, contributing to non-specular areas in that region. On the other hand, complex shapes which do not permit observing clear reflection images, tend to differ in perceived gloss by the amount of caustics and highlights which result from subsurface scattering and back-reflections. This trend is consistent between ***Articles C*** and ***D***, which studied the *Plastique* (Thomas et al. (2018)) female bust and Stanford Lucy (*The Stanford 3D Scanning Repository* (1994)) shapes, respectively. This made us conclude that when studying gloss perception, it is essential to include complex shapes and not generalize the findings based on simple shapes, such as a sphere. It is interesting to identify the source of observer inconsistency reported in ***Articles A*** and ***C***. We believe these differences are produced due to different semantic interpretation of the concept rather than differences in physiological sensory stimuli or the conditions of observation (the trends have been similar for all nine different conditions in which the experiment was conducted in ***Article C***). For instance, we believe experts and non-experts saw the same, but experts tied all stimuli simply because they relied on the official definition of the term *gloss*. Interestingly, the dichotomy in gloss perception between the distinctness-of-image and luminance-based approaches has been also observed by Leloup et al. (2012).

The findings of ***Article D*** are limited by the fact that the dynamic range of the stimuli was small (we use clipping to convert high-dynamic range images to PNG, in order to make them compatible with observer displays). Besides, only still images have been used. Limited dynamic range and absence of motion cues make separation of reflection and transmission components more challenging than it is in real life.

Finally, our studies are limited to finding the correlation between optical properties and the magnitude of perceived glossiness. However, we believe observed phenomena need to be explained from an image statistics perspective. Our studies do not investigate how subsurface scattering affects image structure and which image statistics are variant among the levels of subsurface scattering.

4.1.4 Does the shape of the object impact the perceived magnitude of translucency?

For a given material with given absorption and scattering properties, the likelihood that a photon propagating through it either gets absorbed or scattered, increases with the distance that it needs to travel (Urban et al. (2019)). This means that the luminance distribution, which is supposedly a cue for translucency perception, will vary with object's thickness and size. For example, it has been shown that if the extinction coefficient is high, it is usually the edges which are diagnostic for translucency (Fleming and Bülthoff (2005) and Gkioulekas et al. (2015)). In ***Article A***, we observed that the presence of thin parts can compensate for a higher extinction coefficient in the material. The HVS cannot invert optics (Fleming and Bülthoff (2005)), which makes it challenging to isolate effects of material

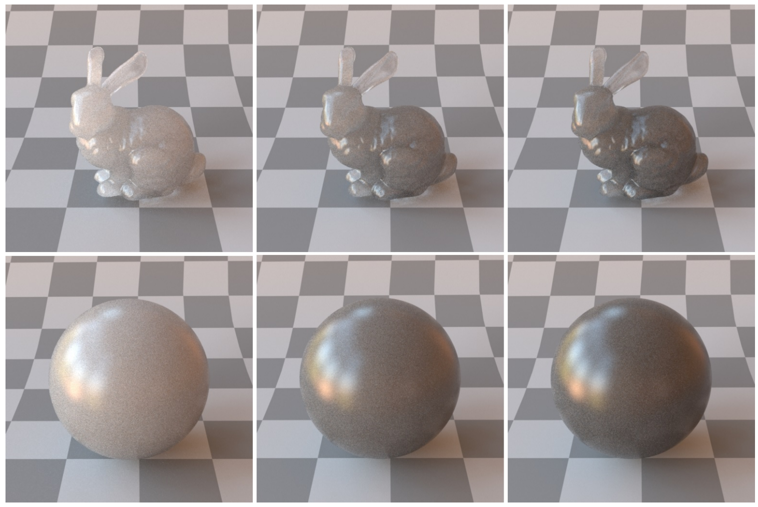

Figure 4.2: Objects in the same column are made of the identical material. However, due to the smaller scale and presence of thin parts, the Bunny has more cues evoking perception of translucency.

properties and illumination from those of object's shape. This is especially true, when the range of structural thicknesses is large and the object has fine details (Xiao et al. (2014)). Objects in each column in Figure 4.2 are made of an identical material. However, the Bunny possesses more translucency cues and evokes a stronger perception of translucency than a thick and compact spherical object.

In addition to the macro-scale shape of the object, we have demonstrated in ***Article F*** that the micro-level composition of the surface also impacts apparent translucency. It turns out that there is a monotonic and linear relationship between translucency and surface roughness (root mean square slope of the microfacets), when subsurface scattering is negligible.

Our works do not attempt to explain how people deal with the conceptual ambiguities between object and material translucency. If an object is composed of parts with varying thickness, some parts of it look translucent, while others look opaque. This confuses observers who cannot decide whether they should assess translucency as a generic property of a material or an object, or translucency for each particular region of the object individually.

4.1.5 Does the shape of the object impact detection of translucency differences?

Article E has provided evidence in support of our hypothesis that it is easier to detect suprathreshold translucency differences on shapes that have thin parts. This is consistent with the state-of-the-art claiming that thin areas contain much information about material translucence (Fleming and Bülthoff (2005), Gkioulekas et al. (2015), Xiao et al. (2014), Sawayama et al. (2019), and Nagai et al. (2013)).

Although the qualitative trend is apparent, we have not been able to model this correlation quantitatively. As we do not know exactly which image cues the HVS relies on, we cannot construct a shape descriptor that could correlate with perceived translucency differences. For instance, we hypothesized that a histogram of surface-to-medial-axis distances might be such a descriptor. However, we do not know whether the abundance of thin parts is needed, or if even a single thin region would suffice for the HVS. The two cases generate contrasting surface-to-medial-axis histograms, but might be identical in terms of detectability of translucency differences.

As shown in Figure 3.7, spherical objects make it difficult to detect translucency difference. According to Marlow et al. (2017), the HVS relies on co-variance between shading and geometric information. However, a compact spherical object which contains no concavities or bumps, leaves less room for observation of this co-variance. We believe that this finding has implication for future studies on translucency perception, as we need to reconsider the common practice of using a sphere as a shape of choice in psychophysics.

4.1.6 Does the magnitude of subsurface scattering impact our ability to detect translucency differences?

The discrimination of the different levels of translucency depends on the magnitude of the subsurface scattering. A similar phenomenon has been described for gloss; Cheeseman et al. (2021) noticed that the HVS is less sensitive to gloss differences in case of the high magnitude specular reflectance. ***Article E*** has shown that changes in subsurface light transport properties are detected faster by humans when it is possible to see the background through this material. First of all, this is consistent with the Steven's and Weber-Fechner laws (Stevens (1960) and Fechner et al. (1966)). Secondly, this demonstrates that transparency and translucency perception mechanisms are fundamentally different and they should be addressed separately. While transparency, on the one hand, is judged based on blur and contrast of the background seen through the object (Singh and Anderson (2002)), translucency perception, on the other hand, involves assessment of luminance distribution and low level image cues (Motoyoshi (2010), Nagai et al. (2013), and Fleming and Bülthoff (2005)). It is intuitive that the impact is larger when direct distortion of the background is visible. This has implication for designing future psychophysical studies. As it turned out, transparent anchor pairs are poor references for translucent test pairs when the method of constant stimuli is used.

4.1.7 Does appearance assessment differ between physical objects and displayed images, and how vital is the direct interaction with the objects when judging their appearance?

Using physical objects and permitting direct interaction introduces multisensory, binocular and motion cues, all of which have been shown to be important for assessing material properties (Fujisaki et al. (2014), Obein et al. (2004), Wendt et al. (2008), Doerschner et al. (2011), and Tamura et al. (2018)). Appearance constancy is unrealistically high for displayed images, while for physical objects they fail faster (Filip et al. (2018)) as more cues and information are available. The viral images that trick the visual system, such as *#TheDress* (Brainard and Hurlbert (2015) and Lafer-Sousa et al. (2015)) and *shiny legs* (Molloy (2016)), would most likely not have happened if they were presented as physical objects rather than images.

We found in ***Article B*** that lack of tactile and auditory information leads to misidentification of materials and misestimation of their mechanical properties, such as elasticity, fragility, softness and hardness. Interaction with physical objects is a natural way to assess material appearance. This is what we experience in daily lives and this is what our visual system is trained on. We observed that all subjects moved and interacted with the objects when they were permitted to do so (however, we cannot rule out a possibility that their behavior was induced by the instructions, which explicitly mentioned that interaction was permitted). Interaction provides a broad range of references, for instance, they can inspect a translucent object on homogeneous and heterogeneous backgrounds, and use the variation in appearance for assessing translucency. In displayed images, observers may lack the proper reference to assess appearance adequately. As noted by Anderson (2011), in such experiments, the experimenter might unwittingly remove information that is fundamental for the HVS - providing little insight into the real mechanisms of appearance perception. Comparison of the results of ***Article A*** and ***C*** with that of ***D***, has shown that observers are more consistent when judging computer generated imagery, while a broader range of cues and opinions emerge when the stimuli are presented in the form of physical objects.

The discussion on this topic in this thesis is fundamentally limited by the fact that we have not studied the physical and digital representation of the identical objects. This makes it impossible to rule out that the difference observed between physical and digital stimuli where rooted in the properties of those stimuli and not in the method of their representation. Future work should compare appearance assessment between a physical object and its digital twin.

4.1.8 Does presence of caustics impact the perceived magnitude of translucency?

We observed in ***Articles A*** and ***B*** that caustics is a widely used and a reliable cue for judging the translucency of a material. Caustics is a familiar cue for the HVS, as we encounter them on a daily basis. For instance, when a wine glass or a vase projects a light pattern onto a table, we understand that the pattern is produced by subsurface transport of light. In ***Article F*** we showed that placing an object on a surface that removes caustics decreases the magnitude of perceived translucency. This can have implications for the retail industry, as the appearance of translucent products might depend on the surface color they are placed on. In some particular

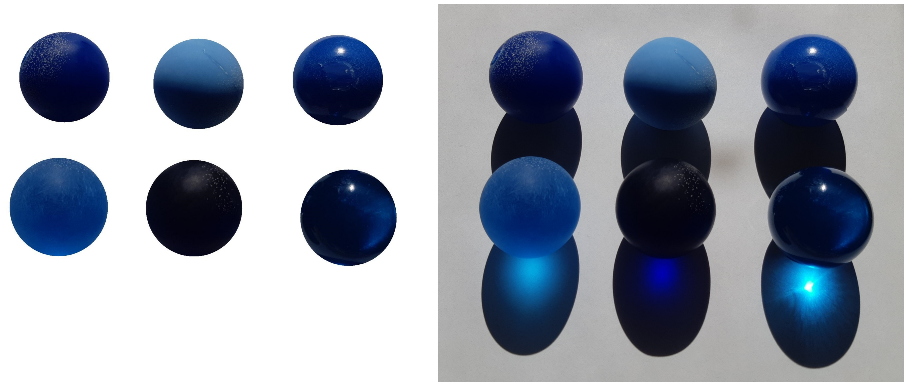

Figure 4.3: While the object's body might look fully opaque (e.g. middle object in the bottom row), caustics provide rich information about subsurface light transport properties of a material.

cases, caustics can be the only indicator of translucency, as in Figure 4.3.

This study comes with particular limitations. First of all, changing floor color to black not only removes caustics from the image structure, but also impacts overall luminance distribution considerably. Objects and scenes become generally darker. Therefore, it cannot be ruled out that the results are impacted not only by the absence of caustics, but other cues that we unwittingly affected with a black floor. Secondly, the study assessed objects with no subsurface scattering. The translucent appearance was generated with surface scattering only. However, the structure of the caustic pattern is impacted by subsurface scattering as well (see the bottom row in Figure 4.3) and studying these kind of materials can reveal more information about caustics as a translucency cue.

4.1.9 Does image blur impact the perceived magnitude of translucency?

Article G has shown that the impact of image blur on the perceived magnitude of translucency is statistically significant, blurrier objects appear less translucent. This is counter-intuitive at first glance, as translucency by definition implies that the image emerging from a translucent object is blurred (ASTM E284-17 (2017), Eugène (2008), and Gerbino et al. (1990)). Furthermore, it has been shown that decreased luminance contrast evokes the perception of translucency (Fleming and Bülthoff (2005), Motoyoshi (2010), and Nagai et al. (2013)). However, these notions need to be taken with care. When discussing decreased luminance contrast, it is usually implied that the luminance contrast between specular and non-specular areas is decreased because luminance increases in non-specular areas and remains the same in the specular highlights (Motoyoshi (2010)). For instance, if concavities are shadowed in opaque objects, they look brighter for translucent ones, as photons reach them via subsurface layers. As shown earlier, local luminance statistics and relation among them are more diagnostic for translucency than global ones (Motoyoshi (2010) and Nagai et al. (2013)). However, when blur is imposed on the entire image, we affect both specular and non-specular areas. This makes it difficult to estimate the spatial information, in particular, surface geometry and texture of the object, which according to Marlow *et al.* (2017) are essential to translucency perception. The fact that perception of translucency implies comprehension of spatial information is intuitive, because a color or a luminance intensity of a single local point can be simply produced by surface reflection, while the spatial variation of colors is what indicates subsurface light transport. Therefore, blurring imposed globally removes information about the spatial distribution of luminance intensities, which means that the image cues evoking perception of translucency disappear. Imagine an extreme case, when a blur produces a homogeneous patch with no intensity variation - indeed, no homogeneous patch can appear translucent, as translucency cues inherently rely on spatial information. If we gradually decrease the blur, more spatial information and more translucency cues can emerge. On the other hand, we understand that the impact of blur might be negligible for see-through objects which lack specular reflections and are placed on a homogeneous background, because the spatial variation of the luminance intensities will be low even for the sharp images.

This finding is important for future psychophysical experiments. We can conclude that visual acuity of the observers and distance to the stimuli have a considerable impact on translucency perception and those factors need to be considered in the experimental design. Additionally, image quality can also impact the results of such studies.

The study has several limitations: first of all, it has been possible to recognize the same object in the images with different degrees of blur, which could tempt the observers to consider them equally translucent. Secondly, the sampling in the blur parameters is very sparse - just two levels of Gaussian blur are applied which differ considerably in the magnitude of imposed blur. This does not permit us to model the impact quantitatively, neither to identify when the impact of blur on translucency becomes noticeable. More levels of blur should be studied for this purpose. Finally, all objects studied in the experiment are glossy and have clearly visible specularities. It is interesting to explore whether the impact would be as strong for objects with no specular regions.

4.1.10 Can the luminance statistics be used for prediction of apparent gloss and translucency?

In ***Article H***, we propose that objects with smooth surface and visible specular reflections produce highly skewed luminance histograms, being consistent with Motoyoshi et al. (2007) and Landy (2007). Additionally, mean luminance might be an indicator of potential contrast gloss, if specularities are present. This is consistent with the state-of-the-art. Leloup et al. (2011) proposed a perceptual gloss metric that estimates gloss by comparing the luminance measured in the specu-

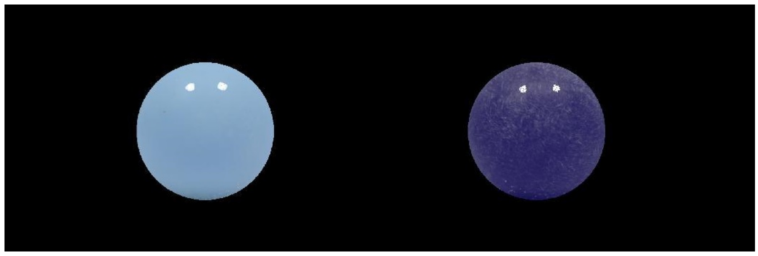

Figure 4.4: Specular highlights are superimposed on the photographs of rough spherical objects. While both highlights are artificial, the left object looks glossier due to the lack of artifacts, while the scratches help us know the right object is not smooth, i.e. not glossy. In the latter case, the highlights look like artefacts rather than specular reflections.

lar and non-specular areas. We also demonstrated that the impact of background change on luminance variance can be used for measuring translucency. However, it is important to consider that both for gloss and translucency, spatial information has critical importance, which cannot be captured with luminance histograms alone. Identical histograms can be produced with images of glossy and matte objects, as well as random re-arrangement of their pixels. Image statistics are usually subject to strict photo-geometric constraints, and perception of gloss, as demonstrated in Figure 4.4, involves complex cognitive understanding of the scene and geometry. When the origin of the detected image structure is not clear (as in ***Article C*** and *shiny legs* image (Molloy (2016))), then given image statistics might produce contrasting perceptions.

Moreover, image statistics are by no means a robust measure of material properties. We have demonstrated that for highly reflective materials, image statistics can vary considerably due to changes in the environment and light field, and meaningful statistics might not be possible to be extracted from planar surfaces and objects with low surface curvature. This is especially true for metallic samples and automotive industry applications.

It is worth mentioning that ***Article H*** correlates physical properties with the luminance statistics. All reasoning on perceptual aspects is based on authors' perception and no psychophysical experiments have been conducted to support those observations. Besides, we studied *JPEG* images instead of *RAW* ones. The JPEG compression together with non-linearity of the acquisition system can be a source of inaccuracies in the luminance information and the luminance intensities recorded by the camera can differ from what the HVS observes.

4.1.11 What are the major obstacles to advancing translucency perception research?

We argue in ***Article I*** that one of the major factors that might undermine the advance of translucency perception research is the problems with the definition and understanding of the term.

Questions about psychophysical mechanisms of translucency perception remain unanswered, which could be cleared up by carefully planned psychophysical experiments. The value of the experimental results greatly depends on observers' adherence to the instructions. Hence, properly formulated instructions have vital importance for the reliability of the experimental data. We have observed that the interpretation of the term *translucency* varies substantially among observers. This fundamental problem might compromise experimental results, lead to miscommunication in the scientific channels and cause misinterpretation of the findings.

For instance, we observed in ***Article A*** that when asked to assess translucency, some subjects rely on preservation of the transmitted image structure, while others try to quantify the radiometric amount of the transmitted light. Unlike *clarity* and *haze*, *translucency* remains largely subjective and the findings are usually limited to particular interpretation in a particular community. We believe that standardization of definitions, measurement and observation conditions are essential for the rapid advancement of this topic.

Additionally, the way visual stimuli are presented to the observers also implies some limitations and the risk of unwitting removal of important translucency cues. For instance, computer-generated stimuli usually limit interactivity and multisensory information.

4.1.12 What is the knowledge status on translucency perception and where should we go next?

The knowledge status and future perspectives in translucency perception research are put together in ***Article J***. Current knowledge on translucency perception is mostly limited to the impact of particular optical parameters. Translucency perception research is in its infancy and partial models proposed by different authors (such as Motoyoshi (2010), Fleming and Bülthoff (2005), and Nagai et al. (2013)) attempt to correlate local image statistics with the perceived magnitude of translucency. However, it remains unknown how the HVS identifies and weights these local regions and why they differ across individuals. The role played by high level vision, memory and cognitive understanding of the scene also remains unclear. Unlike the physiology of color vision, the physiology of translucency perception remains largely unexplored. Thus, neither cross-individual differences in translucency perception, nor the limits of translucency constancy are understood.

4.2 General Discussion

Below we discuss some of the general observations we have made about object appearance throughout the entire project. Each subsection covers one of the key observations.

4.2.1 Image cues and [in]constancy of perception

People effortlessly complete complex visual tasks that remain unattainable for machines and instruments (Sharan et al. (2014)). However, understanding the mechanisms of this ability and predicting human perception based on measurable physical properties remain a challenging task. Although we have identified interesting phenomena about the perception of object appearance, in its broadest sense, we still have not been able *"to obtain numbers that are representative of the way objects and materials look"* (Hunter and Harold (1987)). If we do not comprehend how humans perceive appearance, it will be difficult to mimic this ability with machines.

It is unlikely that the HVS estimates optical properties and light and matter interaction, but it might rather be relying on simple image cues and rules of systematic changes those cues undergo across different conditions (Fleming and Bülthoff (2005) and Fleming (2014)). Moreover, the visual stimuli do not even need to conform to the laws of physics, as long as they generate image cues which are familiar for the HVS – for instance, adept artists can generate vivid impressions of various material appearances simply based on these image rules and recipes, and paintings evoke the perception of translucency, gloss and other appearance attributes in a robust and realistic manner (Cavanagh (2005) and Di Cicco et al. (2020a)). We are not adept at abstracting material properties from the effects of shape and illumination, which makes our gloss (Vangorp et al. (2007) and Wendt et al. (2010)) and translucency constancy (Fleming and Bülthoff (2005) and Xiao et al. (2014)) imperfect. We believe that the constancy of perceived translucency and gloss is limited by the variability of the image cues across different conditions. For judging translucency and glossiness, we seemingly rely on image structures which themselves are not invariant across different shapes and illumination directions. For example, we have observed that objects with thin parts look more translucent (***Articles A*** and ***J***) and materials with identical spectral reflectance differ in apparent gloss (***Articles A, C*** and ***D***), because in both cases, subsurface light transport affects the spatial distribution of the luminance intensities in the proximal stimulus. This once again indicates that our ability to split image structure into reflectance and transmission components, to understand and isolate effects of shape, material and illumination, is substantially limited and we are relying on image structure and statistics.

4.2.2 We rely on references and this can aid metrology

People rely on the references which they extract from the 2D retinal image (***Article B***). When they assess appearance, they are essentially conducting a metrological process where a reference is a unit of measure. If we could identify what these references and their physical correlates are, we could replicate this process with machines. Standardization of reference and units that the HVS quantifies could be fundamental for advancing appearance metrology. However, the high dimensionality of the problem makes it an overly challenging and complicated task. For instance, we need one unit to measure distance and two units to measure velocity (time and distance), while the number of units, or "dimensions" needed for appearance measurement might be impractically high. We have proposed in ***Article B*** that if any appearance ordering system would ever exist, it will be very cumbersome and high dimensional (refer to ***supporting Article K*** for a detailed analysis). This makes us conclude in ***Article I*** that any definition and measurement standards, particularly for translucency, and we believe, for appearance in general, will remain application- and context-specific in the foreseeable future.

4.2.3 Motion leaves less room for uncertainty, which can inspire measurement techniques

Scene dynamics and interaction with the objects play an important role in the assessment of their appearance. The fact that the motion and the ability to inspect a surface from the multiple angles are important for proper estimation of objects' properties has been highlighted by other researchers as well (Ged et al. (2010) and Wendt et al. (2010)). We have observed that humans inspect objects from many different observation and illumination geometries before assessing their appearance. Current measurements of color, gloss and light transmission are however done in predefined geometries (see Pointer (2003) and CIE (2006) for surveys). This further limits the possibility to predict appearance from those measurements. For instance, in order to assess the glossiness of a table, we do not simply rely on specular gloss from the initial observation position, but also move our head to low gazing angles. This means that for quantifying what we perceive, we might need to measure not only specular gloss, but sheen as well. Although we perceive translucency in still images, in real life we rely on scene dynamics and spatio-temporal components – we move objects over different backgrounds (***Article B***) to assess their translucency. We believe that a measurement technique for translucency can be inspired from this observation – we can measure image intensities under different backgrounds and illumination geometries and use variation across them as a correlate for perceptual translucency (as we have proposed in ***Article H***).

A further reason why the temporal component is essential for this kind of metrology is the fact that image statistics are prone to variations, even due to slight changes in the environment (***Article H***). This is especially true for metallic materials, which are used, for instance, in the automotive industry. Therefore, any statistics that are extracted from a single scene and geometry, might not be generalizable and diagnostic enough for translucency and gloss. The commercial relevance of predicting the appearance of the still image scenarios only is limited (e.g. photo-based advertising), while the vast majority of consumer products (e.g. cosmetics, gadgets, 3D printed materials and accessories, video games and computer graphics) are inherently intended for interaction and observation in dynamic scenes. The visual effects that might be achieved in still images, can fail when the interaction is permitted.

4.2.4 It is not just about the low-level vision

What complicates objective measurement and prediction of appearance even further is its multimodal nature. Tactile, olfactory and auditory ("*ah, this sounds like cheap plastic*" phenomenon) information are seemingly used (***Article B***) in yet unidentified ways. Furthermore, in addition to objective visual stimuli that exist in the immersive environment, the perceptual process involves subjective observer-specific factors as well. The knowledge about the translucency perception on the cortical level, as well as the role of cognitive priori, remains virtually non-existent. The only exception is the work by Chadwick et al. (2019), who demonstrated that an observer which suffered from color blindness of a cortical origin was still able to discriminate the levels of translucency, concluding that translucency perception is anatomically independent from color perception on the cortical level. High level cognitive and memory factors ("*this looks like a gummy bear candy I used to have in my childhood*" - ***Article B***) are very difficult to model and quantify. Our ability to unmix absorption and scattering and thus, the accuracy of our perception, might depend on the training our visual system has undergone in the course of the lifetime (Chadwick et al. (2018)). We believe that any statistical model specifying perceived gloss and translucency should be a mixed-effects model, whereas optical properties can be treated as fixed effects, and observer physiology and psychology as a random effect. Intriguingly, the qualitative model proposed in ***Article B*** encapsulates observer characteristics in the conditions of observation, as we observed that the way subjective physiological and psychological aspects contribute to appearance is phenomenologically no different from the contribution of illumination and other extrinsic factors. This might have implication for data analysis and instead of pooling experimental results, appearance perception research might move more towards models tailored to individuals.

4.2.5 Revisiting the qualitative model

The Grounded Theory Analysis permits to return to the original model, and to refine and strengthen it based on new experimental data. We want to highlight that we found ample experimental evidence in support of our research hypotheses. We believe the hypothesis that translucency impacts glossiness perception has the largest generalization potential and it should be scrutinized in future studies. Moreover, the omnipresence of a reference has been observed in the subsequent works as well (e.g. the impact of anchor pair reference in ***Article E*** and the importance of a reference surface scattering in ***Article D***). Additionally, we observed that conditions of observation (***Articles F, G, H*** and ***J***) impact appearance significantly, while task interpretation has affected the methodology selection in ***Article C***. These observations have solidified the qualitative model proposed in ***Article B***. However, a substantial amount of future work is required to achieve general theorization.

4.2.6 Terminology matters: "*material appearance*" versus "*object appearance*"

Finally, we observe that material appearance is a vague and misleading term. Appearance as a visual phenomenon and the respective field studying its nature have been dubbed as *material appearance*. This term has been promoted by a broad range of academic publications (such as Serrano et al. (2018), Lagunas et al. (2019), Dorsey et al. (2010), and Sole et al. (2019)), as well as projects (e.g. MANER[1]) and fora (e.g. MAAP[2]). However, in perceptual experiments we never study materials as abstract entities, but we study objects instead. We need objects to display and represent materials. Indeed, materials possess particular optical properties, which define how light interacts with them. However, we have observed in ***Articles A, B, D, E, F*** and ***J*** that features of an object, such as shape, roughness or its size and thickness also contribute to the visual sensation. In other words, the appearance of a given material might differ considerably across different objects made of this material. This makes us question: can **material appearance** as a term adequately characterize the problem, or should we talk about **object appearance** instead? Moreover, psychophysical studies are usually based on simple shapes, such as spheres (e.g. Pellacini et al. (2000) and Serrano et al. (2018)) or tori (e.g. Fleming and Bülthoff (2005)). Our work indicates that findings based on those shapes might not be generalizable to other, more complex shapes. This means that for appearance modeling and replication tasks each object and shape should be considered individually, which once again brings up a problem of **object appearance** rather than **material appearance**.

4.2.7 Applications

The results obtained throughout this work can not only contribute to the broad range of the industrial applications, where customers' perception of the translucent products is economically important (discussed in detail in section 1.1), but also opens a broad range of new avenues.

Although ***Articles A*** and ***B*** report the hypotheses obtained though the inductive research, which need quantitative validation before being implemented in the industrial applications, the observations obtained from these works can still be applicable to improve the soft and hard metrology techniques of the appearance measurement. For instance, the design of the future psychophysical experiments might be refined with the knowledge on the behavioral observations and the importance of the naturalness of the interaction. We believe this will pave the way for a broader use of the extended reality technologies in psychophysical experiments.

[1]Material Appearance Network for Education and Training. The project funded by the Research Council of Norway. For more details refer to: `https://app.cristin.no/projects/show.jsf?id=675496`

[2]The annual Material Appearance conference at the IS&T Electronic Imaging Symposium. For more details refer to: `https://www.imaging.org/site/IST/IST/Conferences/EI/EI_2021/Conference/C_MAAP.aspx`

Besides, as already mentioned above, the importance of a reference and the scene dynamics might inspire motion-based instrumental measurement techniques for translucency.

The results of the ***Articles C*** and ***D*** have implication for material design in computer graphics and manufacturing. The comparative analyses of these two works can be used for cross-media gloss reproduction – particularly, reproduction of gloss between the physical and the digital light permeable materials.

Article E is important for 3D printing. Cross-shape matching of appearance in the 3D printing applications needs a measure and a space of perceived translucency, which could incorporate the translucency difference metrics. Our results have important implications for the development of such metrics and for cross-shape translucency matching task in general.

Insights from ***Article F*** can be potentially developed into an image-based material measurement technique. Additionally, ***Article F*** along with ***Articles G*** and ***H*** can contribute to the computer vision techniques for the material identification and appearance characterization tasks.

4.3 Limitations

The study comes with a number of limitations that we want to discuss below.

4.3.1 Inconsistent definitions undermine the subsequent analysis

Considering the structure of the work distribution, some trends emerged only at the later stage of the study. The challenge related to the proper definition of translucency, which was discussed in ***Article I***, is the result of the smaller observations collected in the course of the previous studies. Therefore, it has not been until recently that we realized the necessity for consistent instructions and definitions across the experiments. The definition of translucency is inconsistent across our studies. In the experiment reported in ***Articles A*** and ***B***, observers are instructed to rank the objects "*by how the light is going through*", without mention of translucency. In ***Article G***, more translucent was defined as "*transmitting higher amount of light*", while in ***Article F***, least translucent was defined as "*closest to opacity*". The experiment reported in ***Article E*** did not provide any definition for translucency and similarly to Urban et al. (2019) left it for individual interpretation. This inconsistency of the instructions might have affected observers' behavior and the results obtained from different experiments might not be directly comparable.

4.3.2 Our observations might not generalize to all objects, materials and conditions

It is also important to mention that the materials and observation conditions we are testing our hypotheses on represent only a tiny subset of all possible materials and observation conditions that exist around us. In order to keep the length of the experiments within the reasonable range, we limited the number of variable parameters and used only simple materials (e.g. isotropic phase function; wavelength-independent absorption and scattering properties; viewing booth with simple background texture etc.). We are aware that those findings might not generalize well to other conditions and materials and that an extensive amount of future work is needed to determine the limits of our findings.

4.3.3 No method for presenting stimuli is perfect

The findings of our studies are inherently limited by the way the stimuli are presented to the observers. Physical objects and either photographs or computer generated images displayed on a monitor all come with their advantages and limitations (refer to ***Appendix 1 in Article B*** for a detailed analysis). Psychophysical experiments reported in ***Articles A, B*** and ***C*** use physical objects as visual stimuli. Although physical objects permit interactions that are close to what we experience on a daily basis, it is difficult to obtain their optical properties. This limits our ability to conduct quantitative modeling between the physical and perceptual properties, and permits drawing just qualitative conclusions. Besides, objects come with unintended artifacts and are subject to aging effects, which makes it impossible to reproduce the experiments over time. For example, we detected a noticeable change in color of the *Plastique* collection objects due to lengthy exposure to illumination. Maloney and Knoblauch (2020) note that the experiments involving physical objects usually take longer due to the time needed to manually substitute samples from trial to trial. This was especially true for ***Articles A*** and ***B***, which might have caused a loss of concentration among observers.

On the other hand, ***Articles D, E, F*** and ***G*** use images for a psychophysical study. They lack many cues which have been observed to be important in ***Article B***, such as binocular vision, scene dynamics, tactile information and interactivity. In addition to this, all visual stimuli are limited with the specification of the monitor, such as color gamut and dynamic range, which are usually smaller than in real life. Although physically-based rendering made it possible to generate photorealistic stimuli with full control over optical material properties and scene composition (***Articles D, E, F***), Chadwick et al. (2018) have observed that people's performance on synthetic stimuli is not as good as on real ones, proposing that the HVS might be trained on the materials and objects that actually exist around us. The latter problem can be solved by using photographs (***Article G***), but they do not contain the information regarding the optical material properties, limiting the analysis to image statistics extraction (such as in ***Article H***). Finally, each stimulus, either virtual or real, is also limited by the shape it is presented in, since conclusions drawn in our works (***Article A, B*** and ***D***) might not be generalized to other shapes and objects.

4.3.4 Online and physical experiments come with their shortcomings

Conducting an experiment either online or physically also brings additional limitations to the studies. Experiments reported in ***Articles A, B, C, E*** and ***G*** have been conducted with a physical presence of an experimenter and an observer, while those in ***Articles D*** and ***F*** have been conducted online. Online studies enable collecting larger amounts of data in a significantly shorter period of time. For instance, data collection for ***Article B*** took 3 months, while for ***Article D*** it was essentially collected overnight. The diversity of the observers is usually larger in online studies and encompasses a broader part of the general populace. It is difficult to ensure observer diversity in experiments conducted physically. For instance, the observer pool in ***Articles A, B, E*** and ***G*** have been mostly composed of the colleagues from the Norwegian Colour and Visual Computing Laboratory, while the experiment in ***Article C*** was conducted at relevant academic conferences with most observers having expertise in the field. Therefore, the findings of these works might poorly generalize to non-experts and the general populace.

On the other hand, conducting studies with physical presence is advantageous in several ways:

- First of all, unlike online studies, the experimenter has control over observation conditions.
- Physical experiments give the experimenter a choice to select either physical or displayed stimuli, while online studies are limited to displayed stimuli only.
- The information collected about observers is more reliable - e.g. the experimenter can test visual acuity and color vision of the observers.
- We noticed that oftentimes observers need additional clarifications about the task and the instructions, which is usually possible with the physical presence of the experimenter. This is especially true when studying *translucency*, because the term is inherently vague.
- The experimenter can ask observers to reflect on the task and obtain comprehensive explanations for observer's responses when both are physically present.
- We observed that data obtained online was noisier and the overall dedication of the observers was worse. For instance, the experiments conducted on the Amazon Mechanical Turk involved many seemingly random clicks.

4.3.5 The data can be noisy

We also cannot rule out the existence of unintended noise in the data. We have observed in ***Article F*** that the sequence of the comparisons in the course of the experiment might have affected the results in an unintended way. There might be this kind of noise from unidentified sources in the data, especially in the experiments which permitted improvisation by the experimenter (***Articles A-C***).

4.3.6 Semantic communication had to be explored further

Finally, semantic communication has been explored to a very little extent (although see ***supporting Article O***). We observed in the original study that semantic description and communication are essential parts of the appearance assessment process. Understanding the ways to exchange the information about appearance has a considerable economic implication and it ensures the effective communication not only externally, between customers and manufacturers, but also internally, within industrial and academic communities. While this study advanced our knowledge on perception, we believe it could have also explored more on how people express and convey what they perceive. The first step towards this objective can be clearing the ambiguous definitions up.

Chapter 5

Conclusions

We have initially conducted inductive research in order to observe the behavioral traits of material appearance assessment process and to formulate relevant research hypotheses. The study has revealed that the comparison with a relevant reference is at the core of the appearance assessment and multisensory information, motion and scene dynamics are extensively used, making interaction with the objects an important part of the assessment process. Afterwards, we focused on the appearance of translucent materials, as translucent materials represent an important subset of materials we encounter on a daily basis, but are yet mostly understudied. We tested interesting research hypotheses using deductive research methods and found ample evidence that:

- Translucency impacts glossiness perception; surface reflectance distribution functions cannot adequately specify perceived glossiness and subsurface scattering properties need also to be taken into consideration.
- The constancy of translucency appearance is limited with cross-shape and observation condition variations.
- Thin parts facilitate detection of suprathreshold translucency differences when apparent translucency of two materials is compared.
- Translucency and transparency perception cues are essentially different and humans are more sensitive to subsurface scattering changes when the background is seen through the object.
- Caustics encapsulate important information about material translucence and they contribute to the magnitude of perceived translucency.
- Decreasing luminance contrast increases the magnitude of perceived translucency only when specular highlights are kept intact, while blurring the entire image including specular highlights decreases the magnitude of perceived translucency.
- The statistics of the luminance histogram can reflect gloss and translucency properties, but being subject to numerous photo-geometric and environmental constraints, they alone are not reliable predictors of appearance.

In the course of the experiments, we faced substantial challenges due to conceptual ambiguity of translucency and highlight the need for standardization. However, we understand that universal definitions and measurement standards might not be feasible and they could be limited to specific contexts, applications and industries.

Finally, a comprehensive analysis of the state-of-the-art made us conclude that translucency perception research is in the initial stage of its development. An extensive amount of future work is needed to bring those mechanisms to light that are responsible for perceiving translucency. The abundance of translucent objects and materials in our daily lives makes this question economically relevant for a broad range of industries.

Chapter 6

Future Work

Varying distinct optical properties in a systematic manner and measuring how they impact observers' responses provides little understanding of how the HVS functions and what are the actual mechanisms of translucency perception. Although multimodal information contributes to material appearance, we believe the essential portion of the information is encapsulated in a 2D retinal image. The fundamental problem is to identify how the HVS uses and weights image intensities in order to deduce subsurface light transport and surface reflectance properties. Nagai et al. (2013) found that instead of relying on global statistics, the HVS judges translucency based on local informative "hot spots". It was earlier proposed that such regions are usually edges (Fleming and Bülthoff (2005)), but the exact way the color (both intensity and chromaticity), spatial and temporal information is used by the HVS is yet to be understood. We propose that for advancing translucency perception research, and research on the perception of appearance in general, eye tracking experiments should be conducted. Eye tracking will reveal which regions impact observers' decisions. It is especially interesting to conduct it in dynamic scenes, where either the object or the background is in motion. This will help us determine *why* the stimuli differ in translucency and *how* optical and environmental parameters modulate the image cues and thus, the magnitude of perceived translucency. For instance, identifying the regions that contribute to translucency perception will help us construct respective shape descriptors to adequately model the impact of shape on translucency appearance. Additionally, eye tracking will also reveal whether observers actually rely on caustics or other cues located elsewhere in the scene, outside the object's body.

Secondly, machines could assist with the extraction of the relevant image features. For instance, it has been demonstrated recently that unsupervised machine learning techniques outperform image statistics in the prediction of human perception (Storrs and Fleming (2020)). This can be a promising avenue for translucency and in general, appearance research. Extracting perceptually meaningful features using machine learning techniques might provide a deeper insight into the humanly mechanisms of perception than simple handcrafted image metrics.

Thirdly, it is also important to explore how translucency interacts with other appearance attributes and what is the role of cognitive prior, such as material identification and expectations, in translucency perception, which has been demonstrated to be an important factor in appearance in general (Alley et al. (2020)). We believe that a sophisticated and perplexing mechanism of translucency perception cannot be elucidated by psychophysics and image analysis alone without contribution from modern neuroscience. We expect that the research on translucency perception can greatly benefit from studies similar to that by Chadwick et al. (2019). A neuroscientific study should reveal whether the perception of translucency and other attributes are anatomically independent, and in general, which cortical areas are responsible for perceiving translucency. For instance, it is greatly anticipated that translucency perception is interrelated with the perception of shape (Marlow et al. (2017), Chowdhury et al. (2017), and Xiao et al. (2020)). Understanding the physiology of translucency perception on the retinal and cortical levels could aid the definition of a standard observer for translucency.

Besides, it is of particular interest to explore to what extent the HVS can separate surface and subsurface scattering and whether it is feasible to produce translucency metamers with distinct surface and subsurface scattering effects - and if so, in which light field should the object be embedded for this effect. While separation of the two might be easier for smooth, specularly reflecting objects, as observed in ***Article F***, the task can become increasingly difficult with the increase of surface roughness. This research question can have significant economic relevance, as surface manipulation is oftentimes cheaper than that of subsurface scattering properties.

Furthermore, we think that information encapsulated in caustics deserves further attention. Future work should explore to what extent can object and material properties be estimated from a caustic pattern. A cheap and simple image-based measurement technique can be developed, if reliable links are found between caustics and material properties. This measurement technique, however, can be limited with caustic metamers, i.e. different objects and materials producing identical caustics, the potential existence of which is an interesting question itself.

Apart from that, we have discussed a broad range of shortcomings that are associated with the usage of still images. However, the generation of large physical object datasets remains economically inefficient, as well as inconvenient in terms of data sharing and research reproducibility and replicability. We contemplate that future works can find a trade-off using emerging technologies to present stimuli, such as extended reality and programmable matters.

To summarize, the research conducted by us and the open points outlined above show that neither computer science, nor the vision science community is likely to solve the appearance-related problems alone, but rather a multidisciplinary effort and different ways of thinking are needed. Appearance as a concept does not belong to any particular domain and advances in our understanding of it require input from the vision, computer and material science communities, as well as from the visual arts, social science, experimental and cognitive psychology research. We foresee that the key is in interdisciplinary research on appearance.

Bibliography

Alley, Lorilei M, Alexandra C Schmid, and Katja Doerschner (2020). “Expectations affect the perception of material properties.” In: *Journal of Vision* 20.12:1, pp. 1–20.

Amirkhanashvili, Ana and Davit Gigilashvili (2020). “Color Naming and Communication of Color Appearance: Is it Different for Native Georgian Speakers?” In: *10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*. Vol. 2688. CEUR Workshop Proceedings, pp. 1–15.

Anderson, Barton L (2011). “Visual perception of materials and surfaces.” In: *Current biology* 21.24, R978–R983.

Anderson, Barton L and Juno Kim (2009). “Image statistics do not explain the perception of gloss and lightness.” In: *Journal of Vision* 9.11:10, pp. 1–17.

ASTM D 1003 (2003). “Standard Test Method for Haze and Luminous Transmittance of Transparent Plastics.” In: American Society for Testing and Materials. West Conshohocken, PA.

ASTM E2387-05 (2011). “Standard Practice for Goniometric Optical Scatter Measurements. ASTM International Standard.” In: American Society for Testing and Materials. West Conshohocken, PA.

ASTM E284-17 (2017). “Standard Terminology of Appearance.” In: ASTM International, West Conshohocken, PA. URL: `https://doi.org/10.1520/E0284-17`.

Barry, Fabio (2011). “Painting in stone: The symbolism of colored marbles in the visual arts and literature from antiquity until the Enlightenment.” PhD thesis. Columbia University.

Beck, Jacob and Richard Ivry (1988). “On the role of figural organization perceptual transparency.” In: *Perception & psychophysics* 44.6, pp. 585–594.

Beck, Jacob, K Prazdny, and Richard Ivry (1984). “The perception of transparency with achromatic colors.” In: *Perception & psychophysics* 35.5, pp. 407–422.

Beck, Jacob and Slava Prazdny (1981). “Highlights and the perception of glossiness.” In: *Perception & Psychophysics*.

Brainard, David H and Anya C Hurlbert (2015). “Colour vision: understanding #TheDress.” In: *Current Biology* 25.13, R551–R554.

Brunton, Alan, Can A Arikan, Tejas M Tanksale, and Philipp Urban (2018). “3D Printing Spatially Varying Color and Translucency.” In: *ACM Transactions on Graphics (TOG)* 37.4, 157:1–157:13.

BYK Gardner GmbH. Haze-gard Transparency Transmission Haze Meter (n.d.). Accessed on 16/10/20 at: `https://www.byk-instruments.com/us/en/Appearance/haze-gard-Transparency-Transmission-Haze-Meter/c/2345`.

Cavanagh, Patrick (2005). "The artist as neuroscientist." In: *Nature* 434.7031, pp. 301–307.

Chadwick, Alice C, George Cox, Hannah E Smithson, and Robert W Kentridge (2018). "Beyond scattering and absorption: Perceptual unmixing of translucent liquids." In: *Journal of Vision* 18.11:18, pp. 1–15.

Chadwick, Alice C, Charles A Heywood, Hannah E Smithson, and Robert W Kentridge (2019). "Translucence perception is not dependent on cortical areas critical for processing colour or texture." In: *Neuropsychologia* 128, pp. 209–214.

Chadwick, Alice C and Robert W Kentridge (2015). "The perception of gloss: A review." In: *Vision research* 109, pp. 221–235.

Chandrasekhar, Subrahmanyan (1960). "Radiative Transfer." In: Dover Publications Inc. New York, pp. 1–53.

Cheeseman, Jacob R, James A Ferwerda, Frank J Maile, and Roland W Fleming (2021). "Scaling and discriminability of perceived gloss." In: *JOSA A* 38.2, pp. 203–210.

Choudhury, Asim Kumar Roy (2014). *Principles of colour and appearance measurement: Object appearance, colour perception and instrumental measurement*. Elsevier.

Chowdhury, Nahian S, Phillip J Marlow, and Juno Kim (2017). "Translucency and the perception of shape." In: *Journal of Vision* 17.3:17, pp. 1–14.

CIE (1987). *CIE 17.4:1987 International Lighting Vocabulary*. International Commission on Illumination.

CIE (2006). *CIE 175:2006 A framework for the measurement of visual appearance*. International Commission on Illumination. ISBN: 978 3 901906 52 7, 92 pages.

Corbin, Juliet and Anselm Strauss (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications. 4th Edition. Thousand Oaks, CA.

Di Cicco, Francesca, Lisa Wiersma, Maarten Wijntjes, and Sylvia Pont (2020a). "Material properties and image cues for convincing grapes: The know-how of the 17th-century pictorial recipe by Willem Beurs." In: *Art & Perception* 8.3-4, pp. 337–362.

Di Cicco, Francesca, Maarten WA Wijntjes, and Sylvia C Pont (2020b). "If painters give you lemons, squeeze the knowledge out of them. A study on the visual perception of the translucent and juicy appearance of citrus fruits in paintings." In: *Journal of Vision* 20.13:12, pp. 1–15.

Doerschner, Katja, Roland W Fleming, Ozgur Yilmaz, Paul R Schrater, Bruce Hartung, and Daniel Kersten (2011). "Visual motion and the perception of surface material." In: *Current Biology* 21.23, pp. 2010–2016.

Dorsey, Julie, Holly Rushmeier, and François Sillion (2010). *Digital modeling of material appearance*. Elsevier.

Emmert, Ralf (1996). “Quantification of the soft-focus effect: Measuring light-diffusing characteristics of cosmetic pigments and powders.” In: *Cosmetics and toiletries* 111.7, pp. 57–61.

Engeldrum, Peter G (2000). *Psychometric scaling: a toolkit for imaging systems development*. Imcotek.

Eugène, Christian (2008). “Measurement of "total visual appearance": a CIE challenge of soft metrology.” In: *12th IMEKO TC1 TC7 Joint Symposium on Man, Science Measurement*, pp. 61–65.

Faul, Franz (2019). “The influence of Fresnel effects on gloss perception.” In: *Journal of Vision* 19.13:1, pp. 1–39.

Faul, Franz and Vebjørn Ekroll (2002). “Psychophysical model of chromatic perceptual transparency based on substractive color mixture.” In: *JOSA A* 19.6, pp. 1084–1095.

Faul, Franz and Vebjørn Ekroll (2011). “On the filter approach to perceptual transparency.” In: *Journal of Vision* 11.7:7, pp. 1–33.

Fechner, Gustav Theodor, Davis H Howes, and Edwin Garrigues Boring (1966). *Elements of Psychophysics*. Vol. 1. Holt, Rinehart and Winston New York.

Filip, Jiřı, Martina Kolafová, Michal Havlıček, Radomır Vávra, Michal Haindl, and Holly Rushmeier (2018). “Evaluating physical and rendered material appearance.” In: *The Visual Computer* 34.6-8, pp. 805–816.

Fleming, Roland W (2014). “Visual perception of materials and their properties.” In: *Vision research* 94, pp. 62–75.

Fleming, Roland W and Heinrich H Bülthoff (2005). “Low-level image cues in the perception of translucent materials.” In: *ACM Transactions on Applied Perception (TAP)* 2.3, pp. 346–382.

Fleming, Roland W, Ron O Dror, and Edward H Adelson (2003). “Real-world illumination and the perception of surface reflectance properties.” In: *Journal of Vision* 3, pp. 347–368.

Frisvad, Jeppe Revall, Søren Alkærsig Jensen, Jonas Skovlund Madsen, António Correia, Li Yang, Søren KS Gregersen, Youri Meuret, and Poul-Erik Hansen (2020). “Survey of Models for Acquiring the Optical Properties of Translucent Materials.” In: *State of The Art Report, Eurographics 2020* 39.2, pp. 729–755.

Fujisaki, Waka, Naokazu Goda, Isamu Motoyoshi, Hidehiko Komatsu, and Shin’ya Nishida (2014). “Audiovisual integration in the human perception of materials.” In: *Journal of Vision* 14.4:12, pp. 1–20.

Gaucher, Nathalie and Antoine Payot (2011). “From powerlessness to empowerment: Mothers expect more than information from the prenatal consultation for preterm labour.” In: *Paediatrics & Child Health* 16.10, pp. 638–642.

Ged, Guillaume, Gaël Obein, Zaccaria Silvestri, Jean Le Rohellec, and Françoise Viénot (2010). “Recognizing real materials from their glossy appearance.” In: *Journal of vision* 10.9:18, pp. 1–17.

Gerbino, Walter, Casimir I Stultiens, Jim M Troost, and Charles M de Weert (1990). “Transparent layer constancy.” In: *Journal of Experimental Psychology: Human Perception and Performance* 16.1, pp. 3–20.

Giancola, Giorgiana and Mitchell L Schlossman (2015). “Decorative Cosmetics.” In: *Cosmeceuticals and Active Cosmetics*, pp. 191–219.

Gigilashvili, Davit, Lucas Dubouchet, Marius Pedersen, and Jon Yngve Hardeberg (2020). “Caustics and Translucency Perception.” In: *Material Appearance 2020, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 033:1–033:6.

Gigilashvili, Davit, Fereshteh Mirjalili, and Jon Yngve Hardeberg (2019). “Illuminance Impacts Opacity Perception of Textile Materials.” In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 126–131.

Gigilashvili, Davit, Marius Pedersen, and Jon Yngve Hardeberg (2018). “Blurring impairs translucency perception.” In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 377–382.

Gigilashvili, Davit, Weiqi Shi, Zeyu Wang, Marius Pedersen, Jon Yngve Hardeberg, and Holly Rushmeier (2021). “The Role of Subsurface Scattering in Glossiness Perception.” In: *ACM Transaction on Applied Perception* 18.3, 10:1–10:26.

Gigilashvili, Davit, Midori Tanaka, Marius Pedersen, and Jon Yngve Hardeberg (2020). “Image Statistics as Glossiness and Translucency Predictor in Photographs of Real-world Objects.” In: *10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*. Vol. 2688. CEUR Workshop Proceedings, pp. 1–15.

Gigilashvili, Davit, Jean Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2020). “On the Nature of Perceptual Translucency.” In: *8th Annual Workshop on Material Appearance Modeling (MAM2020). Eurographics Digital Library*, pp. 17–20.

Gigilashvili, Davit, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2018). “Behavioral investigation of visual appearance assessment.” In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 294–299.

Gigilashvili, Davit, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (n.d.). “Translucency perception: A review.” In: *Accepted for publication in the Journal of Vision*, 45 pages.

Gigilashvili, Davit, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (2019). “Material appearance: ordering and clustering.” In: *Material Appearance 2019, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 202:1–202:6.

Gigilashvili, Davit, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (2019). “Perceived Glossiness: Beyond Surface Properties.” In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 37–42.

Gigilashvili, Davit, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). “On the appearance of objects and materials: Qualitative analysis of experimental observations." In: *Accepted for publication in the Journal of the International Colour Association (JAIC)*, 33 pages.

Gigilashvili, Davit, Philipp Urban, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2019). "Impact of Shape on Apparent Translucency Differences." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 132–137.

Gigilashvili, Davit, Philipp Urban, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). "The Impact of Optical and Geometrical Thickness on Perceived Translucency Differences." In: *Under review in a journal*, 13 pages.

Gkioulekas, Ioannis, Bruce Walter, Edward H Adelson, Kavita Bala, and Todd Zickler (2015). "On the appearance of translucent edges." In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5528–5536.

Gkioulekas, Ioannis, Bei Xiao, Shuang Zhao, Edward H Adelson, Todd Zickler, and Kavita Bala (2013). "Understanding the role of phase function in translucent appearance." In: *ACM Transactions on graphics (TOG)* 32.5, pp. 1–19.

Gkioulekas, Ioannis, Shuang Zhao, Kavita Bala, Todd Zickler, and Anat Levin (2013). "Inverse volume rendering with material dictionaries." In: *ACM Transactions on Graphics (TOG)* 32.6, pp. 1–13.

Glaser, Barney G and Anselm Strauss (1965). "Discovery of substantive theory: A basic strategy underlying qualitative research." In: *American Behavioral Scientist* 8.6, pp. 5–12.

Glaser, Barney G, Anselm L Strauss, and Elizabeth Strutzel (1968). "The discovery of grounded theory; strategies for qualitative research." In: *Nursing research* 17.4, p. 364.

Goesele, Michael, Hendrik PA Lensch, Jochen Lang, Christian Fuchs, and Hans-Peter Seidel (2004). "DISCO: acquisition of translucent objects." In: *ACM SIGGRAPH 2004 Papers*, pp. 835–844.

Hodgson, Alan (2020). "The viewing of Caustics." In: *The Royal Photographic Society*. Accessed on 17/02/2021 at https://rps.org/news/bristol/2020/october/the-viewing-of-caustics/.

Hunter, Richard S (1937). "Methods of determining gloss." In: *NBS Research paper RP* 958, pp. 19–39.

Hunter, Richard S and Richard W Harold (1987). *The measurement of appearance*. John Wiley & Sons.

Hutchings, John B (1977). "The importance of visual appearance of foods to the food processor and the consumer 1." In: *Journal of Food Quality* 1.3, pp. 267–278.

Hutchings, John B (1999). *Food color and appearance*. 2nd edition. Aspen Publishers, New York.

Hutchings, John B (2011). *Food colour and appearance*. Springer Science & Business Media.

Hutchings, John B (1995a). "The continuity of colour, design, art, and science. I. The philosophy of the total appearance concept and image measurement." In: *Color Research & Application* 20.5, pp. 296–306.

Hutchings, John B (1995b). "The continuity of colour, design, art, and science. II. Application of the total appearance concept to image creation." In: *Color Research & Application* 20.5, pp. 307–312.

Jacob, Jean Daniel and Dave Holmes (2011). "Working under threat: Fear and nurse–patient interactions in a forensic psychiatric setting." In: *Journal of Forensic Nursing* 7.2, pp. 68–77.

Jakob, Wenzel (2010). *Mitsuba renderer*. http://www.mitsuba-renderer.org.

Jensen, Henrik W, Stephen R Marschner, Marc Levoy, and Pat Hanrahan (2001). "A practical model for subsurface light transport." In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 511–518.

Kaltenbach, Frank (2012). *Translucent materials: glass, plastics, metals*. Walter de Gruyter.

Kerrigan, Iona S and Wendy J Adams (2013). "Highlights, disparity, and perceived gloss with convex and concave surfaces." In: *Journal of Vision* 13.1:9, pp. 1–10.

Kim, Juno, Phillip Marlow, and Barton L Anderson (2011). "The perception of gloss depends on highlight congruence with surface shading." In: *Journal of Vision* 11(9).4, pp. 1–19.

Lafer-Sousa, Rosa, Katherine L Hermann, and Bevil R Conway (2015). "Striking individual differences in color perception uncovered by 'the dress' photograph." In: *Current Biology* 25.13, R545–R546.

Lagunas, Manuel, Sandra Malpica, Ana Serrano, Elena Garces, Diego Gutierrez, and Belen Masia (2019). "A similarity measure for material appearance." In: *arXiv preprint arXiv:1905.01562*.

Landy, Michael S (2007). "A gloss on surface properties." In: *Nature* 447.7141, pp. 158–159.

Leloup, Frédéric B, Stefaan Forment, Philip Dutré, Michael R Pointer, and Peter Hanselaer (2008). "Design of an instrument for measuring the spectral bidirectional scatter distribution function." In: *Applied optics* 47.29, pp. 5454–5467.

Leloup, Frédéric B, Gael Obein, Michael R Pointer, and Peter Hanselaer (2014). "Toward the soft metrology of surface gloss: A review." In: *Color Research & Application* 39.6, pp. 559–570.

Leloup, Frédéric B, Michael R Pointer, Philip Dutré, and Peter Hanselaer (2010). "Geometry of illumination, luminance contrast, and gloss perception." In: *JOSA A* 27.9, pp. 2046–2054.

Leloup, Frédéric B, Michael R Pointer, Philip Dutré, and Peter Hanselaer (2011). "Luminance-based specular gloss characterization." In: *JOSA A* 28.6, pp. 1322–1330.

Leloup, Frédéric B, Michael R Pointer, Philip Dutré, and Peter Hanselaer (2012). "Overall gloss evaluation in the presence of multiple cues to surface glossiness." In: *JOSA A* 29.6, pp. 1105–1114.

Liu, Min-Chieh, Steven A Aquilino, Peter S Lund, Marcos A Vargas, Ana M Diaz-Arnold, David G Gratton, and Fang Qian (2010). "Human perception of dental porcelain translucency correlated to spectrophotometric measurements." In: *Journal of Prosthodontics: Implant, Esthetic and Reconstructive Dentistry* 19.3, pp. 187–193.

Lopes Filho, Hibernon, Lúcio EG Maia, Marcus Vinicius A Araújo, and Antônio Carlos O Ruellas (2012). "Influence of optical properties of esthetic brackets (color, translucence, and fluorescence) on visual perception." In: *American journal of orthodontics and dentofacial orthopedics* 141.4, pp. 460–467.

Luckerhoff, Jason and François Guillemette (2011). "The Conflicts between Grounded Theory Requirements and Institutional Requirements for Scientific Research." In: *Qualitative Report* 16.2, pp. 396–414.

Maloney, Laurence T and Kenneth Knoblauch (2020). "Measuring and Modeling Visual Appearance." In: *Annual Review of Vision Science* 6, pp. 519–537.

Marlow, Phillip J and Barton L Anderson (2013). "Generative constraints on image cues for perceived gloss." In: *Journal of Vision* 13.14, pp. 2–2.

Marlow, Phillip J and Barton L Anderson (2021). "The cospecification of the shape and material properties of light permeable materials." In: *Proceedings of the National Academy of Sciences* 118.14, pp. 1–10.

Marlow, Phillip J, Juno Kim, and Barton L Anderson (2011). "The role of brightness and orientation congruence in the perception of surface gloss." In: *Journal of Vision* 11(9).16, pp. 1–12.

Marlow, Phillip J, Juno Kim, and Barton L Anderson (2012). "The perception and misperception of specular surface reflectance." In: *Current Biology* 22.20, pp. 1909–1913.

Marlow, Phillip J, Juno Kim, and Barton L Anderson (2017). "Perception and misperception of surface opacity." In: *Proceedings of the National Academy of Sciences* 114.52, pp. 13840–13845.

Merriam-Webster Dictionary (n.d.). Accessed: 2020-11-06. URL: `https://www.merriam-webster.com/dictionary/translucent`.

Metelli, Fabio (1974). "The perception of transparency." In: *Scientific American* 230.4, pp. 90–99.

Molloy, Mark (2016). "Confused by this shiny leg optical illusion? Here's how it works." In: *The Telegraph*. Accessed on 04/01/2019. URL: `https://www.telegraph.co.uk/news/2016/10/31/confused-by-this-shiny-leg-optical-illusion-heres-how-it-works`.

Motoyoshi, Isamu (2010). "Highlight–shading relationship as a cue for the perception of translucent and transparent materials." In: *Journal of Vision* 10.9:6, pp. 1–11.

Motoyoshi, Isamu, Shin'ya Nishida, Lavanya Sharan, and Edward H Adelson (2007). "Image statistics and the perception of surface qualities." In: *Nature* 447.7141, pp. 206–209.

Murray, Scott (2013). *Translucent building skins: material innovations in modern and contemporary architecture*. Routledge.

Nagai, Takehiro, Yuki Ono, Yusuke Tani, Kowa Koida, Michiteru Kitazaki, and Shigeki Nakauchi (2013). "Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study." In: *i-Perception* 4.6, pp. 407–428.

Nicodemus, Fred E (1965). "Directional reflectance and emissivity of an opaque surface." In: *Applied optics* 4.7, pp. 767–775.

Nicodemus, Fred E, Joseph C Richmond, Jack J Hsia, Irving W Ginsberg, Thomas Limperis, et al. (1977). *Geometrical considerations and nomenclature for reflectance*. Vol. 160. Citeseer, 52 pages.

Nishida, Shin'ya, Isamu Motoyoshi, Lisa Nakano, Yuanzhen Li, Lavanya Sharan, and Edward Adelson (2008). "Do colored highlights look like highlights?" In: *Journal of Vision* 8.6, p. 339.

Nishida, Shin'ya and Mikio Shinya (1998). "Use of image-based information in judgments of surface-reflectance properties." In: *JOSA A* 15.12, pp. 2951–2965.

Nunes, Augusto LP, Anderson Maciel, Gary W Meyer, Nigel W John, Gladimir VG Baranoski, and Marcelo Walter (2019). "Appearance modelling of living human tissues." In: *Computer Graphics Forum*. Vol. 38. 6. Wiley Online Library, pp. 43–65.

Obein, Gaël, Kenneth Knoblauch, and Françoise Viénot (2004). "Difference scaling of gloss: Nonlinearity, binocularity, and constancy." In: *Journal of Vision* 4, pp. 711–720.

Olkkonen, Maria and David H Brainard (2011). "Joint effects of illumination geometry and object shape in the perception of surface reflectance." In: *i-Perception* 2.9, pp. 1014–1034.

Paillé, Pierre (1994). "L'analyse par théorisation ancrée." In: *Cahiers de recherche sociologique* 1.23, pp. 147–181.

Pellacini, Fabio, James A Ferwerda, and Donald P Greenberg (2000). "Toward a psychophysically-based light reflection model for image synthesis." In: *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., pp. 55–64.

Piadyk, Yurii, Yitzchak Lockerman, and Claudio Silva (2020). "Anisotropic Subsurface Scattering Acquisition Through a Light Field Based Apparatus." In: *Electronic Imaging, Imaging Sensors and Systems 2020* 7, 225:1–225:7.

Pointer, Michael R (2003). "Measuring Visual Appearance- A Framework of the Future. Project 2.3 Measurement of Appearance." In: *National Physical Laboratory (NPL) Report: Centre for Optical and Analytical Measurement (COAM) 19*.

Rippon, Daniel, Andrew McDonnell, Michael Smith, Michael McCreadie, and Mark Wetherell (2020). "A grounded theory study on work related stress in professionals who provide health & social care for people who exhibit behaviours that challenge." In: *PLoS ONE* 15.2, pp. 1–23.

Rushmeier, Holly (1995). "Input for participating media." In: *In Realistic Input for Realistic Images (1995), ACM Press, ACM SIGGRAPH '95 Course Notes. Also appeared in the ACM SIGGRAPH '98 Course Notes - A Basic Guide to Global Illumination.* 8:1–8:24.

Sakano, Yuichi and Hiroshi Ando (2010). "Effects of head motion and stereo viewing on perceived glossiness." In: *Journal of Vision* 10.9:15, pp. 1–14.

Sawayama, Masataka, Yoshinori Dobashi, Makoto Okabe, Kenchi Hosokawa, Takuya Koumura, Toni Saarela, Maria Olkkonen, and Shin'ya Nishida (2019). "Visual discrimination of optical material properties: a large-scale study." In: *BioRxiv*, 35 pages.

Serrano, Ana, Diego Gutierrez, Karol Myszkowski, Hans-Peter Seidel, and Belen Masia (2018). "An intuitive control space for material appearance." In: *arXiv preprint arXiv:1806.04950*.

Sharan, Lavanya, Ruth Rosenholtz, and Edward H Adelson (2014). "Accuracy and speed of material categorization in real-world images." In: *Journal of Vision* 14.9:12, pp. 1–24.

Singh, Manish and Barton L Anderson (2002). "Perceptual assignment of opacity to translucent surfaces: The role of image blur." In: *Perception* 31.5, pp. 531–552.

Sole, Aditya, Davit Gigilashvili, Helene Midtfjord, Dar'ya Guarnera, Giuseppe Claudio Guarnera, Jean-Baptiste Thomas, and Jon Yngve Hardeberg (2019). "On the acquisition and reproduction of material appearance." In: *International Workshop on Computational Color Imaging*. Springer, pp. 26–38.

Starrin, Bengt, Lars Dahlgren, Gerry Larsson, and Sven Styrborn (1997). *Along the path of discovery: Qualitative methods and grounded theory*. ISBN: 9144005156. Lund, Sweden: Studentlitteratur.

Stevens, Stanley S (1960). "The psychophysics of sensory function." In: *American Scientist* 48.2, pp. 226–253.

Storrs, Katherine R and Roland W Fleming (2020). "Unsupervised Learning Predicts Human Perception and Misperception of Specular Surface Reflectance." In: *bioRxiv*, 25 pages.

Tamura, Hideki, Hiroshi Higashi, and Shigeki Nakauchi (2018). "Dynamic visual cues for differentiating mirror and glass." In: *Scientific reports* 8.1, pp. 1–12.

The Stanford 3D Scanning Repository (1994). Stanford University Computer Graphics Laboratory. URL: `http://graphics.stanford.edu/data/3Dscanrep/`.

Thomas, Jean-Baptiste, Aurore Deniel, and Jon Y Hardeberg (2018). "The Plastique collection: A set of resin objects for material appearance research." In: *XIV Conferenza del Colore, Florence, Italy*, 12 pages.

Thomas, Jean-Baptiste, Jon Yngve Hardeberg, and Gabriele Simone (2017). "Image contrast measure as a gloss material descriptor." In: *International Workshop on Computational Color Imaging*. Springer, pp. 233–245.

Torgerson, Warren S (1958). "Theory and methods of scaling." In: 1958, Wiley: New York.

Toscani, Matteo, Dar'ya Guarnera, Giuseppe Claudio Guarnera, Jon Yngve Hardeberg, and Karl R Gegenfurtner (2020). "Three perceptual dimensions for specular and diffuse reflection." In: *ACM Transactions on Applied Perception (TAP)* 17.2, pp. 1–26.

Urban, Philipp, Tejas Madan Tanksale, Alan Brunton, Bui Minh Vu, and Shigeki Nakauchi (2019). "Redefining a in RGBA: Towards a standard for graphical 3D printing." In: *ACM Transactions on Graphics (TOG)* 38.3, pp. 1–14.

Van Ngo, Khai, Jehans Jr. Storvik, Christopher André Dokkeberg, Ivar Farup, and Marius Pedersen (2015). "Quickeval: a web application for psychometric scaling experiments." In: *Image Quality and System Performance XII*. Vol. 9396. International Society for Optics and Photonics, 93960O.

Vangorp, Peter, Jurgen Laurijssen, and Philip Dutré (2007). "The influence of shape on the perception of material reflectance." In: *ACM Transactions on graphics (TOG)*. Vol. 26. 3. ACM, 77:1–77:10.

Wendt, Gunnar, Franz Faul, Vebjørn Ekroll, and Rainer Mausfeld (2010). "Disparity, motion, and color information improve gloss constancy performance." In: *Journal of Vision* 10.9:7, pp. 1–17.

Wendt, Gunnar, Franz Faul, and Rainer Mausfeld (2008). "Highlight disparity contributes to the authenticity and strength of perceived glossiness." In: *Journal of Vision* 8.1:14, pp. 1–10.

Wiebel, Christiane B, Matteo Toscani, and Karl R Gegenfurtner (2015). "Statistical correlates of perceived gloss in natural images." In: *Vision Research* 115, pp. 175–187.

Wijntjes, MWA, C Spoiala, and H de Ridder (2020). "Thurstonian Scaling and the Perception of Painterly Translucency." In: *Art & Perception* 8.3–4, pp. 363–386.

Xiao, Bei, Bruce Walter, Ioannis Gkioulekas, Todd Zickler, Edward Adelson, and Kavita Bala (2014). "Looking against the light: How perception of translucency depends on lighting direction." In: *Journal of Vision* 14.3:17, pp. 1–22.

Xiao, Bei, Shuang Zhao, Ioannis Gkioulekas, Wenyan Bi, and Kavita Bala (2020). "Effect of geometric sharpness on translucent material perception." In: *Journal of Vision* 20.7:10, pp. 1–17.

Part II

Article A

Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2018). “Behavioral investigation of visual appearance assessment.” In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 294–299

Behavioral Investigation of Visual Appearance Assessment

Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, Marius Pedersen;
Department of Computer Science, Norwegian University of Science and Technology; Gjøvik, Norway

Abstract

The way people judge, assess and express appearance they perceive can dramatically vary from person to person. The objective of this study is to identify the research hypotheses and outline directions for the future work based on the tasks observers perform. The eventual goal is to understand how people perceive, judge, and assess appearance, and what are the factors impacting their assessments. A series of interviews were conducted in uncontrolled conditions where observers were asked to describe the appearance of the physical objects and to complete simple visual tasks, like ranking objects by their gloss or translucency. The interviews were filmed with the consent of the participants and the videos were subsequently analyzed. The analysis of the data has shown that while there are cross-individual differences and similarities, surface coarseness, shape, and dye mixture have significant effect on translucency and gloss perception.

Introduction and Motivation

Vision is one of the primary senses humans use to perceive and interpret the surrounding. "Visual perception is the ability to interpret the surrounding environment by processing information that is contained in the visible light" [1]. On the other hand, "appearance is the visual sensation through which an object is perceived to have attributes as size, shape, colour, texture, gloss, transparency, opacity etc." [2] Appearance is a complex psychophysical phenomenon that depends not only on the stimuli, but on a broad spectrum of various factors, e.g. memory of the observer [3]. For an easier understanding of appearance, it has been split into several distinct attributes that compose the appearance. CIE defines four major appearance attributes: color, gloss, translucency and texture [2, 4] that interact and influence each other [5, 6, 7].

Advances in computer graphics and simplicity of controlling the parameters have lead to widespread usage of synthetic images for appearance research (e.g. [8, 9]). On the other hand, RGB images of the real objects are frequently used for material appearance analysis, especially in computer vision (e.g. [10, 11, 12]). Despite the clear advantage of using synthetic or real images, the appearance and perception still differ from that of real-life situations. The interaction can be considered less natural due to the presence of the intermediate media and lack of the imperfections in synthetic images [13]. Lack of possibility to touch the objects, limited or no possibility to move them, and lack of the effect of the head movement can be named as further disadvantages of using images for studying appearance.

There has been examples of using real objects for studying appearance [14, 15]. However, experiments were held in controlled laboratory conditions, the observation geometry was fixed and observers were not allowed to touch the objects. This makes the setup artificial and is rarely to be encountered in real life.

Therefore, we decided to use real objects for our study; allowing observers to freely interact with them. The geometry of the measurement can impact the appearance. Bidirectional Reflectance Distribution Functions (BRDF) [16, 17], gloss [18, 19], or color [20, 21] are all measured for predefined geometries. However, observation geometries in real life vary a lot. This is the main reason why we allow the observers to freely interact with the objects. This is primarily a qualitative study to identify traits of appearance assessment by human observers. Analysis of the consistency of human behaviour might potentially outline the directions for further studies, and eventually leading to a better understanding of appearance perception.

The scope of this paper covers the results obtained from the experiments. Particular procedures and processes that lead observers to the results discussed below will be analyzed in the future work. Below we introduce the experimental setup, quantitative results of the experiments followed by the research hypotheses generated from them.

Experimental Setup

In our experiments, we used resin objects of the *Plastique* artwork collection described by Thomas *et al.* [13]. The objects are referred by their codes in task descriptions, as labelled in [13]. The collection of objects is composed of spheres, parallelepipeds, and female bust figures of three levels of surface coarseness and four hues (blue, yellow, white, and achromatic/transparent).

The interviews were conducted in uncontrolled conditions, under a mixture of daylight and artificial fluorescent illumination. The experimenter measured light intensity (in lux) and color temperature of the illumination (in Kelvin) with a light meter before and after the interview. The video and audio of the interview was recorded from two perspectives, front and side. A screenshot from a sample video can be seen in Figure 1. Nine boxes with different sets of the physical objects were used for eleven tasks of the interview (Figure 2). A checkerboard, a pen with text on it, and a white paper were placed on the table without explicit explanations. However, the participants were informed that they could freely interact with the objects. We expected that the white paper, as a homogeneous background, and a checkerboard, as a heterogeneous background, could be used by the observers for judging translucency. Besides, a pen with a text on it could be used to check whether reading through the object was possible. The observers were asked to wear gloves, in order to protect the objects.

17 observers, 11 men and 6 women have been interviewed in total with average age of 35.7 years. 4 out of them were the authors of this paper. 14 observers were experts in the field, while three of them were naïve to visual appearance studies. 2 observers were color deficient.

The interviewees were encouraged to explain their decisions and comment their actions while completing the tasks. The boxes

Figure 1: A screenshot from a sample video.

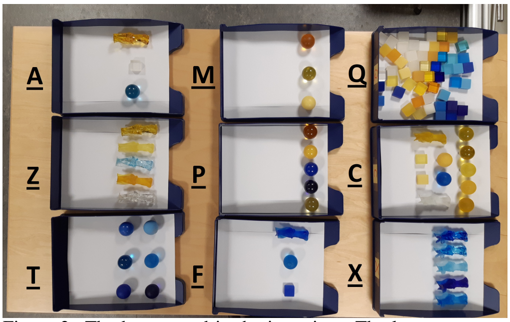

Figure 2: The boxes used in the interview. The letters are randomly assigned to the boxes and are not related to the appearance of the objects.

with respective oral instructions were introduced to the interviewees in the following order:

Task 1 (box Q):

- *Objects*: There are 48 rectangular parallelepipeds of different color, coarseness and translucency in the box.
- *Tasks:* 1) The first task is to cluster the objects into any number of groups the participant considers natural. 2) Experimenter asks the participant to discuss and explain the reasoning of clustering this way. 3) Experimenter asks the participant whether there could be any other way of creating groups that look natural. 4) Experimenter selects one of the groups of the cluster and asks the participant to sub-cluster this group even further.

Task 2 (box C):

- *Objects:* There are 5 yellow spheres of different coarseness and translucency in the box. Besides, there are 6 more objects: two female busts, two spheres and two rectangular parallelepipeds.
- *Tasks:* 1) The first task is to order the 5 spheres in any way the participant considers natural. They are encouraged to use any dimensions they think fit. 2) The participant is given 6 additional objects and is asked to locate the object in relation to the order he/she created with the first five spheres. The observer is expected to fail to order all objects within the created order, and thereby, to generate some questions how to locate the new object. The outcome is to identify potential cues to create an appearance ordering system.

Task 3 (box X):

- *Objects:* There are 5 blue female bust objects from the *Plastique* collection in the box. Object codes: 140, 154, 157, 158, 161.[13]
- *Tasks:* 1) The first task is to describe the appearance of the objects. Besides, the observers are asked, which objects look softer or harder, lighter or heavier, without touching them. 2) The participant can now touch the objects. The participant is asked to rank the object by their gloss/shine.

Task 4 (box M):

- *Objects:* There are 3 yellow spheres of different surface coarseness and translucency in the box. Object codes: 86, 95, 109.
- *Tasks:* 1) The first task is to describe the appearance of the objects with participants' own words. 2) The second task is to rank the object by their gloss/shine.

Task 5 (box P):

- *Objects:* There are 5 spheres of different colors, coarseness and translucency in the box. Object codes: 79, 82, 88, 94, 112.
- *Tasks:* 1) The first task is to describe the appearance of the objects with participants' own words. 2) The second task is to rank the object by their gloss/shine. The goal is to observe, whether difference in color and translucency impacts the result.

Task 6 (box F):

- *Objects:* There are 3 blue objects in the box: one sphere, one rectangular parallelepiped, and one female bust. Object codes: 42, 101, 155.
- *Tasks:* 1) The first task is to describe the appearance of the objects with participants' own words. 2) The second task is to rank the object by their translucency. However, word "translucency" is not be mentioned explicitly throughout the experiment, as it could be ambiguous for some of the interviewees; "how light is going through" is used instead.

Task 7 (box X):

- *Objects:* There are 5 blue female bust objects in the box. Although the box has already been used in the experiment, the experimenter has re-introduced the box in the pile discretely.
- *Tasks:* 1) The first task is to describe the appearance of the objects. 2) The second task is to rank the object by their translucency.

Task 8 (box A):

- *Objects:* There are 3 objects of different shape and color in the box: yellow female bust, achromatic rectangular parallelepiped, and blue sphere. Object codes: 2, 103, 151.
- *Tasks:* 1) The first task is to describe the appearance of the objects. 2) The second task is to rank the object by their translucency. The goal is to observe, whether color and shape impact translucency perception.

Task 9 (box Z):

- *Objects:* There are 5 female bust objects of different colors in the box. Object codes: 115, 152, 160, 163, 167.
- *Tasks:* 1) The first task is to describe the appearance of the objects. Besides, the observers are specially asked, which objects looks softer or harder, heavier or lighter, without touching them. 2) The participant can now touch the objects. The participant is asked to rank the object by their translucency.

Task 10 (box A):

- *Objects:* There are 3 objects of different shape, color, and surface coarseness in the box: yellow female bust, achromatic rectangular parallelepiped, and blue sphere. Although the box has already been used in the experiment, this is not revealed to the participant.

- *Tasks:* 1) The first task is to describe the appearance of the objects. 2) The second task is to rank the object by gloss/shine.

Task 11 (box T):

- *Objects:* There are 6 blue spheres of different surface coarseness and dye mixture in the box. Object codes: 75, 76, 80, 83, 100, 102.

- *Tasks:* 1) The first task is to describe the appearance of the objects. 2) The second task is to cluster them into "opaque" and "non-opaque" categories. We are interested, whether level of light transport is critical for opacity or transparency identification.

The objects used for tasks 3, 4, 5, 7 and 9 are labelled and illustrated on Figure 4.

Analysis and Results

We provide quantitative analysis on 9 boxes, while the first two ones will only be considered in a qualitative way due limited space. Nevertheless, the behavioral patterns and detailed analysis will be considered in a future communication. Behavioral patterns for boxes Q and C are very complex and therefore, left beyond the scope of this paper.

The ranking experiment results are quantified as follows: ranked objects are given points from 5 to 1, where 5 points correspond to the most glossy/translucent one. In case of ties, the average point of the tied objects is assigned to each of them. For instance, if first three objects are tied, each of them gets 4 points, while if only first two are tied, each gets 4.5 points.

The results are visualized as boxplots, given on Figure 3. In order to check statistical significance of the differences, ranked objects were considered as pairs. Afterwards, sign tests have been conducted and Bonferroni correction [22] was applied to avoid the bias due to the multiple testing. Alpha threshold was set to 0.05.

It is worth mentioning that the experimental protocol was not identical for all observers. Some observers were clearly instructed that they could have ties, while in other cases, this was not clearly mentioned by the interviewer. Therefore, the observers might have assumed that they were forced to choose and cross-individual differences might be accounted for this factor.

Task 1 (box Q)

Color or hue was a dominant attribute used by the observers to group the objects. 13 out of 17 participants used this single criterion for clustering, while the criteria used by 4 other observers were the combination of color and translucency, transparency, "surface properties", and "material properties". However, the number of groups created based on color varied, leading to a color naming problem. The second level criteria were mostly gloss and translucency, either separately, or in combination.

Task 2 (box C)

12 observers had 2-dimensional arrangement for defining the space, while 5 observers had 1-dimensional order. 14 observers used translucency as one of the criteria. The dimensions increased in 8 cases after getting access to additional objects. However, 13 observers mentioned that either they would not have changed their space in case they had access to all objects at once, or they were uncertain what they would have done. As suggested by Thomas et al. [13], people usually tend to stick to the standards they create and feel comfortable with.

Task 3 (box X)

The task was reasonably fast taking about 5 minutes on average. Seven observers had binary ranking - grouping the objects into two: "glossy" and "matte" categories. While others had more than two steps with some ties possible. There is very clear separation between the objects, as A, B and C are always considered less glossy than D and E. On the other hand, there is no consensus among observers about ranking within "glossy" and "matte" groups, especially, between D and E. All differences are statistically significant except for that between A and B, and D and E. 5 people considered D more glossy, 5 people ranked E as more glossy, while 7 people tied them. The analysis of their argumentation revealed two different approaches: people opting for D mostly argued that as the object is lighter and more translucent, more light is coming from it and therefore, it appears more glossy. On the other hand, people opting for E argued that it has larger tonal range, as the contrast between brightest and darkest points is larger, and therefore, the object appears more glossy. In the latter case, we can think that people use the contrast gloss (as defined by Hunter [23]) as an additional cue.

Task 4 (box M)

All observers ranked object C as the most glossy one, while the difference between A and B is not statistically significant. Objects A and B have the same level of surface coarseness, while they substantially differ in transparency. In this particular case, we achieved the same gloss perception with the same surface coarseness, even when other material properties are different. We can hypothesize that similar gloss appearance can be achieved with similar surface coarseness. This is in agreement with microfacet BRDF model [24, 25, 26]. However, the limits of this hypothesis need to be understood. As we have demonstrated for Task 3, transparency and lightness can impact gloss perception among some individuals, even when surface properties are the same.

Task 5 (box P)

All five objects have the same surface coarseness. According to the hypothesis drawn from the Task 4, their perceived glossiness is expected to be the same. It is interesting that there is no clear trend in ranking and no statstically significant difference among perceived gloss of the object. The only statistically significant difference was observed between D and E. Five observers decided that all objects have same amount of glossiness. In spite of this, other observers forced themselves to use various cues for ranking. While some used the same argumentation, as in case of the objects from box X (ligher and more translucent ones being more glossy, i.e. objects A and B), others used the clarity of their own image reflected on the surface, listing C, D, and E

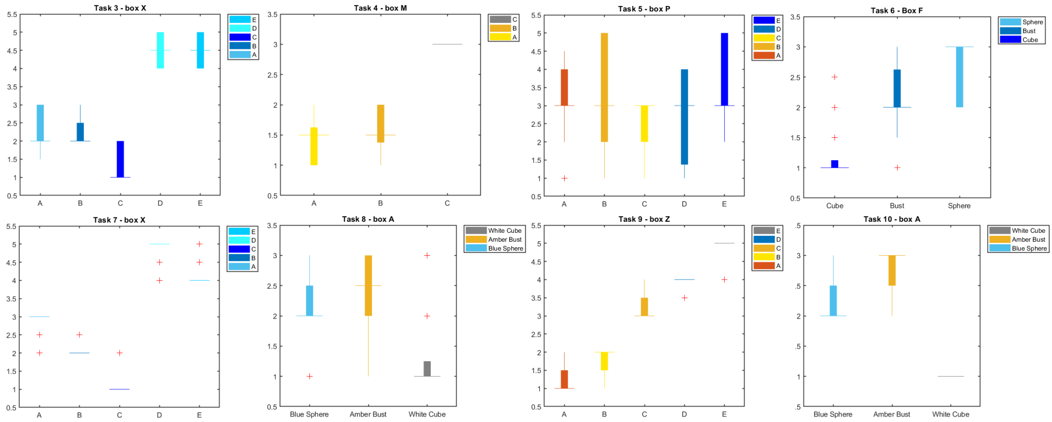

Figure 3: Boxplots for observer scores. Central mark -median; bottom and top edges - 25^{th} and 75^{th} percentiles, respectively; Whiskers extend to the extreme data points excluding outliers; red '+' symbol - outliers. We can observe clear separation for Tasks 3 and 7, clear order can be seen for Task 9, while no difference is significant for Task 5.

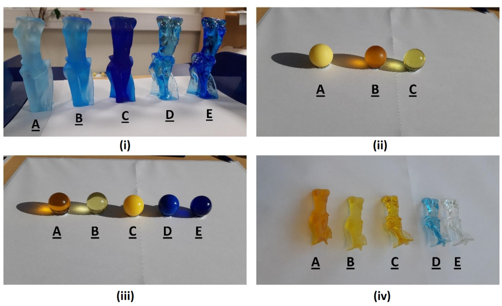

Figure 4: The objects used: **(i)** Tasks 3 and 7 (box X). **(ii)** Task 4 (box M). **(iii)** Task 5 (box P). **(iv)** Task 9 (box Z).

as more glossy ones. This implied that they come a bit closer to the object and then, the intrinsic properties of the material permitted them to infer differences. Hunter [23] defined six types of perceptual gloss. Apparently, specular gloss that is "the most commonly measured parameter in experiments as an approximation for the physical measurement of perceptual gloss" [27] is widely used by the observers. On the other hand, we might argue that distinctness-of-reflected-image gloss is a secondary cue for judgement used by some observers. However, we think that the observers use different reflections from different light sources rather than different types of gloss as a cue. When the reflection of a very intense point light source is equivalent (the sun in our experiments), the observers might have tried to estimate ambient structured light in the room, which was too low to generate a very bright specularity. Therefore, the observers tried to evaluate distinctness of the reflected image. However, considering the data we have at hand, no statistical correlation have been found between average intensity of illumination (mean of the illumination in Lux at the beginning and the end of the experiment) and usage of distinctness-of-reflected-image as a cue. However, illumination has changed rapidly for some experiments due to meteorological conditions and thus, we need more controlled conditions to examine the hypothesis.

Task 6 (box F)

The decisions were very consistent about the rectangular object, 13 observers considering it least translucent and thereby, making the difference statistically significant. However, the difference between the bust and the sphere is statistically negligible. The rectangular object has more coarse surface than other objects, while surface coarseness is the same for the bust and the sphere. On the other hand, the sphere has one-level-less amount of blue dyes than the rectangular object and the female bust. Ranking the cube as least translucent can be accounted for the combination of its compact shape in comparison with the bust, higher amount of dyes in comparison with the sphere, and higher surface coarseness in comparison with both objects. Despite the fact that the bust has higher proportion of the dyes, we still have insignificant difference in perceived translucency with the sphere. This can be explained with the presence of thin areas in the bust, while the sphere is a compact and thick object. Objects of the same shape with varying material properties are often used in appearance studies (e.g. [8, 14, 28]). However, our data has some indications that shape might compensate for the difference in intrinsic material properties and generate the similar translucency perception of the overall object even if the material is less translucent. To test this hypothesis, further experiments are needed using different levels of dye mixture, and same level of surface coarseness.

Task 7 (box X)

In contrast with the first occurrence of this box, the results are very consistent among observers. All differences are statistically significant. 14 people ranked them in the following order from least translucent to the most translucent one: C (least translucent), B, A, E, D (most translucent). In this case, dye mixture and surface coarseness factors do not contradict and compensate each other that makes ranking simple for the observers.

Task 8 (box A)

13 observers considered the cube as least translucent one. Although there have been three observers who ranked this object first. The reason can be the experimental protocol, as the phrase "how light is going through" used by the experimenter was interpreted differently. While some participants understood this phrase as the complexity or simplicity of the light interaction with the objects, others judged simply the amount of light transmitted through them. The ambiguity of the instruction makes sphere and cube the only pair that are significantly different.

Task 9 (box Z)

The observers demonstrated very high consistency when ranking the objects by translucency. While shapes are the same in contrast with Task F, surface coarseness and dye mixing should be impacting perceived translucency. All differences are statistically significant except for that between A and B. Object A and B have a more rough surface than objects C, D, and E. Their surface scatters the light, and blurs the content behind. Besides, Object A has higher portion of yellow dyes, and therefore, considered mostly less translucent than B. However, four observers discarded the "color difference" and ranked them as equally translucent. While object E has smooth surface and no colorants inside, it is intuitive that the object is considered most translucent. There is more neutral transparent material in bluish object D than that in yellowish object C. However, as absorption and scattering properties of the two colorants are different, the effect of their concentrations are not directly comparable. The fact that bluish object is considered more translucent can be accounted for more complex cognitive factors too. Most observers described bluish object as precious and glassy, i.e. something associated with transparent material. On the other hand, yellow one was compared with jelly, less precious plastic, or amber - something to be less prone to transparent. The most interesting case is ranking object C over B, despite having higher concentration of the colorants. We can hypothesize that translucency perception is impacted by surface roughness and lightness of the object. What are the limits of the impact by each factor needs further investigation of the objects with varying surface roughness and dye concentration.

Task 10 (box A)

All observers considered the cube least glossy. However, the difference between the sphere and the bust is not statistically significant. This is an interesting case where objects with similar surface coarseness, but with different shapes and color intensity evoke similar gloss perception.

Task 11 (box T)

There has been interesting inconsistency in what observers consider the limit of being opaque or translucent, as particular objects were sometimes classified as opaque, and sometimes as non-opaque. Even when people observed a certain translucency for some of the opaque spheres, they still classified them opaque. We suggest that opacity does not imply the absence of translucency. However, this topic requires further investigation.

Discussion

After analyzing the data, we can say that expert observers are more scrupulous with taking decisions, judging objects from many different observation geometries, moving objects, trying to look through them and moving head to detect specularities, while non-expert observers decide faster. The interesting trends have been identified in the vocabulary usage, as experts tend more to use common appearance attributes "color", "gloss", "translucency" and "texture". Parallels with familiar objects using words like "icicle","gelatine","amber","milky","honey" etc. have been widely used. This phenomenon has been also observed in the paper by Thomas *et al.* [13] Nonetheless, the full analysis of behavioral patterns and vocabulary statistics will be conducted in further work. On average, each experiment took 1 hour and 7 minutes. Non-expert observers were 16 minutes faster spending 54 minutes on average, while the experiment took 70 minutes for the experts. However, small number of non-expert observers makes difficult to generalize the finding.

The quantitative data has shown that in some cases people are very consistent in what they consider glossy or translucent. Decision making is very easy and the objects are clearly separated. Although in other cases opinions vary a lot and the observers made diametrically different decisions. While poor experimental protocol could impact the result in some cases, there is clear indication that for this dataset cues used by different people vary and that the surface coarseness, dye concentration, and shape of the object play significant role. Furthermore, complex cognitive factors could also contribute to the final outcome.

The major questions can be drawn from above mentioned analyses: whether the trends observed for this dataset can be generalized to other objects and materials, and what are the extent surface coarseness, shape, and dye composition can impact and alter gloss and translucency perception? Considering the dataset, the interview, and the conditions, it is not possible to derive a general model of perception from these data. However, we still could identify some interesting trends to define research hypotheses for our future experiments.

Conclusion and Further Work

We have conducted a set of experiments investigating appearance assessment using real objects in uncontrolled conditions. Quantitative results show interesting cross-individual differences and similarities. We suggest that surface coarseness, material composition, and shape impact gloss and translucency perception.

It is worth mentioning that different tasks generated contradictory research hypotheses. For instance, considering tasks 4, and 10, we demonstrated that similar gloss perception is achieved, when the surface coarseness is nearly identical. On the other hand, task 3 has shown that transparency and lightness also impact gloss perception. Another hypothesis is that shape is significant factor for translucency perception and in some cases, can even outweight the impact from intrinsic material properties. Considering the results of the task 9, we suggest that when the shapes are identical, surface coarseness and dye mixture have most significant impact on translucency perception. The results of task 11 lead us to the hypothesis that opacity does not imply absence of translucency. We plan follow-up experiments to investigate those topics.

Finally, we also plan to conduct a comprehensive study of behavioral patterns and vocabulary better to understand the processes that lead us to given quantitative results. As we are limited to resin objects in this experiment, other materials and computer graphics could be used to generalize the findings.

References

[1] Yasir Nawab, Syed Talha Ali Hamdani, and Khubab Shaker, *Structural Textile Design: Interlacing and Interlooping*, CRC Press, 2017.

[2] Christian Eugène, "Measurement of "total visual appearance": a CIE challenge of soft metrology," in *12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement*, 2008, pp. 61–65.

[3] Thorsten Hansen, Maria Olkkonen, Sebastian Walter, and Karl R Gegenfurtner, "Memory modulates color appearance," *Nature Neuroscience*, vol. 9, no. 11, pp. 1367, 2006.

[4] M Pointer, "A framework for the measurement of visual appearance," *CIE Publication*, pp. 175–2006, 2006.

[5] Yun-Xian Ho, Michael S Landy, and Laurence T Maloney, "Conjoint measurement of gloss and surface texture," *Psychological Science*, vol. 19, no. 2, pp. 196–204, 2008.

[6] Shin'ya Nishida, Isamu Motoyoshi, Lisa Nakano, Yuanzhen Li, Lavanya Sharan, and Edward Adelson, "Do colored highlights look like highlights?," *Journal of Vision*, vol. 8, no. 6, pp. 339, 2008.

[7] Edul N Dalal and Kristen M Natale-Hoffman, "The effect of gloss on color," *Color Research & Application*, vol. 24, no. 5, pp. 369–376, 1999.

[8] Roland W Fleming and Heinrich H Bülthoff, "Low-level image cues in the perception of translucent materials," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 3, pp. 346–382, 2005.

[9] Bei Xiao, Bruce Walter, Ioannis Gkioulekas, Todd Zickler, Edward Adelson, and Kavita Bala, "Looking against the light: How perception of translucency depends on lighting direction," *Journal of Vision*, vol. 14, no. 3, pp. 17–17, 2014.

[10] Lavanya Sharan, Ruth Rosenholtz, and Edward H Adelson, "Accuracy and speed of material categorization in real-world images," *Journal of Vision*, vol. 14, no. 9, pp. 12–12, 2014.

[11] Roland W Fleming, Christiane Wiebel, and Karl Gegenfurtner, "Perceptual qualities and material classes," *Journal of Vision*, vol. 13, no. 8, pp. 9–9, 2013.

[12] Lavanya Sharan, Ce Liu, Ruth Rosenholtz, and Edward H Adelson, "Recognizing materials using perceptually inspired features," *International Journal of Computer Vision*, vol. 103, no. 3, pp. 348–371, 2013.

[13] Jean-Baptiste Thomas, Aurore Deniel, and Jon Y Hardeberg, "The plastique collection: A set of resin objects for material appearance research," *XIV Conferenza del Colore, Florence, Italy*, p. 12 pages, 2018.

[14] Bui Minh Vu, Philipp Urban, Tejas Madan Tanksale, and Shigeki Nakauchi, "Visual perception of 3d printed translucent objects," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2016, vol. 2016, pp. 94–99.

[15] Kevin Smet, Wouter R Ryckaert, Michael R Pointer, Geert Deconinck, and Peter Hanselaer, "Colour appearance rating of familiar real objects," *Color Research & Application*, vol. 36, no. 3, pp. 192–200, 2011.

[16] Yannick Boucher, Helene Cosnefroy, Alain Denis Petit, Gerard Serrot, and Xavier Briottet, "Comparison of measured and modeled BRDF of natural targets," in *Targets and Backgrounds: Characterization and Representation V*. International Society for Optics and Photonics, 1999, vol. 3699, pp. 16–27.

[17] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan, "Efficient isotropic BRDF measurement," 2003.

[18] Frank P Nanna and John Jereb, "Gloss measurement system," Sept. 3 1996, US Patent 5,552,890.

[19] Yoshitaka Kuwada, "Gloss measurement apparatus and gloss measurement method," Mar. 16 2010, US Patent 7,679,747.

[20] Raymond G McGuire, "Reporting of objective color measurements," *HortScience*, vol. 27, no. 12, pp. 1254–1255, 1992.

[21] David J Gozalo-Diaz, Delwin T Lindsey, William M Johnston, and Alvin G Wee, "Measurement of color for craniofacial structures using a 45/0-degree optical configuration," *Journal of Prosthetic Dentistry*, vol. 97, no. 1, pp. 45–53, 2007.

[22] C Bonferroni, "Teoria statistica delle classi e calcolo delle probabilita," *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, vol. 8, pp. 3–62, 1936.

[23] Richard S Hunter, "Methods of determining gloss," *NBS Research paper RP*, vol. 958, 1937.

[24] Robert L Cook and Kenneth E. Torrance, "A reflectance model for computer graphics," *ACM Transactions on Graphics (TOG)*, vol. 1, no. 1, pp. 7–24, 1982.

[25] Addy Ngan, Frédo Durand, and Wojciech Matusik, "Experimental analysis of brdf models.," *Rendering Techniques*, vol. 2005, no. 16th, pp. 2, 2005.

[26] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance, "Microfacet models for refraction through rough surfaces," in *Proceedings of the 18th Eurographics conference on Rendering Techniques*. Eurographics Association, 2007, pp. 195–206.

[27] AC Chadwick and RW Kentridge, "The perception of gloss: a review," *Vision research*, vol. 109, pp. 221–235, 2015.

[28] Philipp Urban, Tejas Madan Tanksale, Alan Brunton, Bui Minh Vu, and Shigeki Nakauchi, "Redefining a in rgba: Towards a standard for graphical 3d printing," *arXiv preprint arXiv:1710.00546*, 2017.

Article B

Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). "On the appearance of objects and materials: Qualitative analysis of experimental observations." In: *Accepted for publication in the Journal of the International Colour Association (JAIC)*, 33 pages

On the appearance of objects and materials: Qualitative analysis of experimental observations

Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen and Jon Yngve Hardeberg

Department of Computer Science, Norwegian University of Science and Technology, Norway
Email: davit.gigilashvili@ntnu.no

Perception of appearance of different materials and objects is a complex psychophysical phenomenon and its neurophysiological and behavioral mechanisms are far from being fully understood. The various appearance attributes are usually studied separately. In addition, no comprehensive and functional total appearance modelling has been done up-to date. We have conducted experiments using physical objects asking observers to describe the objects and carry out visual tasks. The process has been videotaped and analyzed qualitatively using the Grounded Theory Analysis, a qualitative research methodology from social science. In this work, we construct a qualitative model of this data and compare it to material appearance models. The model highlights the impact of the conditions of observation, and the necessity of a reference and comparison for adequate assessment of material appearance. Then we formulate a set of research hypotheses. While our model only describes our data, the hypotheses could be general if they are verified by quantitative studies. In order to assess the potential generalization of the model, the hypotheses are discussed in context of different quantitative state-of-the-art works.

Introduction

We observe the emergence of new ways to fabricate objects and materials, such as 3D printing [1] and advanced surface processing [2, 3]. Object manufacturing is also related to digital edition and design [4]. Both need to be supported by an adequate description of material appearance. This description may be produced with a physical measurement and its correlation with human perception but could also be related to semantic communication. A further challenge comes with the development of programmable matter [5-7]. We foresee that an object's appearance will not be limited to the natural appearance of the material it is made of, but also an object may have an evolving shape, that impacts its appearance. Therefore, description, quantification, and communication appearance is important.

According to the ASTM E284-17, Standard Terminology of Appearance [8], the **appearance of an object** is *"the collected visual aspects of an object or a scene"*; while **perceived appearance** is defined as *"the visual perception of an object, including size, shape, color, texture, gloss, transparency, opacity, etc., separately or integrated."* The same dictionary highlights that *"appearance, including the appearance of objects, materials, and light sources, is of importance in many arts, industries, and scientific disciplines."* Appearance is a complex

 ISSN 2227-1309

phenomenon that is far from being comprehensively understood. Considering its complex nature, it is usually broken down into various attributes that entail only particular dimensions of appearance. The CIE (Commission Internationale de l'Eclairage, International Commission on Illumination) defines color, gloss, translucency and texture as four major appearance attributes [9].

Appearance has long been a point of scholarly interest from physical [10, 11] (e.g. solving radiative transfer equation [12]), psychological [13], and philosophical [14, 15] points of view. Hunter and Harold [10] provided the first significant summary of appearance measurement techniques, which aim *"to obtain numbers that are representative of the way objects and materials look"*. However, they consider that comprehensive analyses of total appearance is impossible and impractical and argue that, at least, *"measurements of specific attributes of appearance can be exceedingly useful and economically important"*. Their work is far from modelling total appearance and provides little guidance on the correlation between metrology and perception.

Practical aspects of total appearance by Hutchings [14, 15] focused on unifying knowledge of appearance from science disciplines and arts, which *"can be based on a quantitative understanding of the basic perceptions of form, colour, translucency, gloss, and movement."* He describes and structures seven factors that influence total appearance [14,16]: appearance images; immediate environment factors; inherited and learned responses to specifics; receptor mechanisms; design; object properties, and light source properties and defines it as: *"total appearance combines a description of the appearance of each element of a scene... with a personal interpretation of the total scene in term of its recognition and expectation".* Eugène [13] highlights the definition recommended by the CIE *"the total appearance points out the visual aspects of objects and scenes"* [9]. On a semantic level, Eugène considers appearance measurement challenging, because it involves subjective judgment and argues that *"a goal of making measurements that ensures appropriate quality control in the manufacturing process is probably achievable, but the measurement process will be multidimensional, product specific and probably application specific"*. Choudhury [11] also reviewed total appearance as a concept and described a four-step flow of total appearance from molecular composition of an object to the high level cognitive interpretation of appearance by a human observer.

Despite those attempts, the objects' total appearance is so difficult that most research focuses on the total appearance of a material. Most recent quantitative studies aim to provide a correlation model between optical properties and perception of a single appearance attribute (e.g. [17]). Works in computer graphics, vision, and metrology focus on very narrow specific cases and provide a quantitative analysis of particular appearance attributes [18-25], or investigate the role of image attributes on appearance, e.g. [26]. Many are based on psychophysical studies with human subject involvement. However, the constraints imposed on the experimental conditions of those works limit, in general, their relevance in real life, such as, the viewing condition in colorimetry. The majority of these studies are based on images, either synthetic [23, 25] or real [27-29], shown on displays with no possibility for physical interaction. Wherever physical samples are used [30, 31], interaction and possible observation geometries are still strictly constrained. While the attributes are studied separately, it is unlikely that individual attributes of appearance are independent, e.g. transparency may impact gloss perception [32]. Furthermore, there is inconsistency in terminology. On the one hand, terminology differs across communities, e.g. *texture* in computer graphics refers to the image mapped on a mesh, while in the context of textiles, *texture* is primarily a tactile attribute describing surface geometry. On the other hand, terminology can also be ambiguous within the field of appearance, e.g. translucency, transparency, perceived translucency or opacity are sometimes used interchangeably, as in [25], which can impact the experimental observations. Further work is needed to develop a quantitative model.

In parallel to the many quantitative studies, we propose building a qualitative model of material appearance outlining general processes to formulate relevant research hypotheses. Analyzing and testing those hypotheses reveals more details of total appearance mechanisms, including people's behavior to assess appearance, the way they perceive and communicate appearance. We hypothesize that appearance is a social interaction, between an object in a scene and a person, or between two persons communicating about one object in a scene. Therefore, we approach the problem from a social science perspective and investigate how subjects interact with objects and communicate with other people. For this purpose, we conducted an experiment and applied the Grounded Theory Analysis [33], derived from the Grounded Theory Approach [34, 35], to the data collected. This method belongs to the class of inductive research methods[1]. We conducted the experiment using physical objects from the *Plastique* artwork [39] comprising resin spheres, cuboids, and complex female bust sculptures with different mixes of colorants and surface roughness properties. The process and the results were videotaped and then analyzed.

In the next section, we introduce the experiment. Then, we develop the qualitative model of our data. From this observation, we formulate research hypotheses and discuss them. We conclude by highlighting the potential limitations of this work.

Materials and methods: the social experiment

We conducted an experiment based on an interview format, which consisted of 11 visual tasks where the observer was asked to interact with physical objects, describe them and explain their choices (both rationales and actions). The experimenter asked additional questions to clarify the motives of particular actions, and to disambiguate the interpretation of the concepts by the participant. The study was reported to and approved by the NSD - Norwegian Centre for Research Data (project number 59754).

Stimuli

Generating the proper visual stimuli for the social interaction was one of the fundamental challenges in the preparation process. This study is based on real physical objects and this choice is discussed in **Appendix 1**. The objects belong to the artwork collection *Plastique* that was commissioned to the independent artist Aurore

[1] An example and method description in English can be found in e.g. [36], many other examples of studies can be found in the literature, focusing on diverse social aspects, such as [37, 38].

Deniel from "Aden Keramikk"[2]. Technical details of production, and a description of the collection and subsequent analysis of the creation process are reported in [39]. The objects in the artwork are made of resin and come in three different shapes (cuboid, spherical, and complex female bust), various colorant mixtures (from achromatic to blue and yellow), and three levels of surface coarseness (also referred to as roughness).

Experimental Protocol

The interviews were held in two rooms with different mixed illuminations from direct sunlight (subject to weather conditions) and artificial fluorescent lighting systems. The illumination was measured with a photometer at the beginning and at the end of the interview to record changes of viewing conditions. The desk, where the objects were introduced to the participant, contained some potential visual references: a white sheet of paper, a checkerboard and a pen with text on it. We expected the observer to use them as a background of reference for appearance assessment. The observers were not explicitly instructed to use these objects to preserve their natural behavior. Additionally, the checkerboard could serve geometric calibration for the camera positions.

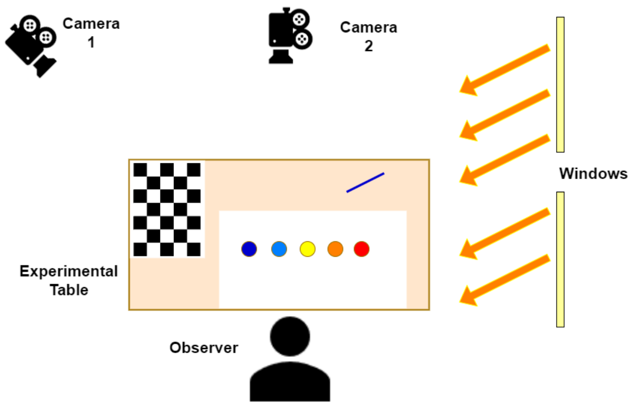

Figure 1: A Bird's-Eye Representation of the Experimental Setup. *The natural illumination incident from the windows is mixed with the artificial light incident from the ceiling (not shown). The different angles of the two cameras helped us analyze the behavior of the observers.*

People had complete freedom to interact with the objects, to touch and move them. The entire process was videotaped by two cameras (Figure 1), from front and side, to detect all potentially interesting movements and facial expressions. 17 observers, 11 males and 6 females, participated in the experiment. All of them were proficient in English. 12 of them had a scientific background related to color, vision, and appearance studies; 2 participants had an artistic background, while 3 observers were considered naïve. Their age ranged between 24 and 60, with 34 being the median age. One participant was color deficient, the others performed the interview with corrected-to-normal vision, when needed. The experiment was conducted between March and

[2] Aden Keramikk website, accessed on the 21/11/2019, https://auroredeniel.wixsite.com/adenraku

May 2018. The experiment was arranged during the day, in order to have direct sunlight in the room. On average, illuminance at the table in the beginning of the experiment was 1512 lux and color temperature was 5306 K, the standard deviation among all experiments was 766 lux and 615 K, respectively. In addition, illuminance difference and color temperature difference between starting and ending point of each interview was on average 683 lux and 497 K, respectively. We assume that some changes in participants' behavior might be related to the amount or quality of incoming light (e.g. using artificial light source for translucency assessment rather than sunlight or vice versa).

12 observers were interviewed by one interviewer and the other 5 by another one. Although the social interaction, particularly the conversation between the participant and the experimenter, was subject to improvisation and individual development, the experiment followed a well-defined routine. The observers went through 11 tasks involving set of objects grouped in 9 boxes (Figure 2). Two boxes were used twice, although this was not revealed to the participants. In the first task (box Q), observers were asked to cluster 48 cuboid objects in any way they considered natural. We wanted to observe whether one particular appearance attribute was predominant in a grouping task. In the second task (box C), observers were asked to arrange five different yellow spheres in a meaningful way, i.e. creating some ordering system for them. Afterwards, they were given additional objects with different shape, color, and other attributes, to be placed into their ordering system. With this experiment, we tried to explore potential appearance ordering systems. Tasks 3 through 10 were composed of two parts. First, observers were asked for a semantic description of the objects without touching them. The second implied ranking them by either glossiness (boxes X, M, P, A) or translucency (boxes F, X, A, Z). It is worth mentioning that the phrase "how light is going through" was used instead of "translucency", to avoid potential confusion by the term. The experiment was concluded with a binary opaque/non-opaque classification of six spherical objects (box T) with and without high intensity directional flashlight.

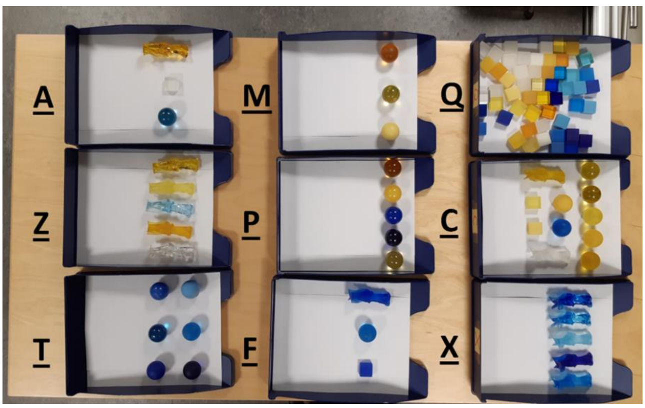

Figure 2: The Nine Sets of Objects. *The nine sets of objects have been used for eleven tasks throughout the experiment. The single letter identifiers of the boxes are completely arbitrary. The figure has been reproduced from [40]. Reprinted with permission of IS&T: The Society for Imaging Science and Technology sole copyright owners of, "CIC26: Twenty-sixth Color and Imaging Conference 2018."*

Data analysis

The data collection process was followed by a thorough data analysis that consisted of three stages:

- Two independent manual transcriptions of the collected data, i.e. more than 20 hours of video materials, were performed. This includes transcribing speech, as well as taking notes on behavior and movements.
- We performed a quantitative study on the results of the tasks by frequency analysis. This analysis was independent from transcription and was based on the task results recorded throughout the experiment. The quantitative data were presented and discussed at conferences [32, 40, 41].
- The qualitative analysis was based on the transcribed material using the Grounded Theory Analysis. Those observations were augmented and strengthened by the results of the quantitative analysis.

Qualitative model of material appearance assessment

We used the Grounded Theory Analysis [33], derived from the Grounded Theory Approach [34], to analyze the data. The method includes a comprehensive description of the observations and labeling them with codes (*coding* step). We watched the videotaped experiments (around 20 hours of video), manually extracted all observations, and labeled them accordingly. Later, conceptually similar observations are grouped into categories (*categorization* step). For instance, we observed that if the object is lit from behind or if it is placed on a textured background, it can look more translucent. These observations are grouped together into the **"Conditions of Observation"** category. Those categories were carefully designed, defined and consolidated - in particular, they were consolidated with the quantification of some of the observations. Afterwards, we identified how different categories interact with each other (*co-linking* step) that eventually leads to *modelling* through the *integration*, where we redefined and refined what we observed. The process led to *theorization*. According to the Grounded Theory Analysis as described in [33], *theorization* is a process that is more advanced than a mere description of observation (more conceptual and better structured), yet still anchored in the observation, but far from a general theory. The potential of generalization towards a theory of our *theorization* is discussed in the next sections. The coding part was performed two times independently by two persons. The categories were consolidated and revised, and the subsequent steps were conducted jointly.

The main reason for choosing this method is that the result, while qualitative, should guarantee to be strongly rooted in the data, and there are security mechanisms that avoid falling into an individual interpretation, e.g. the verification that all the codes are belonging to at least one category. Another reason is that this method is known to allow the experimenter to improve his or her understanding of the phenomenon to be studied, and the authors of this article benefited greatly from this collateral effect.

Definition of categories

We have identified the following categories that encapsulate all the codes observed in the codification step:

1. **Object** is a given sample to be considered for a particular task. It is very stable because its intrinsic parameters are static (e.g. shape, surface, size, but also specific light effect). However, it is dynamic at the same time because its appearance may vary depending on the conditions of observation.
2. **Conditions of Observation** is a set of extrinsic factors that permit the observation, contribute to the appearance of a given object and the communication of it. Conditions of observation is the place and an individual observer (illumination geometry and spectral power distribution, experimental room interior, viewing angle, personal vision, physiological condition and mood, background, vocabulary pool, etc.) We want to highlight that observer is not a separate category but part of the conditions of observation. We are presenting an objective cross-observer generic model representing a task-motivated material assessment process. The way a subjective psychological or physiological condition of the observer contributes to the overall process is by nature no different from illumination geometry or other external conditions of observation.
3. **Methodology** is a stable systematic way to act and make decisions towards completion of a task. Methodology can be based on intuition or experience, and it could converge and be revised after trial and failure (calls **Learning and Adaptation**).
4. **Comparison** is an action that permits judgement of the objects by referring to something else, making assessment relative to a **Reference**. Similarities and differences are judged either with an arbitrarily chosen reference or among different states of the object itself, that becomes the reference.
5. **Reference** is the observation, memory, concept, etc. an object or a set of objects are compared with. This is one of the most important categories when we want to discuss measurement of appearance.
6. **Vocabulary Search** is the process to identify and select the right **Vocabulary** in order to communicate and express the perceived appearance of a given object or set of objects. In the process of **Vocabulary Search**, different methodologies might be applied, including, but not limited to, citing standard definitions from the literature, recalling familiar objects from memory in order to draw parallels, or looking up for proper words on the Internet.
7. **Vocabulary** is a selected set of words, like adjectives, nouns, phrases (e.g. "blown-up glass") - all attributes and labels used to describe the appearance of a given object or set of objects. The selection of this set is derived from the **Vocabulary Search** and serves as a basis for the **Semantic Description**.
8. **Semantic Description** consists of tentatives to name, or to describe the appearance of one given object or a given set of objects.
9. **Completion of a Visual Task** is a process to successfully perform a given mission that relies on the analysis of the visual appearance of a given set of objects but also on the **Task Interpretation**.

10. **Task** is a given mission an observer is instructed to accomplish by an **Experimenter**. We used those tasks to lead the interviews.
11. **Experimenter** is a person, in our case one of the authors of the paper, who introduces tasks to the observers and guides the entire process by oral communication with an observer. The communication and interaction with an observer were subject to individual improvisation by the experimenter. Thus, this impacted the data and made all experiments unique.
12. **Structure Expectation** is an assumption by an observer that there exists a structure in the data. This structure, that may or may not exist, will be used as a cue to perform the task, instead of, or in addition to, relying on visual qualities. This implies that the participant assumes that there is an expectation or a solution known by the experimenter, which was not the case.
13. **Task Interpretation** is a decoding process of the oral description of the task conveyed by the **Experimenter**. The observer tries to understand what they are expected to do and selects a **Methodology** to reach the goal.
14. **Decision-making** is a general approach that leads the observer to the strategy on how to perform a **Task** that involves freedom of interpretation. This was not observed in all experiments, because some tasks were less prone to interpretation.
15. **Learning and Adaptation** is a function of time affecting actions of the observer. It impacts the processes we have observed. As the observer interacts with the corpus of data, their understanding of the data is refined based on the recently acquired experience. Secondary visual attributes, like scratches and imperfections start to be taken into account, leading potentially to refinement in **Methodology**. Observers start recognizing similarities with the part of the corpus already studied and behave accordingly. It can have a positive impact and facilitate the task completion or a negative impact related to exhaustion, shortcut or overconfidence.

Description of the qualitative model

The resulting model of the data is illustrated in Figure 3. The model consists of two blocks. The pivotal visual part unfolds the flow of the process from introduction of the object towards the completion of a particular mission. An auxiliary decision-making part describes all the factors that could impact a methodology selection in the process of task performance. It is worth mentioning that the decision-making part only impacts the result of the experiment, *i.e.* what we observe by the frequency analysis, but does not change the model and the flow of the processes itself. The structure of the model is independent of the observer and the task.

The **Object** is observed in certain **Conditions of Observation**. The combination of both categories creates in fact the core of the sensory perception of the object by a person. While the **Object** has some absolute properties, total appearance is impacted by the various **Conditions of Observation**. Anything that can impact the perception of the appearance of an object is considered a **Condition of Observation**. While usually conditions impact the object appearance, the interaction is both-ways, as an object could also impact the conditions (e.g. produce caustics, evoke particular memories). The category **Methodology** is at the heart of the observation. In fact, we observed how the participants perform the task and describe their actions and decisions. Indeed, the **Object** and **Conditions of Observation** constrain the **Methodology**. However, we observed that there are major contributions from **Comparison** and the **Decision-Making** which define or constrain the **Methodology**, and in our data, they might be as important as the perception part because they are very general. Both of them are induced by the **Task** given to the observer. The **Comparison** is required to analyze the samples, and this is done by **Reference** to *something*. As we shall see, the observation that a reference is systematically used is a crucial piece of information, which is both very positive from a perspective of metrology, but also a great challenge when it comes to selection of an appropriate reference. **Decision-making** is required when a **Task** leaves room for interpretation, and is based on the **Task Interpretation**. It is closely related to the **Task** itself, the way it is conveyed by the **Experimenter**, and constrained by the **Structure Expectation** on the data. The latter was observed in our experiment, but it is hard to anticipate whether this will be observed in a more free context. Observers applied various decision-making models to come up with an efficient strategy and select a particular **Methodology** to complete a mission [41]. Based on the **Methodology**, the visual task is solved and the observer reaches the **Completion of a Visual Task**. We also observe that the **Methodology** is used to structure the **Vocabulary Search**, that led to a selection of **Vocabulary** used to come up with a **Semantic Description**. Several methodologies were observed to be pre-selected, in order to find, choose, and convey the **Vocabulary** necessary for **Semantic Description**. **Semantic Description** can be a substantial prerequisite for the **Completion of a Visual Task**. We observed that subjects tend to describe objects in the process of **Completion of a Visual Task** even if they are not explicitly instructed to do so. In order to assess appearance, they seem to construct a semantic image of the target in their mind with or without explicit oral expression. In addition, the description of the objects might already include the draft solution of the visual task (for instance, object A is described as glossier than B and as less glossy than C, while the visual task is to rank the three by glossiness). Finally, we should highlight that a significant impact of **Learning and Adaptation** was observed throughout the experiment and it impacts all other categories.

Verification and Analysis

In order to demonstrate how the model is rooted in the data, we describe an example case in **Appendix 2**, where the observer is asked to rank five spheres by their glossiness. We recall that this model is a model of our data. However, it is interesting to study how those data compares to general models of material or object appearance by Hutchings [16], Choudhury [11], and Eugène [13]. They all referred to the scene context, supported by the CIE definition that also includes scene concept into the total appearance [9, 13]. In our data we can observe how this context is verbalized by the observers. The context is summarized in the **Conditions of Observation**. These conditions were experienced by the observer, but explicitly mentioned only when these conditions constrained successful completion of the task. Otherwise, the impact of the scene was encapsulated in the **Semantic Description** and in the **Completion of a Visual Task**. For example, observers ranked an object by gloss, using distinctness-of-image gloss when the light was low enough, without further discussing the environment. However, when intense direct sunlight made it impossible to observe distinctness-of-image gloss, the observers discussed the scene and mentioned that the sunlight in the scene made task completion difficult.

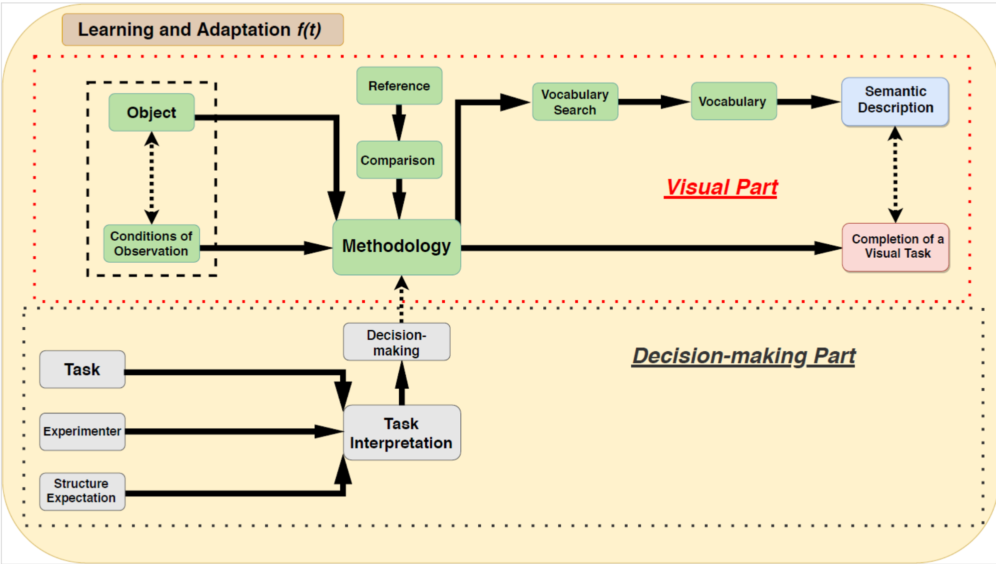

Figure 3: Qualitative Model of Material Appearance Assessment. *The primary Visual Part of the model details the flow of the process from introduction of an object in particular conditions to semantic description of its appearance and completion of a visual task using this object. Auxiliary Decision-making part illustrates categories impacting methodology selection in the Visual Part, while Learning and Adaptation impacts the entire process as a function of time (f(t)).*

Eugène [13] supports the idea of total appearance implying higher level semantics, for instance concepts like, *"visually assessed safety", "visual identification of the scene", "visually assessed usefulness of the scene"* etc. in addition to Hunter's attributes. In our data, this appears in the **Semantic Description** when observers describe the objects as *"like food"*, *"fragile"*, *"pricy"*. In addition to appearance attributes, they also referred to high level semantics, like usefulness ("decoration", "soap"), safety ("fragile"), in order to express and communicate the appearance of the objects and materials.

Apart from that, Hutchings considers that *"there are two classes of appearance images: the impact (or Gestalt) image, and the sensory image. The impact image is the initial perception of the object plus an initial opinion or judgment."* [16] This is also present in our model, where the sensory image is limited by the **Object** and the **Conditions of observation**. This is also the case for Choudhury's model [11], where the three first stages correspond to the sensory image of Hutchings and the fourth one is related to higher cognitive interpretation. Choudhury also emphasizes the physiological phenomena as an explanation of the process, which we do not consider.

To conclude on those comparisons, it appears that the works discussed above focus much more on the sensory analysis, while we observe more on the human behavior, semantic description, decision-making and task-solving than them. Compared to their works on those aspects, which are a formulation of opinions, what we observe is rooted in our data. Our model is centered around the completion of a visual task, while there is no motive of appearance interpretation introduced in those other works. We, however, all agree on the idea that conditions of observation (including environmental or individual background aspects of a human subject) have a tremendous impact on perceived appearance.

Three key behavioral observations

The omnipresence of a reference

Comparison with a reference turned out to be a pivotal point of all methodologies applied for visual task performance, as well as for semantic description. The reference varied and was any of the following, but perhaps not limited to:

a) Comparison to the appearance of another object (e.g. comparing two objects to decide which one is glossier).
b) Comparison of the appearance of the same object under different conditions of observation (e.g. move an object from shadow to direct sunlight to assess its translucency).
c) Comparison of the perception of the background through the object or by direct view (e.g. try to read a text through the object and see how much is it distorted to assess transparency).
d) Comparison to memory of familiar objects (e.g. comparison with an appearance of a favorite childhood candy).
e) Comparison to a hypothetical idealistic object or material (e.g. comparison of a glossy object to a perfect mirror).
f) Comparison to a definition (e.g. "gloss, n. — angular selectivity of reflectance, involving surface-reflected light, responsible for the degree to which reflected highlights or images of objects may be seen as superimposed on a surface" [8] - thus, only the surface is analyzed, rather than the actual sensation of gloss).

Comparison with a reference is a measurement process. The standardization of this reference as a unit of measurement is the fundamental aspect of metrology. In order to quantify and communicate visual appearance, subjects need such a reference that will be used for quantification of the appearance. If one does not exist, we have observed that they try to create one themselves. However, the process to come up with a standard is difficult. For instance, a standard for length implies the usage of one unit, and a standard for speed is based on two units (distance and time), while the standard for appearance should regard many components considering the complicated nature of appearance as a phenomenon. Even though the selection of references is very subjective by nature, the process is still conditioned by the physical world. We have observed that people without much training perform surprisingly well on complex tasks that are impossible nowadays for machines and tools [27, 29, 42]. We believe that in case appropriate physical measures and references are used, we should be able to mimic this ability. Even though Eugène [13] argues that *"it is unlikely that any* *physical scale called "appearance" will be possible"*, he admits that *"it is necessary to find physical parameters that can be measured and the most obvious area for exploitation is that described in terms of the optical properties"*. References vary depending on the context: comparison can be with a local reference (e.g. with another object), or with a global reference (e.g. the appearance of marble according to the subject's memory); comparison can be with objective things (e.g. definition of blue), as well as subjective ones (e.g. a gummy bear that tastes very good). However, communication of appearance requires generalization and some objectivity - in most cases, we have a common understanding and agreement on the definition of the words we use to communicate appearance (e.g. "green" refers to a set of colors most of the general populace agree upon with some marginal exceptions, e.g. [43]).

When **global references** are not enough for a given visual task, the Human Visual System (HVS) might use a **local reference**. Simultaneous contrast and dynamic range adaptation are a good demonstration of this. We have observed in our data that the reference is floating, i.e. varying across situations. We believe that this can be a general pattern for material appearance assessment. In other words, the reference could be application-, material-, or situation-specific. We have observed that references have been selected based on the peculiarities of a given scene. When observers were asked to assess the translucency of an object, they usually looked through the object towards the brightest light source (usually the sun), comparing the original appearance with the appearance of the same object under back-lit illumination geometry (back-lit geometry is typically used for measuring "through translucency" [44] or transmission of translucent materials [45]). When the sunlight was not visible observers tended to use an artificial light source of the room instead. Change of reference depending on the illuminance of the artificial light sources has also been observed in [46]. As this was subject to presence of the bright light source, some observers also moved their fingers behind the object comparing the cues between blocked and non-blocked light source conditions. This supports the notion that illumination and room interior, i.e. *Conditions of Observation*, impact *Methodology*, thus reference selection. Back-lit illumination geometry has been already demonstrated to increase the perceived translucency of the materials [23, 25].

Although the HVS is very sensitive, it is not capable of stand alone quantitative measurements. Humans can discriminate perhaps 5 to 10 million colors when seen side-by-side [47]. However, when the stimuli are seen with long time intervals, it is difficult to tell the difference, unless the difference is very large - proposedly, our memory stores only around 300 colors [10]. While memory as a global reference has limited capacity, presence of a local reference in a particular point of time, could dramatically enhance the discriminative capabilities of the HVS.

For such a high dimensional problem, probably the reference should not be very different from the target. Deborah [48] addresses the importance of reference selection in the context of spectral differences, considering it an important aspect for a metrological hyperspectral image analysis. The author represents an image as spectral sets falling within a convex hull and argues that if the reference is far outside of the convex hull, the distance to all cluster centers will be nearly identical and discrimination will be poor. Drawing a parallel with appearance, we have observed that a transparent reference medium is a poor measure of apparent translucency differences [49].

Fleming discusses "statistical appearance models" as a potential mechanism for material appearance perception [50]. The author argues that instead of estimating physical properties of materials, our visual system identifies salient features of a given material and creates an internal generative model to estimate how these features behave (i.e. vary across conditions), in order to identify a material in different contexts. The model *"seeks to discover in what ways different material samples look different from one another"*, where comparison process and need for a reference seems inevitable. He further argues that our brain tries to characterize systematic changes in the look of materials and the model is *"refined and corrected through experience with other samples"*. This process highlights the importance of reference in material perception, and resembles searching for the optimal reference in our data. The author also describes two pivotal forms of material perception: estimation - assessment of potential characteristics, and categorization – assigning a particular label or material name. Considering his explanation that *"material estimation is the process of establishing the true position of a given sample within the feature space, and material categorization is the process of identifying the boundaries separating different classes of material"*, it becomes obvious that neither process is possible without comparison with a reference. Furthermore, material perception as a categorization process has another interesting aspect - it implies *"access to stored knowledge about other members of the same class"*. This phenomenon has been observed in our data and we describe it as *a reference to memory*.

Multisensory impact on appearance

While reference selection and change might imply direct interaction with the object, the interaction can itself provide additional information for appearance assessment, because relying on visual stimuli might still not be enough for material identification, as demonstrated in [51]. We noticed that observers frequently failed to guess the material without touching the object, even though they could move themselves and inspect fixed objects from various viewpoints. Multisensory information, like auditory (knocking objects on the table), tactile information (examining the surface with a finger), or weighting them by hand, have been used to identify material and to describe it [52]. However, it is worth mentioning that after some time, observers demonstrated adaptation, as they got familiarized with the dataset and concluded that the collection is composed of resin materials only.

Choudhury notes that *"although visual perception apparently seems to be independent of human sensation, some properties are perceived in different ways by more than one sense. Individual visual attributes may arise from combination of signals from different senses."* [11] Limited multisensory interaction in computer graphics might lead to material metamerism and unrealistically large constancy of appearance attributes [52]. This supports our idea that physical objects are important for studying appearance. While we have observed in our data that multisensory information facilitates material identification, neither of the following is clear: whether material identification impacts the perception of the appearance, or whether auditory or tactile information impacts visual appearance. For instance, does the object identified as glass look glossier because this is a typical look for glassy objects? Or if we feel with our finger that the surface of a material is smooth, will it look glossier? It has been shown that priors and expectations regarding familiar-looking materials might actually impact the perception of various mechanical and optical properties of materials [53]. To what extent this applies to visual appearance attributes definitely deserves further study.

Semantic aspects

Analysis of the semantic description has also revealed interesting trends. In [41] we have introduced a hierarchy of the criteria used to assess appearance similarity. Interestingly, it resembles to the vocabulary used for semantic description of the appearance of the objects. The observers have taken different approaches for semantic description that could be diversified into several categories either by tactics, scale, or semantics of the description.

Tactics: 1. Material identification (e.g. amber, ice, silicate, glass, plastic) 2. Attribute-based (glossy, blue, transparent) 3. Familiar object and function identification (e.g. soap, fortune-telling crystal ball, souvenir sold in shops, eraser) 4. Any combination of the previous.

Scale: 1. Absolute (describe just the object) 2. Relative (glossier than this; rougher than that surface).

Semantically: 1. Description as quantification of appearance attributes - the same routine for all objects, e.g. "this object is blue and somewhat glossy". 2. Description as a creative process (comparison with unusual stuff like sorcery; analyzing and describing impact of artifacts on caustic formation; conveying appearance with emotions, like "this looks boring").

All these approaches to semantic description involve comparison with various references. It is worth noting that selecting the attributes to communicate the appearance might be dependent on the similarity or dissimilarity within the corpus. For example, when the shape of all objects under question was identical, shape was mentioned less frequently in semantic description than in the cases, where observers had to describe objects with different shapes.

Formulation of the research hypotheses

While the above discussion refers to our data only, the model and the observations might be general to some extent. We formulate 20 research hypotheses (**H1**-**H20** in the rest of the paper), which, if validated quantitatively, can help us to understand the generality and the limits of our model. The verification of the hypotheses is usually based on quantitative experiments. Some related experiments are already reported in the literature and we use this literature to have a critical reading on those hypotheses. We want to make clear that the verification of the hypotheses do not challenge the existence of the qualitative model, since this is a model of the collected data.

Reference

H1: It is possible to measure and predict perceived appearance. There should be reference(s) and comparison protocol(s), presumably specific to a given material and conditions, that permit objective instrumental measurement of perceived appearance. The critical challenge is to discover these references and comparison protocols.

H2: Human subjects limit one comparison to a single reference at a discrete point of time in appearance assessment process. We have observed that oftentimes, ranking, clustering and ordering visual tasks were broken down into several pair-comparison tasks. For instance, when a subject was asked to rank objects by glossiness, they compared a given object with other objects individually, one by one.

H3: A general appearance ordering system (empirical) cannot exist in sensibly low dimensions. It should be either application specific, local, or most probably unintelligibly high dimensional. If such system would ever exist, it will be strongly non-uniform by nature. There have been several studies in context of material appearance, where *n* manually selected attributes, i.e. features, have been quantified psychophysically to learn how materials relate with one another in a given *n-dimensional* feature space [28, 52, 54]. However, it is observed in [41] that a manually defined system often fails to accommodate new out-of-the-corpus objects.

Conditions of observation

H4: Multisensory information and interaction level impact the robustness of appearance constancy. On multiple occasions we observed multisensory impact on visual assessment. Although visual information is unarguably essential to visual appearance, the role of other senses is yet to be understood. It has been shown that different senses, such as visual, tactile and olfactory impact each other in aesthetics impression [55], object recognition [56], material identification [57, 58] and material perception [59]. However, the exact way multisensory information contributes to visual appearance is not understood yet.

Object

H5: Shape difference can dramatically impact appearance difference even for identical materials. This observation is consistent with the state-of-the-art. Vangorp et al. [60] illustrated that difference in shape, particularly tesselated geometry, diminishes material matching accuracy and comparison is easier between identical shapes. It also impacts perceived translucency differences [49]. As perceptual attributes, such as gloss [61-63] or lightness [64] vary across shapes, it is no surprise that total appearance is also impacted.

H6: Confusion between subsurface and surface scattering might lead to equivalent appearance through different physical material properties. We believe this point boils down to the question whether the HVS can separate contributions of surface and subsurface scattering to the image information. If this is not the case, it could support our proposal that **translucency impacts gloss perception.** We think the confusion can be minimal for gloss if a sharp image of the environment is reflected from the surface, which is subject to presence of well-structured real-world illumination [65]. However, the orientation of the reflected image can also cause confusion between transmission and reflection phenomena [66].

Translucency perception

H7: The amount of transmitted light and preservation of the light structure after transmission are independent, but core dimensions for translucency assessment. From the perspective of hard metrology, this observation can be related to concepts such as, *direct*, *diffuse* and *total transmittance*, as well as *clarity* and *haze* [9, 45]. However, perceptual dimensions of translucency are yet to be understood. In a translucency classification system proposed by Gerardin et al. [67] independent orthogonal dimensions of diffusion and absorption are roughly equivalent to these quantities. However, the authors argue that increasing scattering (i.e. diminishing light structure preservation) makes transparent material to some extent translucent and finally opaque; while increasing absorption (i.e. amount of light) does not cause translucency and ranges from transparency to opacity without translucency in between. This is contradictory to some of our observations that people consider absorbing objects less translucent, even in case of identical scattering properties. We have observed that **the assessment procedure of perceptual translucency difference depends on the subjective interpretation of the term and needs to be standardized.**

H8: A given material looks more translucent when an object made of it has thin parts. This phenomenon is illustrated in Figure 4 below. The observers considering objects with thin-parts more translucent, instead of referring to low level image cues, explicitly mention that they understand and see that the light is being transmitted through the object. This can be an indication that Fleming and Bülthoff's [25] conclusion that the HVS does not invert optics to assess translucency might not hold for thin objects. In general, shorter the distance a photon needs to travel through a medium, easier to detect light transmission. Scale and thickness of the object impact perceived translucency and thin parts, such as edges, are usually informative translucency cues [17, 25]. In addition, thin parts, such as fine surface details and bumps, might blur the background image and make transparent materials appear translucent (Figure 5). Therefore, this hypothesis can be reformulated as a more general statement that **object shape and size impact perceived translucency of the material.**

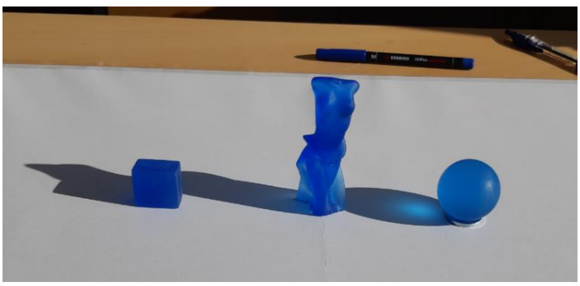

Figure 4: Three Blue Objects Used in the Experiment. *The cuboid and the female sculpture have equal density of the blue colorants, while the sphere has less blue colorants in the volume. On the other hand, the surface coarseness of the sphere and the sculpture is identical, while the cuboid has rougher surface than the other two. Combination of the two factors, led the vast majority of the observers to consider the cuboid least translucent. On the other hand, there was no statistically significant difference in apparent translucency of the sphere and the female sculpture, despite higher density of the colorants in the latter. This can be explained with the fact that a sphere has a dense shape, while the sculpture has thin parts letting the light through.*

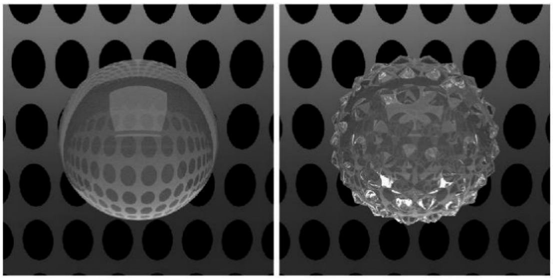

Figure 5: Same Material, Different Transparency. *Although the material is identical in both objects, meso-scale geometry of the right objects removes see-through cues impacting perceived transparency and translucency of the material and object. The images have been reproduced from [49]. Reprinted with permission of IS&T: The Society for Imaging Science and Technology sole copyright owners of, "CIC27: Twenty-seventh Color and Imaging Conference 2019."*

H9: Back-lit is a preferred lighting geometry for translucency assessment. We have observed that observers tend to locate the illumination source in the scene (typically the sun in our context) and look towards it through the object to assess translucency. One interpretation of this behavior can be a potential attempt to invert optics and observe transmission. Xiao et al. [23] have shown that materials typically look more translucent when they are back-lit. The magnitude of difference between translucent and opaque objects is expected to be larger in this condition and moving them from front- to backlight has stronger impact on translucent objects' appearance, as translucent objects, unlike opaque ones, start to shine or glow on the backlight. This is related to the above-discussed notion of *comparison with a reference.* A typical reference can be the appearance of the same object under different illumination conditions. On the other hand, it is worth mentioning that transparent objects might look less transparent on a high-illuminance backlight, as observers do not see the scene through the object due to the limited dynamic range of the HVS. [46]

H10: Dynamic and heterogeneous backgrounds enhance perceived translucency or transparency. We have observed that human observers frequently use object and background relative motion to estimate light transmission properties of a material. This implies both - moving an object over a heterogeneous background, e.g. checkerboard, as well as moving background objects behind a static object, e.g. moving one's own fingers or a pen behind the object. While in a static scene the HVS has a reduced ability to separate reflection and transmission components of the visual stimulus, human subjects try to observe and estimate the magnitude of the changes induced by the background change. Commercial measurement systems measure transmission from a static point perspective (e.g. ISO 13468 for plastics [68]) limiting the capability of measured quantities to adequately describe visual sensation in real life encounters.

H11: Lightness impacts perceived translucency (lighter objects look more translucent). Many translucent materials, such as snow, cream, milk, wax and soap, are typically light-colored and have diffusive, hazy appearance usually described by observers as "milky". Therefore, "milkiness" of light-colored objects might be the cause for perceived translucency (refer to Figure 6). Lightness has been shown to be correlated with luminance [69, 70]. Subsurface scattering can contribute to luminance and highly scattering media usually look lighter. However, lightness information alone cannot be discriminative enough for assessing translucency. Marlow *et al.* [71] demonstrated that if luminance gradients co-vary with surface geometry, surface looks opaque, while if luminance information seems independent from surface geometry, perception of subsurface scattering is evoked. This indicates that in addition to lightness, interpretation of the 3D shape is also involved.

Figure 6: "Milky" Translucent-looking Objects. *With their light and "milky" appearance, the objects evoke perception of translucency in some human observers.*

H12: Glossiness impacts translucency perception. Some of our observers considered glossy objects more translucent. It has been shown that gloss enhances perception of translucency [72] and realism of translucency appearance (refer to Figure 8 in [25]), proposedly because many translucent materials we interact with on a daily basis are glossy and *"the human visual system may "expect" translucent materials to exhibit specular reflections"* [25]. Hence, contribution of gloss to translucency perception might come down to the material identification problem. Schmid et al. [73] propose that neural aspects of gloss perception should be addressed in the context of material identification. However, the role of material association should be taken with care. Some materials (e.g. glass) appear glossy and translucent, but others (e.g. metals) can be glossy and opaque [28, 54].

H13: Presence of caustics is a cue to assess translucency and may increase perceived degree of translucency. We noticed that caustics were often used as a cue for translucency and transparency assessment by the observers, and in some scenes, might be the sole cue about translucency of the material, as illustrated in Figure 7. Caustic pattern projected by an object onto a different surface contains interesting information regarding its properties (refer to the top image in Figure 8). It was shown that when the floor and the caustic pattern projected onto it are removed, the material is judged less translucent. [74]

Gloss perception

H14: Translucency impacts the perceived glossiness of an object. We observed that gloss-based ranking has been possible for the objects with identical surface reflectance but different translucency. It has been demonstrated that translucency can impact gloss and the magnitude of this impact depends on the shape and surface roughness of the object [75]. Translucent objects with complex shape might produce highlights that originate from inside the medium - like, internal reflections, scattering and caustics. Considering the limit of the dynamic range perceived by the HVS, these highlights might be mistaken for specular reflections evoking glossiness perception [32], as shown in Figure 8. Objects can look very glassy and glossy due to internal reflections and caustics even if specular reflections are negligible (refer to Figure 8 in [51]). Additionally, Pellacini et al. [76] have shown that contrast between specular and non-specular regions is an important factor for gloss *"light colored surfaces appearing less glossy than dark ones having the same finish"*. The amount of subsurface scattering can affect lightness of the non-specular regions, while having little impact on specular ones. Hence, for some shapes, they can modulate contrast gloss of translucent objects [75].

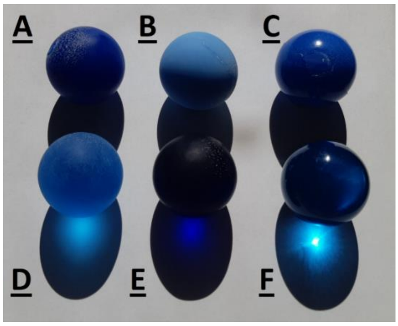

Figure 7: Translucency and Caustics. *Caustic pattern might provide information regarding color and light transmission properties of the material. For object E, it is the sole cue that makes us deduce the material is translucent. The figure has been reproduced from [46]. Reprinted with permission of IS&T: The Society for Imaging Science and Technology sole copyright owners of, "CIC27: Twenty-seventh Color and Imaging Conference 2019."*

Figure 8: Objects Used in Gloss Ranking Experiments. *We identified three groups of people: those who tied all spheres (top image) due to similarity in surface coarseness (35.29% of the observers); those who considered translucent objects more glossy, because of higher luminance and "shininess" (35.29%); and those who considered opaque ones glossier due to higher contrast and more visible distinctness-of-image gloss on them (29.42%). In the follow-up experiment with female sculptures (bottom image) the majority of the observers (78.50%) stated that the transparent ones were glossier. [32] The complex macro-geometry of the surface made it impossible to observe distinctness-of-image gloss, while these objects produced complex caustic patterns that could be mistaken for specular reflections. The top image has been reproduced from [40]. Reprinted with permission of IS&T: The Society for Imaging Science and Technology sole copyright owners of, "CIC26: Twenty-sixth Color and Imaging Conference 2018."*

H15: Complex shape makes materials look glossier. Some observers noted that a complex bust figure looked glossier than a sphere and a cube, because it shines more and has more specular regions. The state-of-the-art shows that shape can considerably impact gloss perception, even if surface reflectance is identical. It has been shown that surface reflectance constancy of the HVS fails across shapes [22] and perceived gloss is correlated with perceived surface bumpiness [62, 63, 77]. However, we see two challenges that need to be addressed:

- What is the threshold between shape change and surface change? What scale do we mean with the hypothesis mentioned earlier? Can we really change a shape without changing a surface, and if so, to what extent can we change shape not to impact the surface?
- All shape changes are due to a manipulation of a controlled parameter (e.g. RMS height deviation). Can we have a shape descriptor statistic that could predict the glossiness of a given material for any random shape?

H16: Motion facilitates gloss perception. We have observed that motion was widely used for glossiness estimation by the observers. They either moved their head or moved the objects to monitor the motion of the highlights. This is consistent with the state-of-the-art. Impact of head motion has been already observed to be important for gloss, as *"temporal changes of the retinal image caused by the observer's head motion"* and *"image differences between the two eyes in stereo viewing"* both significantly increase perceived gloss [78]. Motion seemingly helps the HVS distinguish specular reflections and surface texture. Unlike texture, specular reflections remain static relative to the observer on rotating spheres [79] and *"objects with normal specular motion to appear shinier than those with sticky reflections"* [80]. Motion improves gloss constancy [80] and can even increase the magnitude of perceived gloss [81].

Opacity perception

H17: Opacity does not imply a complete absence of transmission. We have observed that some objects manifesting translucency cues when exposed to high illuminance directional backlight were considered opaque under diffuse and low intensity illumination. While perceived opacity is proposedly impacted by the amount of transmitted light, the latter itself depends on the amount of light incident on the back side of the object. The amount of transmission tolerated for classifying the object opaque varied across observers. We concluded that opacity perception or more likely the interpretation of the concept depends on the thresholds that are floating and subjective by nature. The same trend was observed in [46]. Moreover, Marlow *et al.* [71] argue that the HVS relies on the co-variance between shading and surface orientation for distinction between translucent and opaque objects. They demonstrated that optically translucent object might look opaque *"if the light transported through the material accidentally preserves the co-variation of intensity and surface orientation"*, as if it was a result of reflection rather than transmission which again supports our hypothesis that opacity can be perceived even if subsurface scattering event occurs.

Appearance attributes and subjective material properties

H18: Glossy objects look more fragile and precious. Glossy objects with the complex shape have been described as fragile, expensive and precious. Our observations are partially consistent with the state-of-the-art. Fujisaki et al. [82] found that for wooden materials gloss and expensiveness are positively correlated. Contrasting results have been reported on the correlation between gloss and fragility, which was either positive [28] or negative [54] on different occasions. Additional role can be played with the positive correlation between glossiness and prettiness [52, 54], although some authors found no significant correlation between the two [28, 82]. We believe material identification is also an important factor, as metal, glass, and plastic can all be very glossy, they are not necessarily perceived equally fragile, neither equally precious. Material recognition and semantic interpretation of objects' function have been major contributing factors to subjective perceptual qualities in our experiment. Although observers, by visual inspection, described glossy bust figures as glass or precious stone decorations "found in a fancy store" (per contra, spheres have been described as an "ice ball", "candy", or a "billiard ball"), the auditory and tactile information made them revise their descriptions ("ah, this sounds like a cheap plastic" noted an observer after knocking the figure on the table).

H19: Darker objects look heavier. This phenomenon is correlated with *brightness-weight illusion* meaning that when lifted, a light-colored object feels heavier than a darker object of the same mass, because of the anticipation that darker objects are generally heavier [83]. Bullough [84] demonstrated that darker-colored objects are perceived heavier, proposing an explanation that darker colors evoke a perception of *"more of it"*, potentially referring to "more pigments". Interestingly, our observers provided similar justification. This finding has been supported by numerous studies [85-87]. Another intriguing explanation is that in English the same adjective *light* is used to describe both properties - low weight and high brightness [85].

Artifacts

H20: Complex surface geometry can mask imperfections and artifacts. We have observed that scratches, bubbles and other imperfections were mentioned more often when describing spheres and cuboids, and rarely for a complex bust shape. Considering that the retinal image is actually a 2D projection of the 3D object, we believe this phenomenon is related to the concept of visual masking in image quality, when noise is more apparent in homogeneous parts of the image, while it gets masked in high frequency areas [88].

Conclusion

While the vast majority of appearance studies focus on either instrumental measurement or psychophysics, we analyzed material appearance from a social science perspective. We propose that appearance is a social interaction that implies communication. We have conducted interviews where people were asked to perform visual tasks on objects of different appearances, describe the objects, explain their actions and interact with the interviewer and the objects. Those interviews were videotaped. This large collection of data was analyzed with the Grounded Theory Analysis and we constructed a model to have a structured representation of the observations. This qualitative model and its implications were described in the corresponding section. We conducted an analytical survey of the literature in the perspective of this model, and formalized future research hypotheses. In particular, we found that selecting a reference and the comparison with this reference have been the essential instruments for appearance assessment and communication in our scenario. In this work we addressed the appearance of objects, which have context, rather than the appearance of abstract materials.

Our results are to be taken with care because no level of generalization can be assumed or stated from the specific research methodology we used. Indeed, we used an inductive research method, while deductive research methods are more common in the study of appearance. The observations are limited to the conducted experiment, but when we compare our work with the state of the art, we found encouraging echoes.

Further quantitative verification of the hypotheses is a straightforward follow up of this work. Psychophysical experimental design might also benefit from our behavioral observations on natural ways of object appearance assessment. For instance, the use of extended reality technologies might permit more freedom in future experimental processes.

References

1. Brunton A, Arikan CA, Tanksale TM, Urban P. 3D Printing Spatially Varying Color and Translucency. ACM Transactions on Graphics (TOG). 2018; 37(4):157:1–157:13. doi: https://doi.org/10.1145/3197517.3201349
2. Totten GE, Liang H. Surface modification and mechanisms: friction, stress, and reaction engineering. CRC Press; 2004.
3. Rong MZ, Zhang M, Ruan WH. Surface modification of nanoscale fillers for improving properties of polymer nanocomposites: a review. Materials science and technology. 2006; 22(7):787–796. doi: https://doi.org/10.1179/174328406X101247
4. Dorsey J, Rushmeier H, Sillion F. Digital modeling of material appearance. Elsevier; 2010.
5. McCarthy W. Programmable matter. Nature. 2000; 407(6804):569. doi: https://doi.org/10.1038/35036656
6. Campbell TA, Tibbits S, Garrett B. The Programmable World. Scientific American. 2014; 311(5):60–65. Available from: https://www.jstor.org/stable/26041830
7. Campbell TA, Tibbits S, Garrett B. The next wave: 4D printing programming the material world. Atlantic Council, Washington, DC, Technical Report. 2014; p. 15 pages.
8. ASTM E284-17 Standard Terminology of Appearance. ASTM International, West Conshohocken, PA; 2017. Available from: https://doi.org/10.1520/E0284-17
9. CIE 175:2006 A framework for the measurement of visual appearance, International Commission on Illumination, 2006. 92 pages. ISBN: 978 3 901906 52 7
10. Hunter RS, Harold RW. The measurement of appearance. John Wiley & Sons; 1987.
11. Choudhury AKR. Principles of colour and appearance measurement: Object appearance, colour perception and instrumental measurement. Elsevier; 2014.
12. Ishimaru A. Wave propagation and scattering in random media and rough surfaces. Proceedings of the IEEE. 1991; 79(10):1359– 1366. doi: https://doi.org/10.1109/5.104210
13. Eugène C. Measurement of ”total visual appearance”: a CIE challenge of soft metrology. In: 12th IMEKO TC1 TC7 Joint Symposium on Man, Science Measurement; 2008. p. 61–65.

14. Hutchings J. The continuity of colour, design, art, and science. I. The philosophy of the total appearance concept and image measurement. Color Research & Application. 1995; 20(5):296–306. doi: https://doi.org/10.1002/col.5080200507
15. Hutchings J. The continuity of colour, design, art, and science. II. Application of the total appearance concept to image creation. Color Research & Application. 1995;20(5):307–312. doi: https://doi.org/10.1002/col.5080200508
16. Hutchings JB. Food color and appearance; 2nd edition. Aspen Publishers, New York, 1999.
17. Urban P, Tanksale TM, Brunton A, Vu BM, Nakauchi S. Redefining A in RGBA: Towards a Standard for Graphical 3D Printing. ACM Transactions on Graphics (TOG). 2019 (preprint available at https://arxivorg/abs/171000546); 38(3):21:1–21:15. doi: https://doi.org/10.1145/3319910
18. Hunter RS. Methods of determining gloss. NBS Research paper RP. 1937; 958:19–39.
19. Motoyoshi I. Highlight–shading relationship as a cue for the perception of translucent and transparent materials. Journal of Vision. 2010; 10(9:6):1–11. doi: https://doi.org/10.1167/10.9.6
20. Motoyoshi I, Nishida S, Sharan L, Adelson EH. Image statistics and the perception of surface qualities. Nature. 2007; 447(7141):206–209. doi: https://doi.org/10.1038/nature05724
21. Nicodemus FE. Directional reflectance and emissivity of an opaque surface. Applied optics. 1965; 4(7):767–775. doi: https://doi.org/10.1364/AO.4.000767
22. Nishida S, Shinya M. Use of image-based information in judgments of surface-reflectance properties. JOSA A. 1998; 15(12):2951– 2965. doi: https://doi.org/10.1364/JOSAA.15.002951
23. Xiao B, Walter B, Gkioulekas I, Zickler T, Adelson E, Bala K. Looking against the light: How perception of translucency depends on lighting direction. Journal of Vision. 2014; 14(3:17):1–22. doi: https://doi.org/10.1167/14.3.17
24. Chowdhury NS, Marlow PJ, Kim J. Translucency and the perception of shape. Journal of Vision. 2017;17(3:17):1–14. doi: https://doi.org/10.1167/17.3.17
25. Fleming RW, Bulthoff HH. Low-level image cues in the perception of translucent materials. ACM Transactions on Applied¨ Perception (TAP). 2005; 2(3):346–382. doi: https://doi.org/10.1145/1077399.1077409
26. Tanaka M, Nakayama D, Horiuchi T. Analysis of factors affecting the contrast effect for total appearance. Journal of the International Colour Association. 2020; Vol. (25):1–11.
27. Sharan L, Rosenholtz R, Adelson EH. Accuracy and speed of material categorization in real-world images. Journal of Vision. 2014; 14(9:12):1–24. doi: https://doi.org/10.1167/14.9.12
28. Fleming RW, Wiebel C, Gegenfurtner K. Perceptual qualities and material classes. Journal of Vision. 2013;13(8:9):1–20. doi: https://doi.org/10.1167/13.8.9
29. Sharan L, Liu C, Rosenholtz R, Adelson EH. Recognizing materials using perceptually inspired features. International Journal of Computer Vision. 2013; 103(3):348–371. doi: https://doi.org/10.1007/s11263-013-0609-0

30. Vu BM, Urban P, Tanksale TM, Nakauchi S. Visual perception of 3D printed translucent objects. In: Color and Imaging Conference. 1. Society for Imaging Science and Technology; 2016. p. 94–99. doi: https://doi.org/10.2352/ISSN.2169-2629.2017.32.94
31. Smet K, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. Colour appearance rating of familiar real objects. Color Research & Application. 2011; 36(3):192–200. doi: https://doi.org/10.1002/col.20620
32. Gigilashvili D, Thomas JB, Pedersen M, Hardeberg JY. Perceived Glossiness: Beyond Surface Properties. In: Color and Imaging Conference. 1. Society for Imaging Science and Technology; 2019. p. 37–42. doi: https://doi.org/10.2352/issn.2169-2629.2019.27.8
33. Paille P. L'analyse par théorisation ancrée. Cahiers de recherche sociologique. 1994; 1(23):147–181. doi: https://doi.org/10.7202/1002253ar
34. Glaser BG, Strauss AL, Strutzel E. The discovery of grounded theory; strategies for qualitative research. Nursing research. 1968;17(4):364.
35. Corbin J, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory. Thousand Oaks, CA: Sage publications; 1998.
36. Jacob JD, Holmes D. Working under threat: Fear and nurse–patient interactions in a forensic psychiatric setting. Journal of Forensic Nursing. 2011;7(2):68–77. doi: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1939-3938.2011.01101.x
37. Gaucher N, Payot A. From powerlessness to empowerment: Mothers expect more than information from the prenatal consultation for preterm labour. Paediatrics Child Health. 2011 12;16(10):638–642. doi: https://doi.org/10.1093/pch/16.10.638
38. Rippon D, McDonnell A, Smith M, McCreadie M, Wetherell M. A grounded theory study on work related stress in professionals who provide health social care for people who exhibit behaviours that challenge. PLOS ONE. 2020 02;15(2):1–23. doi: https://doi.org/10.1371/journal.pone.0229706
39. Thomas JB, Deniel A, Hardeberg JY. The Plastique collection: A set of resin objects for material appearance research. XIV Conferenza del Colore, Florence, Italy. 2018; p. 12 pages.
40. Gigilashvili D, Thomas JB, Hardeberg JY, Pedersen M. Behavioral investigation of visual appearance assessment. In: Color and Imaging Conference. 1. Society for Imaging Science and Technology; 2018. p. 294–299. doi: https://doi.org/10.2352/ISSN.2169-2629.2018.26.294
41. Gigilashvili D, Thomas JB, Pedersen M, Hardeberg JY. Material appearance: ordering and clustering. In: Material Appearance 2019, IS&T International Symposium on Electronic Imaging. Society for Imaging Science and Technology; 2019. p. 202:1–202:6. doi: https://doi.org/10.2352/ISSN.2470-1173.2019.6.MAAP-202
42. Sharan L, Rosenholtz R, Adelson E. Material perception: What can you see in a brief glance? [Abstract]. Journal of Vision. 2009; 9(8):784. doi: https://doi.org/10.1167/9.8.784
43. Koren M. What Color Is a Tennis Ball? An investigation into a surprisingly divisive question. The Atlantic. Accessed on 23/02/2020 at https://wwwtheatlanticcom/science/archive/2018/02/what-color-tennis-ball-green-yellow/523521/ .

44. Bucknell SP, Winsey NJP, Gale DR. Translucency measurement. Google Patents; 2000. US Patent 6,111,653.
45. Tirpak A, Young R. Accurate Transmission Measurements of Translucent Materials-The light scattering of translucent materials means that two sources are needed for the best transmission measurements. Photonics Spectra. 2008;42(2):1–6.
46. Gigilashvili D, Mirjalili F, Hardeberg JY. Illuminance Impacts Opacity Perception of Textile Materials. In: Color and Imaging Conference. Society for Imaging Science and Technology; 2019. p. 126–131. doi: https://doi.org/10.2352/issn.2169-2629.2019.27.24
47. Thomas JB, Colantoni P, Trémeau A. On the Uniform Sampling of CIELAB Color Space and the Number of Discernible Colors. In: Tominaga S, Schettini R, Trémeau A, editors. Computational Color Imaging. Berlin, Heidelberg: Springer Berlin Heidelberg;´ 2013. p. 53–67. doi: https://doi.org/10.1007/978-3-642-36700-7_5
48. Deborah H. Towards Spectral Mathematical Morphology [PhD thesis]. Norwegian University of Science & Technology, University of Poitiers; 2016. Available from: http://goo.gl/AsfCjj
49. Gigilashvili D, Urban P, Thomas JB, Hardeberg JY, Pedersen M. Impact of Shape on Apparent Translucency Differences. In: Color and Imaging Conference. Society for Imaging Science and Technology; 2019. p. 132–137. doi: https://doi.org/10.2352/issn.2169-2629.2019.27.25
50. Fleming RW. Visual perception of materials and their properties. Vision research. 2014; 94:62–75. doi: https://doi.org/10.1016/j.visres.2013.11.004
51. Todd JT, Norman JF. Reflections on glass. Journal of Vision. 2019; 19(4:26):1–21. doi: https://doi.org/10.1167/19.4.26.
52. Filip J, Kolafová M, Havlıček M, Vávra R, Haindl M, Rushmeier H. Evaluating physical and rendered material appearance. The Visual Computer. 2018;34(6-8):805–816. doi: https://doi.org/10.1007/s00371-018-1545-3
53. Alley LM, Schmid AC, Doerschner K. Expectations affect the perception of material properties. bioRxiv. 2019; p. 36 pages. doi: https://doi.org/10.1101/744458
54. Tanaka M, Horiuchi T. Investigating perceptual qualities of static surface appearance using real materials and displayed images. Vision research. 2015; 115:246–258. doi: https://doi.org/10.1016/j.visres.2014.11.016
55. Fiore AM. Multisensory integration of visual, tactile, and olfactory aesthetic cues of appearance. Clothing and Textiles Research Journal. 1993;11(2):45–52. doi: https://doi.org/10.1177/0887302X9301100207
56. Liu H, Yu Y, Sun F, Gu J. Visual–tactile fusion for object recognition. IEEE Transactions on Automation Science and Engineering. 2016; 14(2):996–1008. doi: http://doi.org/10.1109/TASE.2016.2549552
57. Liu H, Sun F. Material identification using tactile perception: A semantics-regularized dictionary learning method. IEEE/ASME Transactions on Mechatronics. 2017; 23(3):1050–1058. doi: https://doi.org/10.1109/TMECH.2017.2775208

58. Fujisaki W, Goda N, Motoyoshi I, Komatsu H, Nishida S. Audiovisual integration in the human perception of materials. Journal of Vision. 2014; 14(4:12):1–20. doi: https://doi.org/10.1167/14.4.12
59. Wongsriruksa S, Howes P, Conreen M, Miodownik M. The use of physical property data to predict the touch perception of materials. Materials & Design. 2012; 42:238–244. doi: https://doi.org/10.1016/j.matdes.2012.05.054
60. Vangorp P, Laurijssen J, Dutre P. The influence of shape on the perception of material reflectance. In: ACM Transactions on´ graphics (TOG). vol. 26. ACM; 2007. p. 77:1–77:10. doi: https://doi.org/10.1145/1275808.1276473
61. Olkkonen M, Brainard DH. Joint effects of illumination geometry and object shape in the perception of surface reflectance. i-Perception. 2011; 2(9):1014–1034. doi: https://doi.org/10.1068/i0480
62. Ho YX, Landy MS, Maloney LT. Conjoint measurement of gloss and surface texture. Psychological Science. 2008; 19(2):196–204. doi: https://doi.org/10.1111/j.1467-9280.2008.02067.x
63. Marlow PJ, Kim J, Anderson BL. The perception and misperception of specular surface reflectance. Current Biology. 2012; 22(20):1909–1913. doi: https://doi.org/10.1016/j.cub.2012.08.009
64. Schmid AC, Anderson BL. Do surface reflectance properties and 3-D mesostructure influence the perception of lightness? Journal of Vision. 2014; 14(8:24):1–24. doi: https://doi.org/10.1167/14.8.24
65. Fleming RW, Dror RO, Adelson EH. Real-world illumination and the perception of surface reflectance properties. Journal of Vision. 2003; 3:347–368. doi: https://doi.org/10.1167/3.5.3
66. Kim J, Marlow PJ. Turning the world upside down to understand perceived transparency. i-Perception. 2016; 7(5):1–5. doi: https://doi.org/10.1177/2041669516671566
67. Gerardin M, Simonot L, Farrugia JP, Iehl JC, Fournel T, Hébert M. A translucency classification for computer graphics. In: Material Appearance 2019, Electronic Imaging. Society for Imaging Science and Technology; 2019. p. 203:1–203:6. doi: https://doi.org/10.2352/ISSN.2470-1173.2019.6.MAAP-203
68. Plastics — Determination of the total luminous transmittance of transparent materials — Part 1: Single-beam instrument. Geneva, CH: International Organization for Standardization; 2019.
69. Gilchrist A, Kossyfidis C, Bonato F, Agostini T, Cataliotti J, Li X, et al. An anchoring theory of lightness perception. Psychological review. 1999; 106(4):795–834.
70. Land EH, McCann JJ. Lightness and retinex theory. JOSA. 1971; 61(1):1–11. doi: https://doi.org/10.1364/JOSA.61.000001
71. Marlow PJ, Kim J, Anderson BL. Perception and misperception of surface opacity. Proceedings of the National Academy of Sciences. 2017; 114(52):13840–13845. doi: https://doi.org/10.1073/pnas.1711416115
72. Yu H, Liu PX, Hu L. A Highlight-Generation Method for Rendering Translucent Objects. Sensors. 2019; 19(4):860:1–860:15. doi: https://doi.org/10.3390/s19040860
73. Schmid AC, Barla P, Doerschner K. Material category determined by specular reflection structure mediates the processing of image features for perceived gloss. bioRxiv. bioRxiv preprint doi: https://doiorg/101101/20191231892083 , 2020; p. 45 pages.

74. Gigilashvili D, Dubouchet L, Pedersen M, Hardeberg JY. Caustics and Translucency Perception. In: Material Appearance 2020, IS&T International Symposium on Electronic Imaging. Society for Imaging Science and Technology; 2020. p. 033:1–033:6. doi: https://doi.org/10.2352/ISSN.2470-1173.2020.5.MAAP-033
75. Davit Gigilashvili, Weiqi Shi, Zeyu Wang, Marius Pedersen, Jon Yngve Hardeberg, and Holly Rushmeier (2021). "The Role of Subsurface Scattering in Glossiness Perception." ACM Transactions on Applied Perception. 2021; 18(3): 10:1-10:26. doi: https://doi.org/10.1145/3458438
76. Pellacini F, Ferwerda JA, Greenberg DP. Toward a psychophysically-based light reflection model for image synthesis. In: Proceedings of the 27th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co.; 2000. p. 55–64. doi: https://doi.org/10.1145/344779.344812
77. Qi L, Chantler MJ, Siebert JP, Dong J. Why do rough surfaces appear glossy? JOSA A. 2014; 31(5):935–943. doi; https://doi.org/10.1364/JOSAA.31.000935
78. Sakano Y, Ando H. Effects of head motion and stereo viewing on perceived glossiness. Journal of Vision. 2010; 10(9:15):1–14. doi: https://doi.org/10.1167/10.9.15
79. Chadwick AC, Kentridge R. The perception of gloss: A review. Vision research. 2015; 109:221–235. doi: https://doi.org/10.1016/j.visres.2014.10.026
80. Doerschner K, Fleming RW, Yilmaz O, Schrater PR, Hartung B, Kersten D. Visual motion and the perception of surface material. Current Biology. 2011;21(23):2010–2016. doi: https://doi.org/10.1016/j.cub.2011.10.036
81. Wendt G, Faul F. Can color and motion information be used to disentangle the influence of multiple light sources on gloss perception? i-Perception. 2018; 9(5):1–26. doi: https://doi.org/10.1177/2041669518803964
82. Fujisaki W, Tokita M, Kariya K. Perception of the material properties of wood based on vision, audition, and touch. Vision research. 2015; 109:185–200. doi: https://doi.org/10.1016/j.visres.2014.11.020
83. De Camp JE. The influence of color on apparent weight. A preliminary study. Journal of experimental psychology. 1917;2(5):347– 370. doi: https://doi.org/10.1037/h0075903
84. Bullough E. On the apparent heaviness of colours. A contribution to the aesthetics of colour. British Journal of Psychology. 1907; 2(2):111–152.
85. Walker P, Francis BJ, Walker L. The brightness-weight illusion. Experimental psychology. 2010; 57(6):462–469. doi: https://doi.org/10.1027/1618-3169/a000057
86. Walker P. Cross-sensory correspondences and naive conceptions of natural phenomena. Perception. 2012; 41(5):620–622. doi: https://doi.org/10.1068/p7195
87. Hagtvedt H. Dark is durable, light is convenient: Color value influences perceived product attributes. ACR North American Advances. 2014; 42:27–31.
88. Chandler DM. Seven challenges in image quality assessment: past, present, and future research. ISRN Signal Processing. Article ID 905685, 2013;p. 53 pages. doi: http://dx.doi.org/10.1155/2013/905685

Appendix 1

Different ways to display stimuli in appearance research

There are three ways to generate the visual stimuli: direct view to the real physical objects, photographing the real objects, and using computer graphics to generate synthetic images. However, the ways to present them to the observer are two: either present the object directly, or to display it through an intermediate medium - e.g. computer display or VR headset. By presenting the stimulus on an intermediary display the dimensionality of the stimulus reduces (e.g. from infinite dimensions in a natural scene to 5D in 2D displayed color image). Therefore, the way of stimuli introduction should be carefully chosen. The advantages and disadvantages of different methods for displaying the stimuli is summarized in the table below.

Table 1. Advantages and disadvantages of using tangible and displayed stimuli.

	Advantages	Disadvantages
Physical Objects	• Subjects can freely interact with the physical objects - i.e. possibility to apply all behavioral patterns we use in our daily lives for appearance assessment (move head, move object). • Multisensory information is present (e.g. tactile, auditory). • Binocular vision. • Realistic environment. • Artifacts make objects realistic. • In the real world we have access to full scene context that is often not possible in graphics.	• Difficult to model, measure, and replicate. • High cost of manufacturing. • Unpredictable effects of aging. • Unwanted artifacts. • Risk of damaging. • Limited access across the scientific community. • Limited reproducibility of the experiments (due to access, aging).
Displayed Images	• Full control of the material parameters (e.g. phase function, absorption and scattering) and scene (illumination, background). • Simplicity of manipulation of any material or scene parameters. • Relatively low cost of production/generation. • Better reproducibility. • Easier to share the data across the scientific community. • Realistic photographs can be used. • Free from aging effects.	• Graphic rendering is based on a model that might be limited and might significantly impact result of the experiment. Physically based rendering is extremely time-consuming. • It is very difficult to relate a radiate image and stimuli to the optical model due to digitization of the information and calibration of the display. If it is relative to display (and full calibration, even though might be reproducible), it is still not correlated to the optical model. • Many factors, like resolution, color gamut or heterogeneity of the display might impact the results. • Dynamic range of the displays are lower. • Interactivity is limited in computer graphics. • Multisensory information is absent, or extremely limited. • Often no stereo vision is possible. • The environment is often unrealistic in computer graphics (e.g. neutral gray background). • No virtual system replicates fully the complex lighting environments we
Virtual Reality	• All display-related advantages apply to VR as well. • VR might enable binocularity and motion. • More realistic interactivity than in case of displays. • Not affected by the ambient illumination. • Less distraction from the ambience.	

Advantages	Disadvantages
	encounter in real lives, especially characterizing directional spectral variation in natural environments. • Lack of imperfections in computer graphics not only reduce naturalness of the stimuli, also undermines robustness of the models built based on them. • While photographs are realistic and superior to synthetic stimuli in several above-mentioned aspects, they do not contain the information regarding the physical material properties, and we are limited to image statistics extraction.

Appendix 2

An example of the observations, with the transcript, the action performed and their interpretation within the model

We introduced 15 categories that unify conceptually similar observations. Afterwards, we also presented the qualitative model that not only shows how the categories relate with one another, but also explains the entire pipeline of the material appearance assessment in context of our tasks. At first glance, it might be ambiguous in what way the videotaped experiment is processed using the Grounded Theory Analysis. In order to illustrate exactly how the model is rooted in the data, below we present a detailed transcript of the 6.5-minute excerpt from the actual experiment where the observer tries to rank five spheres by glossiness (refer to Table 2). The first column shows the time frame (in mm:ss format) the comments in the corresponding row are referring to. The second column contains the speech from a given time frame - either quoted, or paraphrased. The third column describes the actions happening within a given time frame. The fourth column comments the content and explains the process in context of our model.

Table 2. An example task transcript illustrating how the model describes the data.

Time	Speech	Action	Comment
00:00 to 00:20	The experimenter introduces objects to the observer.	The experimenter puts objects in front of the observer.	***Object*** enters the scene under given ***Conditions of Observation***. The ***Experimenter*** starts impacting the process.
00:20 to 00:30		The observer starts inspecting the objects.	The appearance perception is evoked by the combination of two factors: characteristics of the ***Object***, and the ***Conditions of Observation***, like illumination conditions.
00:30 to 00:45	The experimenter says that as the observer has got used to this dataset, he can again describe them by appearance.		The ***Experimenter*** contributes to ***Task Interpretation***. The experimenter means that the observer has already seen similar objects in previous tasks, and

			Learning and Adaptation facilitates the process.
00:45 to 01:20	Observer describes: ”even without taking them and looking through them towards the sun, which is an usual way for translucency, even without that, I see that this is yellowish and very translucent, these are opaque, opaque I do not know color, bluish and somewhat translucent, orange and very translucent”.	Observer moves his head to the sides while examining objects. Points one by one to the caustics of the objects with an index finger, while describing the appearance. The judgement is based solely on the caustic pattern projected onto the table.	The observer has come up with a particular ***Methodology*** (that involves assessment of the caustic pattern). He needs a ***Reference*** for ***Comparison***. In this case, he compares appearance of the two objects between the two observation geometries (when moving the head), where the ***Reference*** is the appearance in normal sitting condition that is compared with the appearance of the same object seen with a head tilted to the side. For ***Semantic Description***, the observer needs ***Vocabulary Search***. His professional background in material appearance is a ***Condition of Observation*** that contributes to his ***Methodology*** and ***Vocabulary Search***, coming up with a particular ***Vocabulary*** that is composed of appearance attribute terminology related to color, and light transmittance properties. When exact word was not found with ***Vocabulary Search***, the ***Comparison*** with the nearest ***Reference*** is used to express uncertainty, like words ”yellowish”, ”bluish”, and ”somewhat translucent”.
01:25 to 01:35	The experimenter asks: ”so, you put them against light, so you can see the shadow in front of you as a color palette?”. The observer confirms.		The ***Experimenter*** clarifies the ***Task Interpretation*** and selected ***Methodology***.
01:35 to 01:47	The observer continues description: ”well, they are pretty glossy. No texture, they have all spherical shape”.	The observer moves his head to the sides, looks from the top to observe the image in the reflections.	The ***Vocabulary*** is still strongly impacted by the ***Conditions of Observation*** - the background of the observer, and the illumination conditions in the room. The observer continues using ***Comparison*** between two observation geometries.
01:47 to 02:28	The observer continues description: ”I see some kind of artifacts. Here the scratches are deeper. This one has more severe artifacts. Apart from	Observer picks one object and looks closely. Then picks the next one.	As the time passes, ***Learning and Adaptation*** helps the observer to include more details in the ***Semantic Description***.

	artifacts, they are all glossy, those three are translucent, those two are opaque. They differ in color, yellowish, this is kind of yellow too, orange, dark blue, light blue. ”		
02:28 to 02:45	The experimenter introduces the visual task: ”now I will ask you a very specific task. Rank them by glossiness again. As you said, they are very glossy, so it might be more difficult. ”		The ***Task*** is presented. ***Experimenter*** conveys the message and the observer starts ***Task Interpretation***.
02:45 to 02:55	”That’s true” - the observer admits the task is difficult.	The observer picks two objects up, and looks at them from the side, holding them next to each other.	The observer has ***Structure Expectation***. The task is considered ”difficult”, because the observer assumes the ranking should be possible and there is the ”right answer”, even though all objects look ”very glossy”. This impacts the rest of ***Task Interpretation***. After ***Task Interpretation***, the observer has taken his time for ***Decision-Making*** and came up with a ***Methodology*** (that will be refined over time due to ***Learning and Adaptation***). The observer clearly needs a ***Reference*** for ***Comparison*** to quantify appearance of a particular object. So, he picks two objects and compares them against each other.
02:55 to 03:16	Experimenter gives further instructions: ”one thing you could consider is artifacts, if you can’t find any other difference; but, first of all, I want to ask you to classify without taking them into account.”		This is a pure improvisation by the ***Experimenter*** that impacts ***Task Interpretation*** and further ***Decision-making***.
03:16 to 03:21		The observer continues picking pairs of objects and inspecting them. Comparing each other.	***Comparison*** with a ***Reference***.
03:26 to 03:31		The observer puts two spheres next to the third one, and compares the three.	***Comparison*** with a ***Reference***.
03:31 to 03:41		The observer moves his hand atop the objects, and looks at the reflections.	New details appear in selected ***Methodology***. In addition to picking objects up and comparing them, the observer starts a

			different kind of ***Comparison*** with a different ***Reference*** - he compares reflection image on the same sphere among several conditions - among several positions of his hand. According to the selected ***Methodology***, better the hand movement is depicted in the surface reflection image, glossier the object.
	Observer: ”this is I think the most glossy one, without considering the artifacts.”	The observer picks the dark blue object and examines from close. Then puts it on the right hand side of the table, as being ranked most glossy.	The ***Semantic Description*** is regularly used for ***Completion of a Visual Task***.
03:41 to 03:46		Puts his hand close to the sphere surface and observes closely. Then puts the blue one next to the one ranked first. Then chooses the third one.	
03:46 to 04:02	The observer explains his decisions: ”these specular reflections look the same on all of them. Except for the damaged areas. The way I am going to classify them is whether I see myself on them. Whether it has a mirror effect or not. ”		The observer explains the ***Methodology***, and the ***Decision-making*** process that lead him to this particular ***Methodology***.
04:02 to 04:08	Experimenter: ”so you are not using specular effect, but how you can use them as a mirror.”		The ***Experimenter*** clarifies the ***Task Interpretation*** and selected ***Methodology***.
04:08 to 04:18	Observer: ”yes, I tried to use specular reflections, but they all look the same.”		The combination of ***Object*** and ***Conditions of Observation*** have impacted ***Methodology*** selection.
04:18 to 04:28		The observer blocks direct sunlight with his hands towards two translucent spheres, and looks at them in the shadow. Then picks them up and inspects closely.	Again, ***Comparison*** with a ***Reference*** in several conditions.
04:28 to 04:35		The observer takes decision one of them is glossier. Puts it on the fourth place, while the last one is put on the fifth place.	The ***Comparison*** with a ***Reference*** using particular ***Methodology*** leads to ***Visual Task Completion***.

04:35 to 04:40	Experimenter: artifacts would have changed this order, or not?		
04:40 to 06:14	The observer explains the process: ”it depends how you look at it. At first, I did not pay attention to them, because I know they are not intended to be there. So, I judged just the normal part. But between this two”, - points to the last two ones - ”when I did not have any other choice, because I couldn’t use them as a mirror, and specular reflections are same, so I look at them and decided which one has more damaged areas that reflects less light. It’s very very last cue, I looked specular reflections first of all, but they are the same. Then I saw my gloves on this one [most glossy one], here it’s a bit blurry second and third ones]”, - moves his hand atop the object. ”And here [two least glossy ones] very little bit. Here (first two ones), I even see my face, while here [last two ones], I just see my gloves when I bring it very close to the surface.”	Picks the two objects again and shows the areas, which do not reflect in a specular direction due to scratches.	The observer explains the ***Methodology***, and the ***Decision-making*** process that lead him to this particular ***Methodology***. Also names particular **References** used.
06:14 to 06:30		The experimenter thanks the observer, the result is recorded (photographed),and they switch to a new task.	

Article C

Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (2019). “Perceived Glossiness: Beyond Surface Properties.” In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 37–42

Perceived Glossiness: Beyond Surface Properties

Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, Jon Yngve Hardeberg;
Department of Computer Science, Norwegian University of Science and Technology; Gjøvik, Norway

Abstract

Gloss is widely accepted as a surface- and illumination-based property, both by definition and by means of metrology. However, mechanisms of gloss perception are yet to be fully understood. Potential cues generating gloss perception can be a product of phenomena other than surface reflection and can vary from person to person. While human observers are less likely to be capable of inverting optics, they might also fail predicting the origin of the cues. Therefore, we hypothesize that color and translucency could also impact perceived glossiness. In order to validate our hypothesis, we conducted series of psychophysical experiments asking observers to rank objects by their glossiness. The objects had the identical surface geometry and shape but different color and translucency. The experiments have demonstrated that people do not perceive objects with identical surface equally glossy. Human subjects are usually able to rank objects of identical surface by their glossiness. However, the strategy used for ranking varies across the groups of people.

Introduction

Appearance is a complex psychovisual phenomenon that is defined as ”the visual sensation through which an object is perceived to have attributes as size, shape, colour, texture, gloss, transparency, opacity, etc.” [1] Due to its multiplex nature appearance is usually split into distinct attributes. According to CIE, there are four major appearance attributes: color, gloss, translucency and texture [1, 2]. Eugène [3] cites CIE definition of gloss as: ”the mode of appearance by which reflected highlights of objects are perceived as superimposed on the surface due to the directionally selective properties of that surface” and adds that ”gloss perception is particularly depending on the way that light is reflected from the surface of the object at and near the specular direction.” [1] ASTM Standard Terminology of Appearance [4] defines gloss as ”angular selectivity of reflectance, involving surface-reflected light, responsible for the degree to which reflected highlights or images of objects may be seen as superimposed on a surface.” In computer graphics the Phong reflection model [5] (that is a simplification of bidirectional reflectance distribution function - BRDF) is widely used to model glossy appearance. The component responsible for this effect is the ratio of specularly reflected and incident light. However, the model does not account for transmission or subsurface scattering and no translucency is considered. Ho *et al.* [6] have demonstrated correlation between perceived glossiness and perceived bumpiness, describing gloss as a ”surface property”, while Hunter [7] distinguishes six different types of gloss: 1. **Specular gloss** - ”identified by shininess”; 2. **Sheen** - ”identified by surface shininess at grazing angles”; 3. **Contrast gloss** - ”identified by contrasts between specularly reflecting areas of surfaces and other areas”; 4.**Absence-of-bloom gloss** - ”identified by the absence of reflection haze or smear adjacent to reflected high lights”; 5. **Distinctness-of-reflected-image gloss** - ”identified by the distinctness of images reflected in surfaces”; 6. **Absence-of-surface-texture gloss** - ”identified by the lack of surface texture and surface blemishes.” He proposes that glossiness might be correlated with surface specular reflectance and concludes that reflectance distribution functions *”offer the only means by which the reflectance properties of surfaces responsible for their glossiness may be completely specified.”* On the other hand, Motoyoshi *et al.* [8] propose that simple image statistics, like skewness of luminance histogram or similar metric of histogram assymetry, are used by the human visual system to assess surface properties and glossiness without knowledge of the reflectance distribution function [9]. The authors explicitly mention gloss as a surface-related property without discussing the possibility that the histogram might be affected by transmission or sub-surface scattering of the light. They further conclude that average luminance has a significant impact on perceived lightness, but not on perceived glossiness and demonstrate the two images of Michelangelo’s St Matthew sculpture that have identical mean luminance but substantially differ in perceived glossiness, while comparing grayscale images of the opaque surfaces. Nishida and Shin’ya [10] propose that a combination of mean luminance, luminance contrast, maximum and minimum luminance, as well as spatial structure of luminance gradients, might be cues for perception of surface properties. They also demonstrate that surface-reflectance constancy of the human visual system fails when shape is changed. Chowdhury *et al.* [11] have shown that perceived mesoscopic shape differs between translucent and opaque objects due to difference in luminance gradients.

Pellacini *et al.* [12] have explored dimensionality of gloss perception, introducing a percetually uniform gloss space and psychophysically-based light reflection model that should enable cross-object description and matching of apparent gloss. Using mutlidimensional scaling the authors came up with a 2-dimensional space with orthogonal axes that are *”qualitatively similar to the contrast gloss and distinctness-of-image gloss attributes”*. They also claim that CIELAB ligthness parameter impacts apparent gloss and demonstrated that *”apparent gloss is*

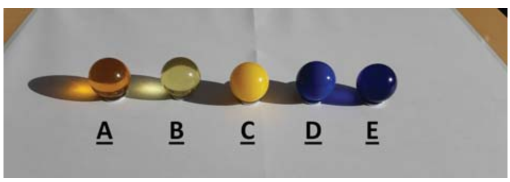

Figure 1. The objects used for the preliminary experiment.

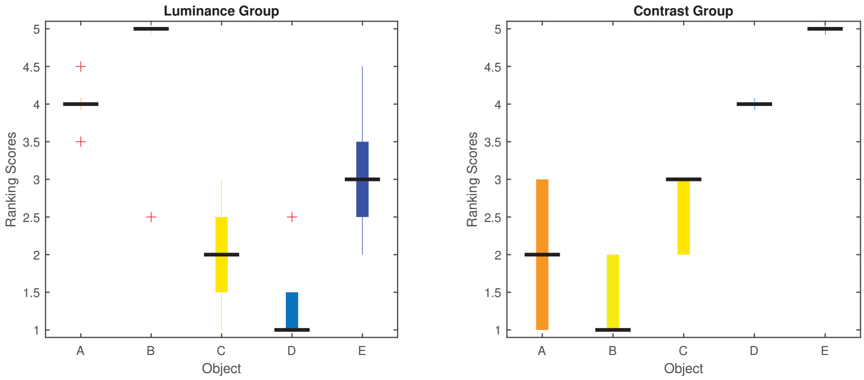

Figure 2. *Boxplots for observer scores showing how observers ranked the five (A, B, C, D, and E) objects (Figure 1). 1 means least glossy, while 5 means most glossy. In case of ties, the mean score was taken. Central mark -median; bottom and top edges - 25^{th} and 75^{th} percentiles, respectively; Whiskers extend to the extreme data points excluding outliers; red '+' symbol - outliers. We can observe clear separation for both groups.*

affected by the diffuse reflectance of a surface, with light colored surfaces appearing less glossy than dark ones having the same finish". Although the proposed framework performed well for their dataset, the study is limited to opaque spherical objects assuming that chromaticity and apparent gloss are independent, without mention of any possible impact from translucency.

The paper is organized as follows: in the next section background information is provided. Afterwards, we conduct detailed analysis of the first experiment [13] followed by the experimental setup of the new one. Subsequent section covers results and discussion. Finally, we conclude and outline the future work.

Background and Motivation

In an earlier paper [13] we summarized a psychophysical experiment where observers were asked to rank five spheres by their glossiness which had identical surface smoothness but different color and translucency (Figure 1). Aggregate frequency analysis did not show statistically significant differences in observer scores, making us hypothesize that similar gloss perception can be achieved with similar surface smoothness, but more thorough insight into the interviews of the observers has outlined three groups of people of roughly same size: 1. Subjects who considered all spheres to be equally glossy; 2. Subjects who ranked the spheres considering translucent ones more glossy. Those people mentioned shininess of the translucent spheres as the reason for their apparent glossiness. In this case brightness was the cue for them; 3. Subjects who ranked the spheres considering opaque ones more glossy. Those observers used distinctness-of-image gloss and contrast gloss (for the dark ones) as a cue. The three groups used different cues to reach the conclusion, and some of those cues may be impacted by other material properties, not only the shape and surface geometry.

In this paper we want to challenge the established opinion that gloss perception is solely surface-based quality. While translucency and color can contribute significantly to the cues like mean luminance as well as luminance contrast and luminance histogram, associated with perceived gloss in the literature [8, 10], it has been proposed [8, 14] that the human visual system has poor ability, if any, to invert the optics. Therefore, we propose that translucency and color, particularly lightness, have significant impact on perceived glossiness. Translucency is a point of particular interest due to two reasons: first of all, light transmission and back-reflections increase overall luminance and shininess of the object that might be consciously or subconsciously associated with gloss; and secondly, caustics could play significant role too. According to Lynch [15], caustic is "three dimensional envelope of imperfectly focused rays" or "two-dimensional pattern formed when a caustic falls on a surface." Internal and external caustics and the glittering effect of the caustic highlights might be mistaken for specular highlights and thus, for gloss, considering their similarity in luminance, and proposedly poor optics inversion ability of the human visual system. We conducted series of psychophysical experiments asking people to rank objects by their glossiness. The objects had nearly identical surface smoothness but different color and translucency. As the observers were explicitly instructed that they could have ties among objects including tying all of them, if our hypothesis is false and perceived glossiness depends solely on the surface geometrical properties, the vast majority of them should have said that all objects have the same glossiness. In the previous paper [13] different cues used by subjects in opaque and translucent spheres compensated each other leading to statistically insignificant difference among perceived glossiness when analyzed the aggregated data. In order to clear up this ambiguity, we: 1. Analyzed the data from the first experiment [13] separately for different groups of people. 2. Replaced spheres with a complex object shape that decreases predictability of caustics and makes it impossible to observe distinctness-of-image gloss. As the cross-shape failure of reflectance constancy has been shown in [10], we used objects with an identical shape.

Group-based analysis of the first experiment

The first experiment using five spheres is discussed in [13]. The observers were asked to rank five spheres by their glossiness. Although the spheres had different colors and translucency, surface geometry among them was nearly identical. While aggregate analysis of the overall data did not illustrate statistically significant differences in perceived glossiness, more thorough insight in the data revealed three different groups of the people using different strategies. Below we will illustrate group-based analysis of the data. The spheres used in the experiment are shown in Figure 1. 17 observers participated in the experiment. Six observers concluded that all spheres have the same glossiness; six people used luminance-based strategy (later referred as ”luminance group”), and five people used distinctness-of-image gloss or contrast gloss-based strategy (”contrast group”). The boxplots for the latter two groups are illustrated in Figure 2. Due to low number of tests, it is difficult to assess statistical significance of the differences. However, the boxplots show very interesting trends. The ”luminance group” has a very clear separation between shiny transparent A and B spheres, and opaque C and D spheres. The dark blue but semi-transparent sphere E has overlaps with both groups as it demonstrates characteristics of the both. On the other hand, for the ”contrast group” there is a clear separation between A, B, and C spheres on the one hand, and D and E spheres, on the other hand. Dark blue and fully or significantly opaque spheres are considered more glossy, because this group of the subjects used a combination of distinctness-of-image gloss and contrast gloss that are stronger than in case of translucent or opaque but very light yellow spheres. Nevertheless, it is impossible to draw solid conclusions due to low number of subjects and test objects. We conducted a second experiment to verify the results.

Experimental Setup & Methodology

Task and Stimuli

The subjects were introduced to nine plastic female sculpture objects placed on an A3 white paper with a printed scale and two extremes: *”Least Glossy”* and *”Most Glossy”* points. Afterwards the following instruction was given: *”Please, rank the objects by their glossiness: from the most glossy to the least glossy. You can have any number of ties, including the case, when all objects are tied and no ranking is possible.”* The observers were allowed to interact with the objects, touch and move them freely. No explicit definition has been given for gloss. However, they were allowed to check the definition in case of uncertainty. We used a subset of the *Plastique* artwork collection [16]. The collection has been created by **an independent artist** Aurore Deniel to support research on material appearance. The samples are illustrated on Figure 3.

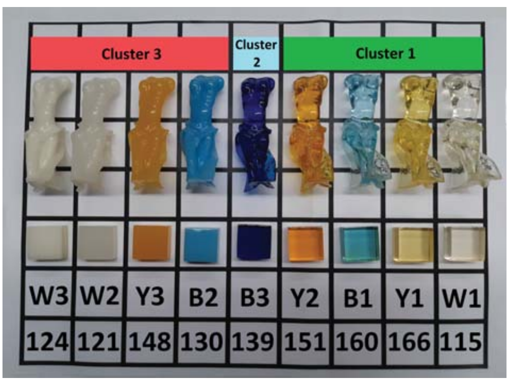

Figure 3. *The female bust objects used for the the experiment. The corresponding 2-symbol codes are for reference purposes only. 3-digit codes are their IDs used by Thomas et al. Cuboid objects have been used for transmittance and relative radiance measurements discussed below. [16].*

Experimental Conditions

We made an assumption that impact of the illumination conditions is less than that of cross-individual differences. Psychophysical experiments have been conducted on several occasions in controlled and uncontrolled conditions, and similar trends have been revealed under all conditions. In total, 107 observers participated in the experiments. 7 experiments were conducted in uncontrolled conditions, namely: 1. 2018 Color and Imaging Conference, Demonstration Session (8 observers, attendees of the conference); 2. 2019 IS&T International Symposium on Electronic Imaging, Demonstration Session (17 observers, attendees of the conference); 3. Material Appearance 2019 Conference (8 observers, attendees of the conference); 4. Internal academic activity at the Norwegian University of Science and Technology (NTNU), Trondheim (5 observers, master and PhD students); 5. Internal academic activity at NTNU, Gjøvik (11 observers, high school students); 6.Internal academic activity at NTNU, Gjøvik (7 observers, bachelor students); 7. Internal academic activity at NTNU, Gjøvik (7 observers, bachelor, master and PhD students). In addition, two experiments took place in controlled conditions, in two different viewing booths with a distance of roughly 50 cm: 8. VeriVide Color Assessment Cabinet 60-5 under D65 illumination with 1392 lux and 6180K color temperature (30 observers of mixed backgrounds). 9. GretagMacbeth Spectralight III viewing booth under Ultralume 30 (U30) illumination with 665 lux and 2865K color temperature (14 observers of mixed backgrounds). The experiments were anonymous and no further demographic information has been collected.

Analysis of the Collected Data

The rank order of the object is recorded as a numerical value. For instance, if the object was ranked most glossy, it was assigned ”1”; in case it was ranked second most glossy, the object was assigned ”2”, and so on. In case of ties, a mean score was assigned to all objects. For example, if the second and third objects were tied, each objects got rank equal to 2.5. If no ranking was done, each object was assigned ”5”. For visualization’s sake, results of similar ranking strategies were grouped together, and the ranks given to the each object by different observers were plotted as a graph to visualize the variation of a position for a particular object among different trials (Figure 4). Besides, the rank scores for each object are illustrated as box-plots (Figure 5). An alternative method for analyzing the ranking could be considering each experiment a pair-comparison among all objects, where selected object gets 1, the other one gets 0, and both objects get 0.5 in case of a tie. As the both methodologies lead to nearly identical results, we report the former for consistency’s sake with [13]. Afterwards, k-means clustering was conducted using MATLAB *kmeans()* function [1] to identify which objects were ranked together. That could help us to identify the right attributes that

[1]MATLAB R2017b version.

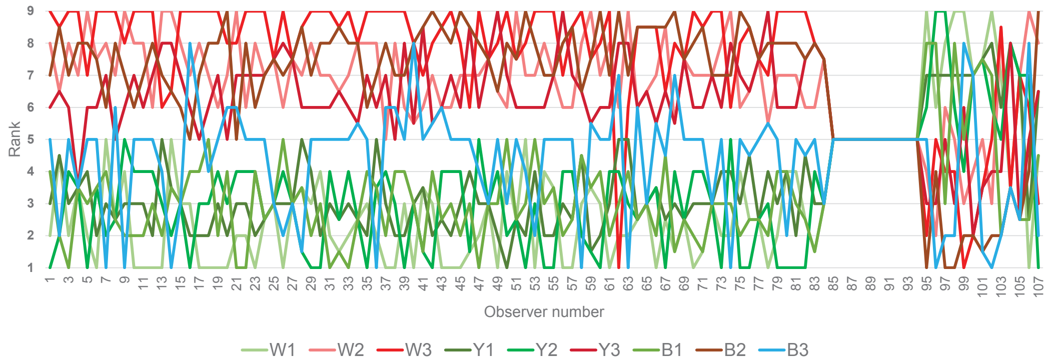

Figure 4. *The aggregate results from all individual experiments. Each colored line corresponds to a particular object. For the majority of the subjects, we can see a clear separation between more transparent (marked with green hue lines), and more opaque objects (marked red hue lines).*

made observers rank objects in a similar manner. Observations in this case were nine objects and variables were 107 ranks from 107 experiments. The cluster was defined as the centroid being the mean of all points in that particular cluster. Maximum number of iterations was set to 1000. Cluster centroids were initialized using *k-means++ algorithm* [17]. Finally, material luminance has been measured and correlated with mean ranking scores.

Results & Discussion

Graph Results

Identically to our previous experiment, three different ranking strategies have been observed:

1. 10 people (9.35%) mentioned that gloss was identical among objects, and thus, considered ranking impossible.
2. 84 people (78.50%) ranked more transparent objects over the ones closer to opacity.
3. 8 people (7.48%) opted for the objects closer to opacity.

The ranking of five people (4.67%) did not fit in any of the above-mentioned categories. It is worth mentioning that the trend has been similar in all illumination conditions. Clusters of the objects ranked similarly by each group of the people is further substantiated below by *kmeans clustering* results. The overall results are illustrated in Figure 4. The graphs for transparent objects are coded with the greenish hue, while the ones with more opacity are represented by reddish hue, and the dark blue translucent object that stands out from the rest of the dataset is represented by light blue graph. Each object can be identified with its two-symbol code from Figure 3. For clarity's sake, similar results are grouped across the horizontal axis. There is a very clear separation between green and red graphs for the vast majority of the cases, while blue graph oscillates between the two. In the majority of the cases, transparent objects have lower rank orders, i.e. are ranked more glossy. This group of observers is followed by the group of observers that have tied all objects. Finally, the red and green parts, still clearly separated, swap places. This part corresponds to the observers, who considered objects with more opacity being more glossy. By the right extreme of the plot, some chaotic arrangements are illustrated that did not follow transparency-opacity cue. On the other hand, it is difficult to see patterns within transparent and opaque groups that makes us think that impact of chromatic information might be negligible.

Clustering

Clustering has been repeated 1000 times by new centroid initialization and the solution with the least sums of point-to-centroid distances was selected out of the 1000 trials. By observation of the graphs above, the number of clusters was set to 3. This lead us to the following clusters (illustrated in Figure 3):

1. Transparent and shiny objects: W1, B1, Y1, Y2.
2. Dark blue translucent object: B3.
3. Objects with more opacity and less shine: W2, W3, Y3, B2.

Rank scores and statistical properties

Rank scores have been illustrated as boxplots (Figure 5) for two major group of the observer population, and as an aggregate for all 107 observers. Objects from the same cluster are coded with the same hue. We can observe a very clear separation between transparent-shiny and more opaquish objects both for "Luminance Group" as well as for "Contrast Group" of the people, with a few outliers included, while object B3 from a separate cluster has some overlap with both clusters. In case of aggregate results, separation remains visible due to significantly higher number of observers in the "Luminance Group" and number of outliers increases due to inclusion the observers making no ranking or doing that with unique strategies. Statistical properties for each cluster of objects for each group of population are illustrated in Figure 6. For the Luminance group, as well as for the entire population, mean and median observer scores for more transparent objects are lower. Standard deviation of B3 for luminance and contrast groups is higher, as it oscillates between the two groups.

Transmittance Measurements

Transmittance spectra for each material has been measured in backlit illumination geometry and relative colorimetric values have been calculated. Due to the complexity of the surface of the female bust objects, measurements have been conducted on cuboid shapes of the identical material shown in Figure 3. The

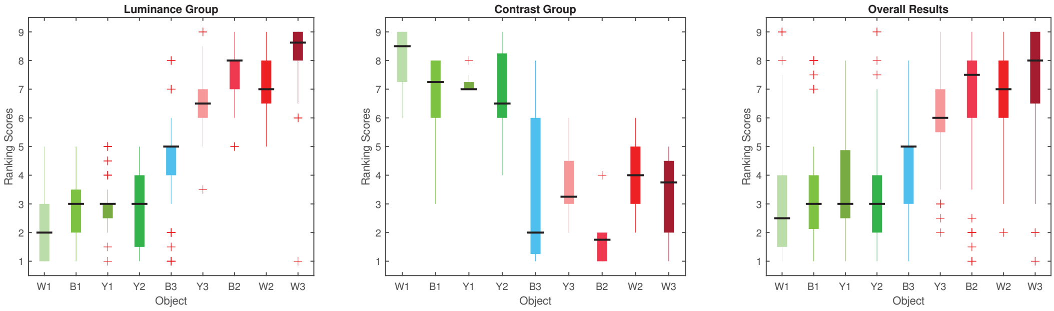

Figure 5. *Boxplots for observer scores showing how each group of the observers ranked the objects. More transparent objects are given with greenish hue, objects closer to opacity are illustrated with reddish hue, while the object B3 is sky blue. 1 means most glossy, while 9 means least glossy. In case of ties, the mean score was taken. Central mark -median; bottom and top edges - 25^{th} and 75^{th} percentiles, respectively; Whiskers extend to the extreme data points excluding outliers; red '+' symbol - outliers. We can observe clear separation for both groups.*

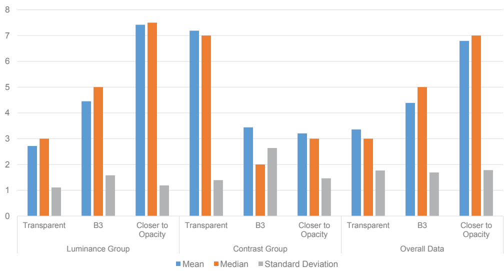

Figure 6. *Statistical properties by cluster for each group of observers.*

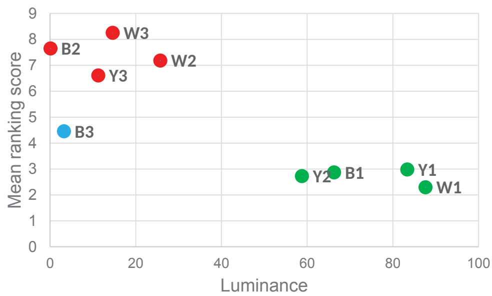

Figure 7. *Mean ranking score as a function of transmittance expressed as luminance value.*

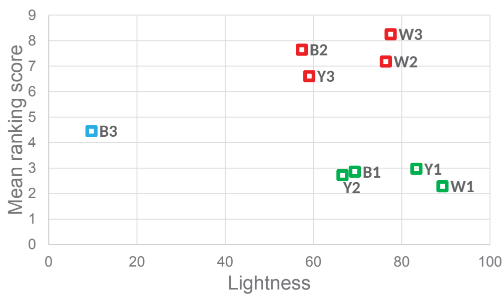

Figure 8. *Mean ranking score as a function of relative radiance expressed as CIELAB L* value.*

white paper seen through the object and caustics should have contributed to shiny appearance. Hence, transmitted luminance information, (Y from measured CIE XYZ), is seemingly correlated with mean rank scores for the "luminance group". This can be seen in Figure 7, where separation among high and low luminance objects is apparent, also supported by k-means clustering. Although luminance for B3 dark blue object is low, it has very high contrast gloss, observers explicitly mentioning that *"highlights are more clearly visible on this object"*. Figure 8 illustrates mean ranking scores as a function of relative radiance expressed as a CIELAB L* value measured in reflectance setup, where cuboid objects were placed on the white background. This enables us to draw parallels with Pellacini's statement that objects with higher lightness in diffuse areas appear less glossy.

Discussion

While the impact of illumination conditions is still to be studied, cross-individual differences might have significantly affected the results. The most obvious illustration of this fact is abundance of "no ranking" scenario for Material Appearance and Electronic Imaging Demonstration Session experiments, where the majority of the subjects had expertise in color, vision, or related fields. Those who considered all objects equally glossy were explicitly asked to justify their decision. All of them defined gloss as surface-only property, limiting themselves to surface judgment. In general, still 97 out of 107 observers were able to rank the objects even though they had explicitly given a possibility not to. After analyzing the data, three groups of people pop out: the ones that judge surface only; people who consider transparent-shiny objects more glossy; and the people who considered objects with more opacity being more glossy. The justification of ranking more opaque ones more glossy were clarity of the highlights and higher contrast gloss, while people opting for transparent ones associate gloss with overall shine and high brightness without scrupulous study of the details. In contrast with the previous experiment, where the two groups were of the equal size, here shininess-based decisions prevail significantly. This could be explained by the absence of distinctness-of-image gloss on the complex surface of the female bust objects, in contrast with a sphere. This confirms Nishida's claim [10] that perceptual surface-reflectance constancy fails when shape is changed, and challenges Pellacini's sphere-based model [12]. Clustering supports our hypothesis that translucency-related attributes as transmittance-measured luminance are common within a cluster. This leads us to hypothesize that gloss and translucency might impact each other. Several observers explicitly complained that it was impossible to isolate translucency/transparency and gloss for above-mentioned objects and thus, to judge them independently. Translucency difference between the two clusters was very large making it challenging to discard its effect. However, the ranking pattern for the B3 object was more irregular. In some cases it was ranked most glossy, justified by high contrast gloss. We can draw a parallel with the first experiment, where sphere "E", made of the similar material, also had substantial confidence interval overlaps with the both groups. This is in agreement with Pellacini's [12] finding that *"for the same specular energy, contrast gloss is smaller for lighter objects"*. Assuming that specular reflections are identical, higher relative radiance in the diffuse part (Figure 8) leads to higher perceived lightness in non-specular areas, and thus, lower contrast gloss. Contrast and clarity of the highlights were mentioned as a cue when they came from surface reflection only, while being less reliable in case of ambiguity whether the light originated from surface reflection or from sub-surface scattering. In total, light transmission properties have impacted perceived gloss in several ways. While contributing to specular gloss by transmission and caustics, contrast gloss is impacted by lightness of the diffuse areas in opaque materials.

Conclusion and Future Work

We have observed that glossiness perception function varies among subjects. While some people try to stick to the literature definition, the vast majority of them ignore surface similarity and sort out objects by gloss using their own criteria. Whether they completely ignore the surface similarity, or they consider it but look for the additional criteria, needs to be explored in the future. There is a very clear indication that perceptual gloss cannot be estimated by surface properties only and light transmission among others might have impact on it. However, the data at hand does not enable us to analyze what is the exact way translucency strengthens glossiness perception and whether the effect comes from overall increase in luminance after light transmission, or due to internal and external caustics that are mistaken for the specular reflections. In future work we should isolate those phenomena and study their impact separately. The hypothesis needs further investigation with more dense sampling across translucency-opacity scale possibly using computer graphics. However, it comes with the compromise that tactile information - a widely-used cue for surface estimation will be lost. Although darker colors enable higher contrast-gloss and contribute to gloss perception, the role of chromatic information is still to be determined by measuring and studying reflection properties. We hypothesize that higher transmittance will lead to stronger gloss perception in the majority of the naïve observers, but low brightness/shine for dark opaque objects could be compensated with increased contrast gloss. Particular interest will be measurement of scattering coefficient and inclusion of the multi-material objects with an eventual goal to model a correlation between material properties and perceived gloss. Development of this work will be reported in the future.

References

[1] Christian Eugène, "Measurement of "total visual appearance": a CIE challenge of soft metrology," in *12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement*, 2008, pp. 61–65.

[2] Michael Pointer, "A framework for the measurement of visual appearance," *CIE Publication*, pp. 175–2006, 2006.

[3] "CIE 17.4:1987 international lighting vocabulary," International Commission on Illumination, 1987.

[4] "ASTM E284-17 standard terminology of appearance," ASTM International, West Conshohocken, PA, 2017.

[5] Bui Tuong Phong, "Illumination for computer generated pictures," *Communications of the ACM*, vol. 18, no. 6, pp. 311–317, 1975.

[6] Yun-Xian Ho, Michael S Landy, and Laurence T Maloney, "Conjoint measurement of gloss and surface texture," *Psychological Science*, vol. 19, no. 2, pp. 196–204, 2008.

[7] Richard S. Hunter, "Methods of determining gloss," *NBS Research paper RP*, vol. 958, 1937.

[8] Isamu Motoyoshi, Shin'ya Nishida, Lavanya Sharan, and Edward H Adelson, "Image statistics and the perception of surface qualities," *Nature*, vol. 447, no. 7141, pp. 206–209, 2007.

[9] Fred E Nicodemus, "Directional reflectance and emissivity of an opaque surface," *Applied optics*, vol. 4, no. 7, pp. 767–775, 1965.

[10] Shin'ya Nishida and Mikio Shin'ya, "Use of image-based information in judgments of surface-reflectance properties," *JOSA A*, vol. 15, no. 12, pp. 2951–2965, 1998.

[11] Nahian S Chowdhury, Phillip J Marlow, and Juno Kim, "Translucency and the perception of shape," *Journal of vision*, vol. 17, no. 3, pp. 1–14, 2017.

[12] Fabio Pellacini, James A Ferwerda, and Donald P Greenberg, "Toward a psychophysically-based light reflection model for image synthesis," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 55–64.

[13] Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen, "Behavioral investigation of visual appearance assessment," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2018, pp. 294–299.

[14] Roland W Fleming and Heinrich H Bülthoff, "Low-level image cues in the perception of translucent materials," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 3, pp. 346–382, 2005.

[15] David K Lynch, William Charles Livingston, and William Livingston, *Color and light in nature*, Cambridge University Press, 2001.

[16] Jean-Baptiste Thomas, Aurore Deniel, and Jon Y Hardeberg, "The plastique collection: A set of resin objects for material appearance research," *XIV Conferenza del Colore, Florence, Italy*, p. 12 pages, 2018.

[17] David Arthur and Sergei Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.

Article D

Davit Gigilashvili, Weiqi Shi, Zeyu Wang, Marius Pedersen, Jon Yngve Hardeberg, and Holly Rushmeier (2021). “The Role of Subsurface Scattering in Glossiness Perception.” In: *ACM Transaction on Applied Perception* 18.3, 10:1–10:26

The Role of Subsurface Scattering in Glossiness Perception

DAVIT GIGILASHVILI, Norwegian University of Science and Technology, Norway
WEIQI SHI and ZEYU WANG, Yale University, USA
MARIUS PEDERSEN and JON YNGVE HARDEBERG, Norwegian University of Science and Technology, Norway
HOLLY RUSHMEIER, Yale University, USA

This study investigates the potential impact of subsurface light transport on gloss perception for the purposes of broadening our understanding of visual appearance in computer graphics applications. Gloss is an important attribute for characterizing material appearance. We hypothesize that subsurface scattering of light impacts the glossiness perception. However, gloss has been traditionally studied as a surface-related quality and the findings in the state-of-the-art are usually based on fully opaque materials, although the visual cues of glossiness can be impacted by light transmission as well. To address this gap and to test our hypothesis, we conducted psychophysical experiments and found that subjects are able to tell the difference in terms of gloss between stimuli that differ in subsurface light transport but have identical surface qualities and object shape. This gives us a clear indication that subsurface light transport contributes to a glossy appearance. Furthermore, we conducted additional experiments and found that the contribution of subsurface scattering to gloss varies across different shapes and levels of surface roughness. We argue that future research on gloss should include transparent and translucent media and to extend the perceptual models currently limited to surface scattering to more general ones inclusive of subsurface light transport.

CCS Concepts: • **Computing methodologies** → *Computer graphics*; *Graphics systems and interfaces*; *Perception*; *Rendering*;

Additional Key Words and Phrases: Material appearance, gloss perception, translucency perception, subsurface light transport, MTurk

ACM Reference format:
Davit Gigilashvili, Weiqi Shi, Zeyu Wang, Marius Pedersen, Jon Yngve Hardeberg, and Holly Rushmeier. 2021. The Role of Subsurface Scattering in Glossiness Perception. *ACM Trans. Appl. Percept.* 18, 3, Article 10 (May 2021), 26 pages.
https://doi.org/10.1145/3458438

1 INTRODUCTION

Humans are adept at identification of materials [2, 60] and can easily characterize their appearance [2, 19]. A typical human with normal vision does not need much effort or prior training to tell the difference between

This work was supported in part by NSF Grant No. IIS-2007283. The work was also supported by MUVApp (Grant No. 250293) and MANER (Grant No. 288187) projects of the Research Council of Norway

Authors' addresses: D. Gigilashvili, Norwegian University of Science and Technology, 22 Teknologiveien, Ametyst-bygget, Gjøvik, 2815, Norway; email: davit.gigilashvili@ntnu.no; W. Shi, Z. Wang, and H. Rushmeier, Graphics Lab, Department of Computer Science, Yale University, P.O. Box 208285, New Haven, CT 06520, USA; emails: {weiqi.shi, zeyu.wang, holly.rushmeier}@yale.edu; M. Pedersen and J. Y. Hardeberg, Norwegian University of Science and Technology, 22 Teknologiveien, Ametyst-bygget, Gjøvik, 2815, Norway; emails: {marius.pedersen, jon.hardeberg}@ntnu.no.

1544-3558/2021/05-ART10
https://doi.org/10.1145/3458438

Fig. 1. Examples of materials and shapes used in the study of the impact of subsurface scattering on gloss.

shiny and matte objects, or whether a material transmits light. Assessment of material appearance has a vital importance in our daily lives—just by visual inspection, we know whether food is edible or spoiled, whether the road is slippery or not. Tactile expectations derived from the visual appearance can guide our haptic interaction with the surrounding objects—for instance, we touch glossy, transparent crystal-looking objects with more care than we do for jelly-looking, matte objects; expecting the latter to be soft and elastic, while the former is deduced to be fragile. How the **human visual system (HVS)** calculates these appearance properties from the physical stimulus is far from being fully understood. Comprehending the physical processes and inverting optics [51], as well as the calculation of image statistics by our brain [44] have been named among the potential explanations, both criticized on several grounds [13, 14, 32].

Gloss is among the most important visual attributes of a material [11, 20]. It is usually associated with shininess [21] due to the specular reflection and is formally defined as an *"angular selectivity of reflectance, involving surface reflected light, responsible for the degree to which reflected highlights or images of objects may be seen as superimposed on a surface"* in the ASTM Standard Terminology of Appearance [1]. The six distinct dimensions of gloss—specular gloss, contrast gloss, distinctness-of-reflected-image gloss, absence-of-bloom gloss, absence-of-surface-texture gloss, and sheen—have been proposed by Hunter [29] back in 1937. Since then, gloss has been accepted as a surface-related quality, and perception of gloss has been studied in the context of surface scattering models [49, 62, 71]. Various image cues have been proposed to be used by the HVS for gloss perception (for instance, the total area covered by specular reflections, contrast between specular reflections and surrounding areas, the sharpness of the edges of the specular regions [38, 39]). Although it has been demonstrated that shape and illumination co-vary with the image cues proposedly used for gloss estimation [38], these cues can also be affected by the subsurface light transport (See Figure 1).

When a light ray reaches a boundary between two media with mismatching indices of refraction, part of it is reflected specularly (i.e., the light re-emerges back toward the incidence hemisphere but on the opposite side of the surface normal) or refracted (i.e., changes the direction and continues propagation inside the new medium). The light can either get absorbed or scattered by scattering particles when propagating through a medium. An average distance a photon travels before it gets either absorbed or scattered depends on the extinction coefficient of the material. Many rendering techniques use the concept of diffuse reflectance (i.e., scattering the incident light from a surface into many different angles) for modeling opaque media. However, the optical phenomenon known as "diffuse reflectance" actually involves subsurface scattering of light—a photon penetrates the superficial layer of the material, where it quickly gets either absorbed by the pigments or scattered backwards toward the incidence hemisphere, defining the color of the material and generating an opaque appearance. However, if the extinction coefficient is low or the object is thin enough, then a photon might re-emerge from a different side of the object—generating transparent or translucent appearance. The process is illustrated in Figure 2. While primarily specular reflection has been thought to be responsible for glossy appearance (see Reference [36] for a review), diffuse reflection has been shown also to be playing a role [49]—assuming negligible or non-existent

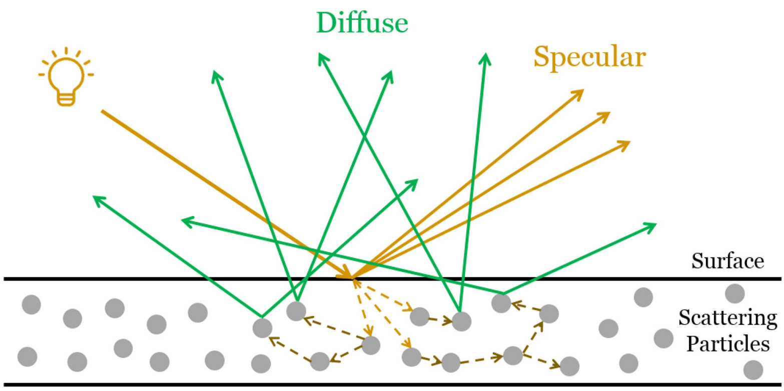

Fig. 2. Light gets either reflected specularly or refracted at the boundary of the two media with mismatching indices of refraction. What is known as diffuse reflection is actually light scattered backwards from the superficial layers of the subsurface due to high extinction coefficient. However, if the extinction coefficient is low, then light can re-emerge far from the point of incidence, considerably affecting the visual appearance.

subsurface light transport most of the time. In other words, the studies addressing gloss perception have been traditionally limited to surface reflection and fully opaque media (e.g., References [15, 38, 39, 47, 49, 53, 54, 62, 64, 66, 69–71]), while a lot of materials we interact with on a daily basis, are both glossy and light-transmissive—water, glass, marble, or human skin can be named among many. The knowledge about the peculiarities of gloss perception on transparent and translucent materials is very limited.

In this article, we hypothesize that subsurface scattering impacts glossiness perception. The hypothesis is reasoned from the following notions:

(1) Due to the limited dynamic range and poor capability of the HVS to comprehend and invert the complex optical path of the light [14], human observers might have difficulty unmixing transmitted and surface-reflected light. Hence, caustics, direct transmission or volume scattering can be mistaken for specular reflections. Imagine a transparent crystal vase with a complex shape. It shines, has sparkles and highly luminant areas. Is it possible to tell whether the highlights are due to the reflection, direct transmission, or subsurface scattering of light? Do not all these shiny parts evoke a feel of glossiness regardless of their origin?
(2) It has been demonstrated that darker objects look glossier than lighter ones [49, 62] due to higher contrast between specular and diffusely-reflecting areas (Hunter's contrast gloss [29]). As volume scattering and absorption can impact the contrast between specular and non-specular areas, they might also impact apparent gloss.
(3) Observation of the mirror-like reflection image on the surface has been identified to be a strong glossiness cue [21] (Hunter's distinctness-of-reflected-image gloss). While it has been thought to be correlated with surface roughness only [49], the distinctness of the reflected image can be dependent on light transmission properties as well. The same applies to the sharpness of the highlights, which is another glossiness cue [38, 39].
(4) Subsurface light transport can influence the size of the highlights on complex-shaped objects. It has been demonstrated multiple times that the size of the highlights is correlated with perceived glossiness [4, 31, 38, 39].

(5) For transparent objects, as the transmitted and reflected light integrate, overall luminance reaching the human retina is higher and the object shines more [19, 21]. Overall shine as an inherent characteristic for gloss, might evoke a perception of glossiness.
(6) Finally, caustics and light transmission might facilitate material identification. If a stimulus is associated with a familiar, usually glossy material, then the expectations about this material can impact the perception of glossiness [58].

To test this hypothesis, we have conducted a series of pair-comparison experiments. In the first (pilot) experiment, we studied how surface and subsurface scattering affect gloss perception on the example of spherical objects. The results of the pilot experiment have indicated that the impact of subsurface scattering on gloss varies among different levels of microfacet-scale surface roughness. This can be explained by the fact that glossiness cues vary dramatically between mirror-like and Lambertian-like surfaces [29, 53, 70]. We have interviewed several participants (members of our lab) in the pilot study. They noted that if the shape of the stimulus were different, it could have affected their answers. This correlation was deemed reasonable by us, as the macro-scale shape of the object can impact translucency and subsurface light transport [14, 19, 22]. To investigate further, the second experiment was arranged, studying objects with five different shapes each with five different levels of surface roughness. We analyzed the depth and curvature of object shapes and identified interesting trends in how the contribution of subsurface scattering to gloss varies among object shapes. Our contributions in this article are the following:

- We experimentally test the hypothesis that subsurface scattering impacts gloss perception for materials with identical shape and identical surface scattering.
- We identify whether the contribution of subsurface scattering to the glossiness perception varies among different macro-scale and micro-scale (microfacet-level) shapes, and characterize this impact qualitatively.
- We discuss the need for inclusion of subsurface scattering in future studies, opening a new avenue in gloss perception research.

The article is organized as follows: in the next section, we summarize the related work. In Sections 3 and 4, we present the two experiments and their results, respectively, followed by the Discussion section. Finally, we summarize the conclusions and overview the open points for future work.

2 RELATED WORK

The perception of gloss and translucency has attracted scholarly interest in vision, psychology, and computer graphics alike. While substantial progress has been achieved on both topics, the two attributes have usually been studied separately from each other.

2.1 Gloss Perception

One of the most widely discussed hypotheses about gloss perception is that the HVS calculates skewness of luminance histogram or a similar measure of asymmetry when assessing gloss [18, 35, 44]. Interestingly, many glossy objects have positively skewed histograms. However, it has been shown by Anderson and Kim [3] that non-glossy images can also produce similar histograms and image statistics do not fully explain the complex neurophysiological processes of gloss perception (e.g., References [18, 32, 37]). Other widely studied image metrics that are proposedly related to gloss are contrast [38, 39, 49, 62], sharpness [38, 39, 49], and coverage area [4, 31, 38, 39] of the highlights. The glossiness of a given material has been demonstrated not to be constant and can vary to a great extent, e.g., across different shapes [39, 48, 66]. In some particular cases, even Lambertian surfaces are capable of evoking gloss perception [52, 53, 70]. Gloss has been shown also to be impacted by illumination geometry [15, 48], motion [9, 56, 69], and color [46, 69]. Pellacini et al. [49] have used **multidimensional scaling (MDS)** and identified two perceptual dimensions of gloss that are similar to contrast and distinctness-of-image. They conclude that *"darker objects look glossier than lighter ones."* Wills et al. [71] tried to embed **bidirectional reflectance distribution functions (BRDFs)** into the perceptual space. These perceptual dimensions have been modeled with physical material properties in Ward's reflectance model [68], ignoring subsurface light transport. Toscani et al. [64] have recently proposed that surface reflection has at least three perceptual dimensions: lightness, gloss, and metallicity. However, the authors did not address how these dimensions behave on highly transparent and translucent media.

2.2 Translucency Perception

Translucent appearance is a result of subsurface scattering for the materials where the light can penetrate into the volume. Although Chadwick et al. [5] have reported yet imperfect still reasonable perceptual unmixing of absorption and scattering by humans in "milky tea" images, Fleming and Bülthoff [14] argued that the HVS has poor ability to reconstruct complex processes of light and matter interaction and instead it relies on simple image cues to perceive translucency. These cues co-vary with various properties of an object. Image cues as well as the amount of light exiting the volume depend on the shape complexity and thickness of a given object. For instance, it has been shown that sharp geometric details of the object impact apparent translucency [74] and the other way round, translucency affects perception of geometric edge sharpness [6]. Sawayama et al. [57] have reported that *"sensitivity to translucent discrimination was high when the object has rugged surfaces."* Furthermore, Gigilashvili et al. [19] have observed that objects with thin parts look more translucent and that the HVS is more sensitive to translucency differences when an object has thin parts [22]. Motoyoshi [43] observed that luminance statistics of the non-specular regions are essential for apparent translucency and that decreasing local contrast in these regions of an opaque material renders translucent appearance. Nagai et al. [45] discussed luminance statistics of potential "hot spot" image regions that are especially informative about translucency. Later, particularly edges have been proposed to contain a vital portion of the information for translucency assessment [23]. Similar to gloss, the translucency of a material is not constant either. It has been shown to be dependent on the illumination geometry [17, 73] and shape [14, 19]. Gkioulekas et al. [24] have examined translucent appearance in the context of computer graphics and found that the phase function of volume scattering affects translucent appearance.

2.3 Impact of Translucency on Gloss

Gigilashvili et al. [19] reported no significant differences in gloss perception of five physical spherical objects with identical surface roughness but different translucency and color. The authors revisited the study in Reference [21] and after analyzing the observer interviews, they discovered that different people rely on different cues. The authors have identified three groups of people with different approaches to solve the gloss-based ranking task. While objects with identical surface were automatically considered equally glossy by some subjects, two other groups used different cues for ranking, either overall shininess of the object—mostly present in transparent and translucent spheres, or distinctness-of-image and contrast—that were higher for more opaque ones. When the experiment was conducted using complex-shaped objects instead of spherical ones [21], the majority of the observers considered translucent objects glossier than their opaque counterparts. The authors hypothesize that this happens due to the complex shape, which generated more caustics and back-reflections for translucent and transparent materials, while lacking distinctness-of-image for the opaque ones. They refer to the reasoning by Fleming and Bülthoff [14] about poor optics inversion ability of the HVS and propose that subjects might have mistaken caustics for specular reflections. If that is possible for physical objects during direct interaction, then confusion can be even larger in computer graphics, where haptic interaction is impossible and tactile information is absent. It is worth mentioning that these works have been primarily of a qualitative nature. To the best of our knowledge, this is the first work quantitatively evaluating the impact of translucency on gloss.

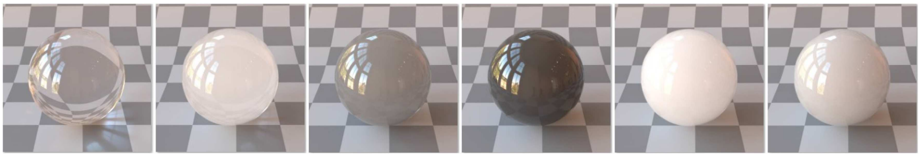

Fig. 3. Spheres with the same surface roughness (*alpha* = 0) but different subsurface scattering properties. [σ_t, albedo] parameters of these spheres are equal to [0.10,0.50]; [1.00,0.90]; [2.00,0.60]; [3.00,0.30]; [3.00,0.95]; [4.00,0.90], from left to right, respectively.

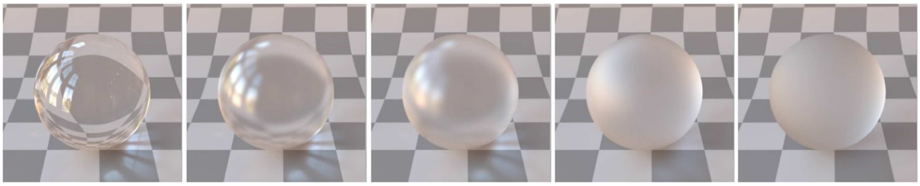

Fig. 4. Spheres with the same subsurface scattering properties (σ_t = 0.10 and albedo = 0.50) but different surface roughness, with *alpha* equal to 0.00, 0.05, 0.10, 0.25, 0.50, from left to right, respectively.

3 EXPERIMENT 1: PILOT STUDY

3.1 Methodology

3.1.1 Objectives. The objectives of this experiment are twofold: first, we test a hypothesis that subsurface scattering impacts gloss perception when surface scattering and object shape are identical; second, we observe how surface and subsurface scattering impact perceived gloss together.

3.1.2 Stimuli. We began our study by considering different scenes to use for our experiments. For illumination, we followed the previous work [24] using the side-lighting by rotating Bernhard Vogl's museum environment map provided by Mitsuba [30] to a proper angle. We created synthetic images of spherical objects using a physically based rendering in Mitsuba. Spheres have been widely used in the past for studying gloss perception (e.g., References [15, 19, 49, 62, 72]). For surface reflectance, we used an isotropic rough dielectric microfacet model with the Beckmann distribution [30]. The model is defined by roughness *alpha* (the root mean square slope of microfacets) and an index of refraction *IOR*. As we restrict our attention to subsurface scattering effects, we use a fixed *IOR* of 1.5, which is typical for translucent media such as glass, wax and polymeric materials [42, 59]. All objects were placed on a Lambertian checkerboard. It is important to highlight that the rendering technique we used [30, 67] has accounted for Fresnel effects. Fresnel effects imply that the amount of observed reflectance varies with the observation angle, which have been shown to be important for gloss perception [12] and for appearance of dielectric materials, in general [26]. The experiment was conducted in two rounds: Since our primary goal was to explore whether subsurface light transport influences gloss perception, in the first round, we compared objects with an identical surface roughness parameter (also referred to as *alpha*) and different parameters of subsurface scattering. To explore how the impact of volume scattering on gloss perception varies among the different levels of surface roughness, we have repeated the experiment for the different *alphas* separately. In the second round, we compared the stimuli with different *alphas*. We select roughness from the set $\{0, 0.05, 0.1, 0.25, 0.5\}$ to cover a wide range of surface reflectance behavior. Some of the stimuli are illustrated in Figures 3 and 4.

We used a homogeneous isotropic subsurface scattering model to simulate the translucent appearances. For this pilot, we assume an isotropic phase function and wavelength-independent scattering and absorption for subsurface light transport. The subsurface scattering parameters are the extinction coefficient σ_t and *albedo*. For the extinction coefficient, we found through experimentation that increasing σ_t over 10 does not yield significant differences in appearance for our shape, because the material becomes opaque. Therefore, we selected $\sigma_t \in \{0, 0.1, 0.5, 1, 2, 3, 4, 5, 10\}$. For albedo, we selected $albedo \in \{0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$. Such a dense sampling of parameters covers a wide range of appearance but would require an enormous number of comparisons to be evaluated. Although the parameters have been selected based on visual inspection in a trial-and-error manner, many pairs of parameter values still lead to indistinguishable appearances, which are redundant for the user study. To select a smaller set of parameter combinations for stimuli with the same surface reflectance, we used the K-means clustering algorithm to find six distinctive clusters based on different subsurface scattering parameters. We used the averaged Euclidean distance of pixels from the rendered images as a metric to perform K-means clustering. We have explored other clustering algorithms, such as affinity propagation [16], but K-means has provided the best clustering results according to the silhouette coefficient. We used the cluster center as our stimulus for the user study. Since the K-means has been conducted separately on different groups of surface roughness, the cluster centers were not identical for all surface roughness levels. The variation in the cluster centers was small, however, and so we selected identical subsurface scattering parameters for all levels of surface roughness. Thirty different stimuli were used in total (five different levels of surface roughness $\{0.00, 0.05, 0.10, 0.25, 0.50\}$ and six different combinations of σ_t and albedo, where $[\sigma_t$,albedo$]$ $\in \{[0.10, 0.50]; [1.00, 0.90]; [2.00, 0.60]; [3.00, 0.30]; [3.00, 0.95]; [4.00, 0.90]\}$). We used the volumetric path tracing integrator of Mitsuba to render the stimuli with 512×512 pixel resolution and 16,384 samples per pixel. The tonemapped (clipped) low-dynamic-range images have been used to ensure the compatibility with the user displays. All images can be found in supplementary materials (Figure 23).

3.1.3 Experimental Design. We considered two different designs of two alternative forced-choice task: either displaying two stimuli and asking the subjects (also referred to as users) to select a glossier stimulus, or displaying three stimuli and asking to select two stimuli closer to each other in terms of gloss (a setup similar to Wills et al. [71]). We ran a preliminary study with both designs. Eight members of our lab completed the tasks and participated in informal post-experiment interviews. Seven of eight subjects mentioned that selecting a glossier stimulus between the two was an easier task than comparing the three by similarity. They also admitted that oftentimes they had found it difficult to isolate gloss from total appearance and were tempted to judge similarity by overall appearance or lightness. Therefore, we selected the former option for the task design.

First, we conducted separate paired-comparison experiments for each level of *alpha*. The users were shown two spherical objects with the same surface roughness and different subsurface scattering parameters. They were asked to select the one with a glossier appearance. The user interface is illustrated in Figure 5. The proper command of English among subjects was ensured with the Amazon Mechanical Turk average approval rate filter (see Section 3.1.6). Only the users with a positive track record of similar tasks were allowed to participate. The following instruction was given to them: ***"Click on the image that contains the glossier object. You can click after taking two seconds to look at the images."*** No further definition or guidance was provided. The reason for abstaining from a definition is the following: any particular definition for *gloss* could have biased subjects' decisions. For instance, as mentioned above, the ASTM Standard Terminology of Appearance [1] defines *gloss* as *"angular selectivity of reflectance, involving surface reflected light, responsible for the degree to which reflected highlights or images of objects may be seen as superimposed on a surface."* Reference to the definition that highlights gloss as a *reflectance* property might have had an implication for some subjects that *subsurface scattering effects* should be ignored. This contradicts the objective of this experiment. The research objective of this study was the identification of the factors impacting the overall sensation of gloss, not the psychometric measurement of an internal function for a given visual cue. It is worth mentioning that seminal works on gloss perception

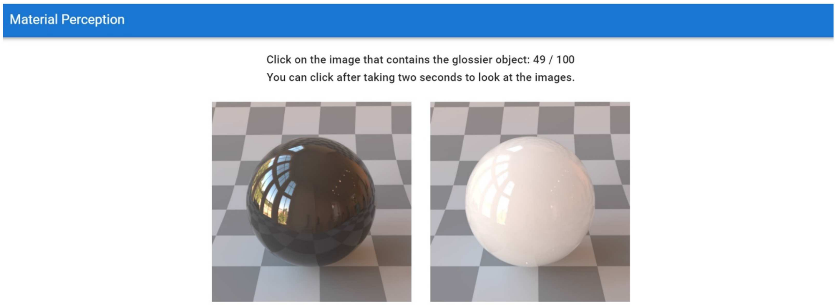

Fig. 5. The user interface identical to this one has been used to conduct the experiments on the Amazon Mechanical Turk.

(e.g., References [49, 71]) usually have no mention that the term was defined for the subjects, unless the objective is a psychophysical measurement of a particular, explicit cue (such as *specular contrast* and *specular sharpness* in Reference [38]).

There was no time limit for each trial. Each user was asked to complete 100 trials in random order, of which 75 were unique trials (6 different materials yield 15 trials for each roughness level, totalling to 15×5) and 25 were repeated trials with images in reverse order. We used the repeated trials to assess *intra-rater reliability* by counting the number of pairs (of 25) the subject selected the same stimulus on both trials. We designed our system with a delay mechanism: the users could only select the candidate image two seconds after the pair was displayed. This mechanism makes sure that users take time to examine the images. The users, on average, took about 5 min to assess 100 comparisons.

To understand how surface reflectance and volume scattering influence gloss perception together, we conducted a second round of paired-comparison experiments, where the two candidate images had different surface roughness. Instead of dividing the 30 stimuli into five groups and conducting experiments separately for each roughness level, this time the users had to compare the stimuli from different roughness groups, yielding 360 unique pairs in total (each of the 30 stimuli was compared with other 24 stimuli of different *alpha*; from the first round of the experiment, we already had the data for the objects with the same *alpha*). Twenty-five percent of the pairs were shown twice for controlling *intra-rater reliability*.

3.1.4 Analysis: Hypothesis Testing. We formulate a null hypothesis that subsurface light transport has no impact on gloss perception. To test the null hypothesis, we conducted Binomial exact statistical significance tests, as our outcome is binary. Under the null hypothesis, the expected probability of each stimulus being considered glossier is 0.50. We assess observed frequencies and calculate the probability of observing those frequency values when the null hypothesis is true. As it is not important at this stage which of the two stimuli is glossier (we just want to show that subsurface scattering makes them look different in terms of gloss), we conduct a two-tailed test—i.e., it does not matter whether the observed frequency is larger or smaller than the expected one. If the probability of observing given frequencies is less than 0.05 under the null hypothesis, then the difference is deemed significant and the null hypothesis is rejected. To avoid falsely rejecting the null hypothesis due to multiple testing (type I error), we applied Holm-Bonferroni [28] correction to the data.

3.1.5 Analysis: Z-scores. A further method to analyze the pair-comparison data is Z-scores (Standard scores) [10, 65]. It is based on Thurstone's law of comparative judgment [63]—assuming that each sample has a quality that is being assessed by a subject and these qualities are Gaussian random variables. Each time a subject compares the two samples, realizations from both random variables are drawn and compared, selecting the one with higher quality. The probability of selecting a given option is found using the standard normal **cumulative distribution function (CDF)**. The inverse CDF of the standard normal is a Z-score showing how many standard deviations away is a given option from the mean. Usually, Thurstone's simplified Case V model is used assuming that all samples are independent and have equal variance [65]. For all samples, we present the mean Z-scores and their 95% confidence intervals as error bars (calculated using MATLAB Colour Engineering Toolbox [25]). The mean Z-score shows how far a given stimulus is from the mean of the set of stimuli being assessed. If the 95% confidence intervals of the Z-scores do not overlap, then we can tell with 95% confidence that the qualities of the two stimuli are significantly different.

3.1.6 Subjects. The sample size is found by desired statistical power, significance level and effect size for the Binomial null hypothesis testing. The desired statistical power was set to 0.8 (the probability of rejecting the null hypothesis when the alternative hypothesis is true) and the significance level was set to 0.05 (the probability of falsely rejecting the null hypothesis when it is actually true). As per the null hypothesis two stimuli are equally glossy, the expected probability is 0.5. To decide on alternative proportion, two different effect size metrics [55] were used: Cohen's g—usually used for the cases where the expected proportion is 0.5 and simply found as a difference between the proportions, and Cohen's h—that is found as

$$h = 2\left(\arcsin\sqrt{p_1} - \arcsin\sqrt{p_2}\right), \tag{1}$$

where h is Cohen's h (sometimes reported as an absolute value) and p_1 and p_2 are the two proportions. Under an alternative proportion of 0.75, $g = 0.25$ and $h = 0.52$, being interpreted by Cohen [7] (cited in Reference [55]) as large and medium effect sizes, respectively. Thus, we set an alternative proportion to 0.75. Considering these values, the needed sample size was approximated as 29.

We conducted our experiments on Amazon **Mechanical Turk (MTurk)** and collected responses from 50 users per pair. In total, around 250 subjects participated in both rounds. The users were compensated for participation. The compensation varied from experiment to experiment and was within the range of 2–3 USD per 100 comparisons. To ensure the reliability of the users, two filters were applied: first, only the MTurk users with an average approval rate above 50% were allowed to participate; and second, the participants were ranked by their performance in the *intra-rated reliability* test, i.e., by the consistency of their responses on the validation set (how many times they selected the same stimulus in the pairs shown twice). Eventually, 30 most consistent subjects were considered per stimuli pair, around 150 subjects in total. The reason for users' inconsistency can be not only their inattentiveness but also the stimuli that are visually indistinguishable. The number of such pairs is unknown before the experiment and hence, it is not possible to set a threshold for "acceptable consistency" in advance. For this reason, we had to rely on ranking instead of absolute values of consistency. Interestingly, the top 30 users turned out to be consistent in at least 70% of the cases. Finally, it is worth mentioning that the results with concurrent clicks from the same IP address were discarded, because it was impossible to calculate their intra-rater reliability and to identify how many unique subjects were responding.

3.2 Results

The results for the fixed roughness experiment are shown in Figures 6 and 7. Figure 6 shows that the difference is significant and the null hypothesis can be rejected for a substantial number of image pairs. This is especially true for smooth objects. The number of pairs that are significantly different gradually decreases, but for *alpha* = 0.50 it starts increasing again. While the two-tailed Binomial tests can just tell whether the difference is significant, the Z-score plot in Figure 7 illustrates which stimuli have been deemed glossier. If the null hypothesis were true, then all stimuli were expected to end up with similar Z-scores. However, the observed trend is consistent with the Binomial tests—the difference among some stimuli is significant and it is large for smooth objects while the difference gradually diminishes but starts increasing again for the highest *alpha*. The materials either with low σ_t

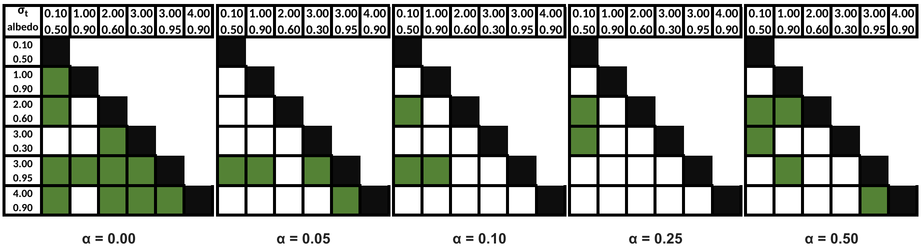

Fig. 6. Significance tables for each roughness level. Each lower triangular matrix shows which of the stimuli pairs are significantly different. Green cells—statistically significant difference; white cells—no statistically significant difference. The number of significantly different pairs is larger for smooth objects (alpha equal to zero).

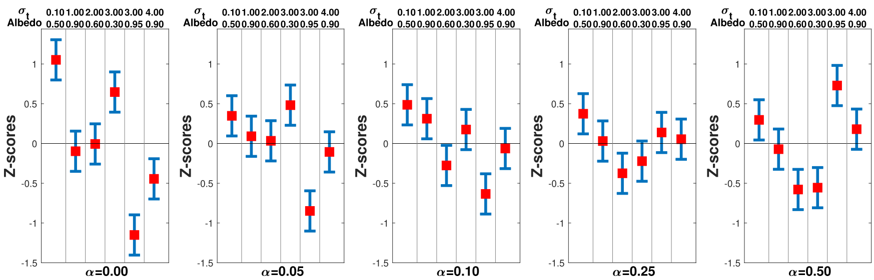

Fig. 7. Z-scores for fixed roughness experiments. A red cube corresponds to the mean Z-score for a given object, while the error bar corresponds to a 95% confidence interval. The variation among Z-scores decreases with the increase of roughness, i.e., Z-scores of five different materials are more equal when *alpha* is high. However, this trend is not monotonic and it does not hold for *alpha* = 0.50.

or albedo were considered glossiest, while the ones with high albedo turned out less glossy. The results including all comparisons among the 30 stimuli are shown in Figure 8 and 9. The significance table shows that the vast majority of the differences between different roughness levels are significant, while no significant differences are usually observed among the objects with the same roughness. However, there are a few exceptional instances—the materials with high albedo (0.95) are not significantly glossier than some other objects with a rougher surface (Figure 8). Examples of the objects with different surface roughness but equivalent (not significantly different) apparent gloss are illustrated in Figure 10.

A clear trend is visible in Z-score plots (Figure 9)—with the increase of surface roughness, the perception of glossiness is decreasing monotonically, being consistent with the prior works [27, 54]. It is worth noting that although it is the identical data, the Z-score differences among the stimuli within each roughness group decreases when considered together with all other stimuli (compare Figures 7 and 9). This can be explained by the fact that a Z-score for a given stimulus is relative and depends on the judgment against all other stimuli in the set. Within a larger pool of stimuli and various *alphas*, the subjects tend to focus more on the surface reflectance instead of the subtle effects of subsurface light transport. All these observations demonstrate that even though the subsurface light transport has an impact, the surface reflectance still plays a major role in the perception of glossiness.

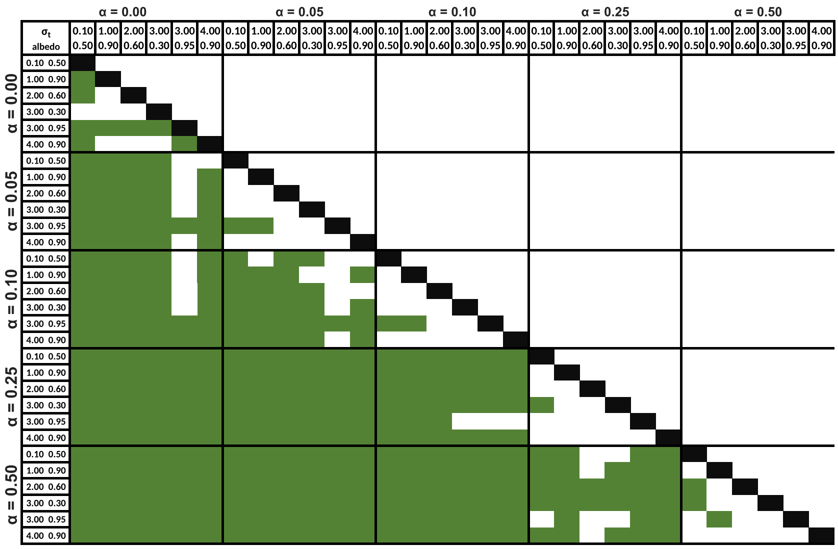

Fig. 8. The significance table for all 30 stimuli. The lower triangular matrix marks the stimulus pairs with statistically significant difference. Green cells—statistically significant difference; white cells in the lower triangle—no statistically significant difference.

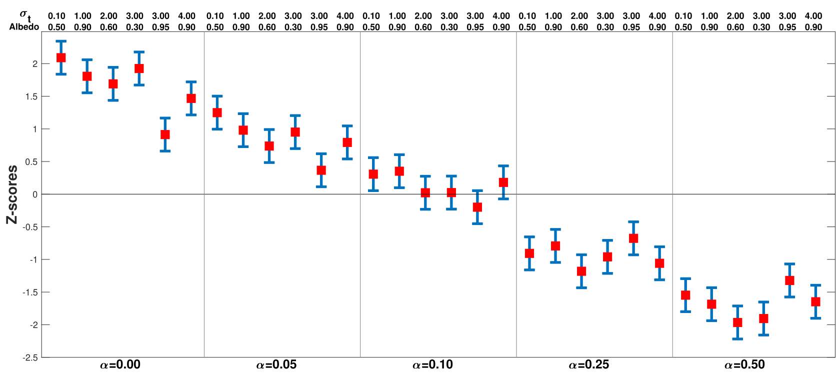

Fig. 9. Z-scores for the comparisons of all 30 stimuli. As we observe, surface scattering is dominant over subsurface scattering and smoother objects usually look glossier. However, in some cases, high albedo makes objects no glossier than some of the rougher ones.

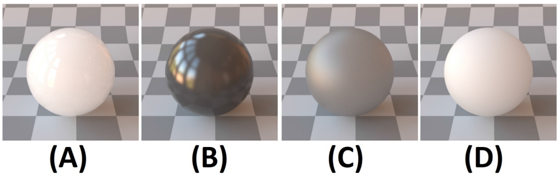

Fig. 10. The difference between apparent gloss of the objects A and B, as well as between C and D, has been shown not to be significant. We can consider them having equivalent apparent gloss. Even though A has smoother surface (*alpha* = 0.00) than B (*alpha* = 0.05), low albedo of the latter compensates for the difference in surface scattering. Similarly, C has relatively smoother surface (*alpha* = 0.25) than D (*alpha* = 0.50), but in this case, it is the high albedo of the latter that is responsible for the equivalent apparent gloss despite substantial difference in surface scattering.

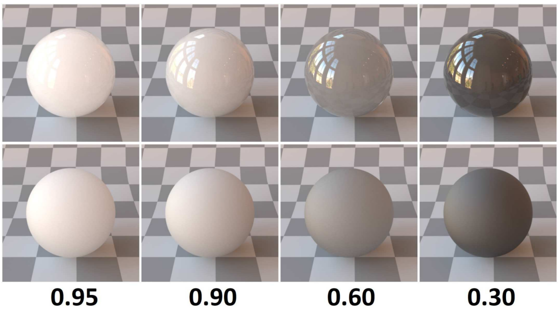

Fig. 11. The number below the images corresponds to their albedo. Although all of the objects have identical surface roughness in each row (*alpha* = 0.00 in the top row; *alpha* = 0.50 in the bottom one), the users have distinguished them in terms of glossiness. According to user responses, the top row can be ranked in terms of apparent gloss, from left to right, the rightmost one being the glossiest. The bottom row can be ranked in the opposite way—the leftmost one being glossiest (but the difference between the two rightmost ones is not significant).

3.3 Discussion

While surface roughness has a strong negative impact on gloss (being consistent with References [49, 62]), for numerous pairs of the stimuli with identical surface roughness, we have rejected the null hypothesis and observed a significant gloss difference induced by subsurface scattering of light. The way subsurface scattering impacts gloss perception differs among different levels of surface roughness and changes non-monotonically.

When *alpha* is low and σ_t is high, gloss increases as the albedo decreases. This phenomenon is illustrated in Figure 11 (also supported by the plot in Figure 16). With a high extinction coefficient, the subsurface light penetration is reduced, yielding appearance closer to diffuse reflectance. This scenario can be paralleled with a diffuse component in Ward's surface reflectance model: decreasing the diffuse reflectance leads to glossier appearance—proposedly due to increased contrast, making our observations consistent with that of Pellacini et al. [49].

When the stimuli are rough (high *alpha*) and do not have strong glossiness cues (such as specular highlights), caustics or the overall shinier look created by high volume scattering could potentially be considered a glossiness cue. This might explain why people can still tell the difference between the stimuli with high *alpha* in our experiments, and why Lambertian surfaces are capable of evoking perception of glossiness [53, 54]. In general, the stimuli with low σ_t and smooth surface ($alpha = 0$) were selected as the glossiest (see the leftmost image in Figure 3). The caustics and back-reflections from the background might be reasons for this (a similar trend has been observed for some subjects in Gigilashvili et al. [19, 22]). Furthermore, the glass-like appearance can also evoke a stronger perception of glossiness due to material identification and the association with the properties of a familiar material, as proposed by Schmid et al. [58]. Several important points have been learned from this experiment that guided the subsequent experiments:

- Since the way subsurface light transport contributes to gloss depends on the surface scattering, we decided to study this contribution for each surface roughness level individually.
- If the change in surface scattering induced by subtle changes in microfacet slopes has a dramatic impact on the behavior of subsurface scattering, then we believe the same will be true for macro-scale changes of the object shape. Therefore, we decided to study the contribution of subsurface scattering for multiple different shapes individually and to compare the trends among them.

4 EXPERIMENT 2: IMPACT OF SHAPE

4.1 Methodology

4.1.1 Objectives. Experiment 1 provides evidence that subsurface scattering can impact gloss perception for spherical objects, and this impact depends on the amount of surface scattering. The objective of Experiment 2 is to quantitatively study whether subsurface scattering impacts glossiness perception in shapes other than a sphere, and to explore qualitatively how these effects vary with the shape complexity expressed in depth and curvature.

4.1.2 Stimuli. The same scene and rendering technique was used as in Experiment 1. To study a broad spectrum of stimuli, we varied the same three parameters as in Experiment 1 and also the shape of the object, where $shape \in \{sphere, spiky\ sphere, Stanford\ Lucy, low\ resolution\ Lucy, cylinder\}$ and $alpha \in \{0, 0.05, 0.1, 0.25, 0.5\}$.

The sphere had already been studied in Experiment 1, while Experiment 2 was conducted on four new shapes. Several factors were considered when selecting the shapes: we need a shape that differs from a sphere by surface complexity and curvature, i.e., does not have large curved areas and does not reflect the mirror image of the environment (if you pick it up, you cannot see yourself); has many fine details; is not compact, has thin parts that transmit light well; we selected the *Lucy* from the Stanford 3D Scanning Repository [33], as it satisfies these conditions and has been used in other works for studying the appearance of translucent materials (e.g., Reference [24]). Afterwards, we wanted to isolate several features and selected the following objects: is as thick as a sphere but has more complex surface geometry—*spiky (bumpy) sphere*; has little thickness, similar to Lucy, has thin parts, but lacks fine details, has relatively simple surface geometry and lower curvature—the *low-resolution Lucy*; the main body is as thick as that of Lucy, but lacks thin parts and has very simple surface geometry and a very low curvature—a *cylinder*. The objects are illustrated in Figure 12.

We defined the initial pool of subsurface scattering properties as $\sigma_t \in \{0, 0.1, 0.5, 1, 2, 3, 4, 5, 10\}$ and $albedo \in \{0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$. We performed a clustering process similar to that used in Experiment 1 (described in Section 3.1.2). As the clustering was conducted for each individual shape and surface roughness, the cluster centers were not identical among them. Although the difference was negligible among the surface roughness levels, it was substantial between the sphere and the Lucy. Therefore, we selected two sets of [σ_t-*albedo*] pairs, $\{[0.1, 0.5]; [1.0, 0.9]; [2.0, 0.6]; [3.0, 0.3]; [3.0, 0.95]; [4.0, 0.9]\}$ for spiky sphere (identical parameters had already been used for a sphere in Experiment 1), and $\{[0.5, 0.8]; [1.0, 0.4]; [3.0, 0.4]; [3.0, 0.7]; [3.0, 0.9]; [5.0, 0.1]\}$ for the Lucy, low-resolution Lucy and the cylinder. All images can be found in the supplementary materials (Figures 23–27).

4.1.3 Experimental Design. The experimental design was identical to the first round of Experiment 1. The objects were compared only with the objects of similar shape and *alpha*.

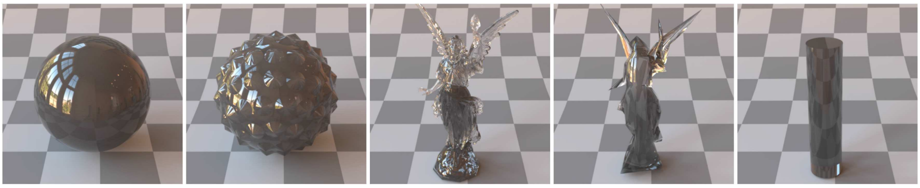

Fig. 12. Five different shapes have been studied throughout the experiment. Left to right: sphere (3.00; 0.30), spiky sphere (3.00; 0.30), Stanford Lucy (5.00; 0.10), low-resolution Lucy (5.00; 0.10), and cylinder (5.00; 0.10). The numbers given in the parentheses are σ_t and albedo, respectively. *Alpha* = 0.00 for all of them.

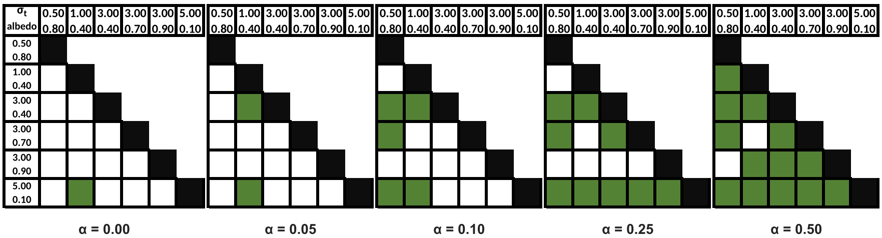

Fig. 13. The results for Lucy. Significance tables for each roughness level. Each lower triangular matrix shows which of the stimuli pairs are significantly different. Green cells—statistically significant difference; white cells—no statistically significant difference. The number of significantly different pairs is larger for rough objects.

4.1.4 Analysis. Similarly to Experiment 1, Binomial tests were conducted to test the null hypotheses for each pair, and Z-scores were calculated to assess the big picture. In addition to this, a scatter plot of Z-scores as a function σ_t and albedo was plotted to identify how these individual parameters of subsurface light transport affect gloss. Finally, we used the variance of the Z-scores and the number of significantly different pairs for a given shape and *alpha*, to compare the magnitude of the subsurface scattering impact on perceptual gloss. The shapes have been quantified in terms of *depth* (thickness) and *surface curvature*. The 3D models were presented in dimensionless units—the radius of a sphere was considered 1, and all other shapes were quantified relative to that. *Depth* was defined as a range of coordinates in all three dimensions separately, covered by the point cloud of a given object. *Local surface curvature* (Gaussian and mean) has been calculated for all points on the object surface [8, 41] and average values have been reported.

4.1.5 Subjects. The procedure was identical to Experiment 1.

4.2 Results

With this experiment, we wanted to answer three questions:

(1) Does subsurface scattering affect gloss for object shapes other than a sphere?
(2) How does the impact of subsurface scattering on gloss co-vary with surface roughness for object shapes other than a sphere?
(3) How do σ_t and albedo relate with the perceived glossiness and how does this differ across the shapes?

4.2.1 Does Subsurface Scattering Impact Gloss? In Experiment 1, we demonstrated with spherical objects that subsurface scattering impacts gloss perception. The results for the Lucy are shown in Figures 13 and 14. Although

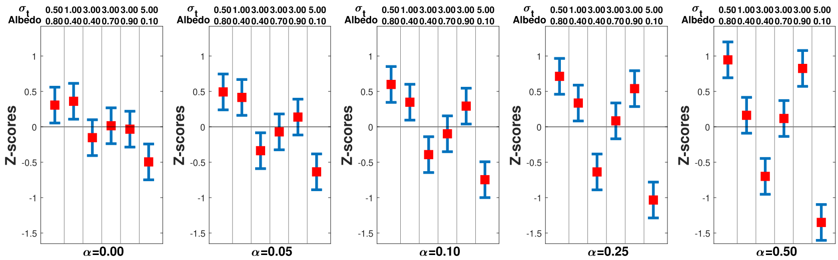

Fig. 14. Z-scores for Lucy. A red cube corresponds to the mean Z-score for a given object, while the error bar corresponds to 95% confidence interval. The difference among Z-scores grows with the increase of roughness.

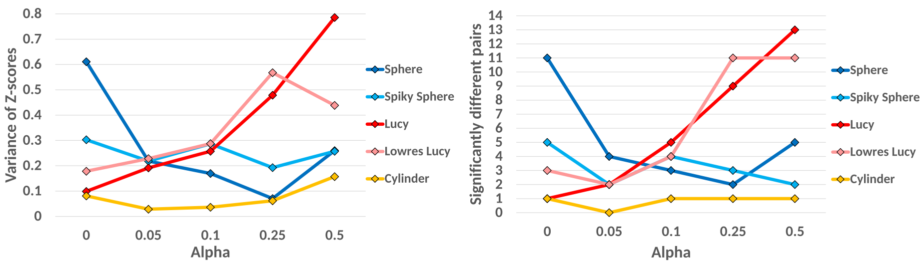

Fig. 15. The variance (left) of the mean Z-scores and the number of significantly different pairs (right). The two metrics are consistent.

the results are not one-to-one comparable with that of a sphere due to the differences in subsurface scattering parameters, the following contradiction in the overall trends still stands out (compare with Figures 6 and 7): the impact is subtle for smooth Lucy objects and the contribution of subsurface scattering increases with *alpha*, while the opposite is true for spherical objects. The null hypothesis was rejected for 13 of 15 pairs when $alpha = 0.5$, while it was rejected for one pair only when $alpha = 0$. The results for the spiky sphere and low-resolution Lucy closely follow the trends of a sphere and Lucy, respectively. Interestingly, a cylinder was the least affected object by the change in subsurface scattering. The detailed results for those shapes can be found in the supplementary materials (refer to Figures 28–36 for all results).

4.2.2 Impact of Alpha Across Different Shapes. We compared the variance of the mean Z-scores, as well as the number of statistically significantly different pairs (of 15) for each shape and *alpha*. The results are shown in Figure 15. As expected, the results are very consistent between the two metrics. The large variance of the Z-scores or the higher number of significantly different pairs means that the variation in subsurface scattering leads to larger gloss differences. The σ_t and albedo parameters used for rendering, although subtly, still differ between a sphere and spiky sphere, on the one hand, and the Lucy, the low-resolution Lucy and the cylinder, on the other hand. This makes it challenging to directly compare the results between the two groups. However, we can still observe how the variance changes with *alpha* for a given shape. For spherical objects, the impact of subsurface scattering on gloss is larger when $alpha = 0$. The impact gradually diminishes as *alpha* increases, but interestingly, the impact starts climbing again when $alpha = 0.5$. Conversely, the impact of subsurface scattering on Lucy-shaped objects increases with the *alpha*. It is also worth noting that the cylinder remains the least affected object for all *alphas*.

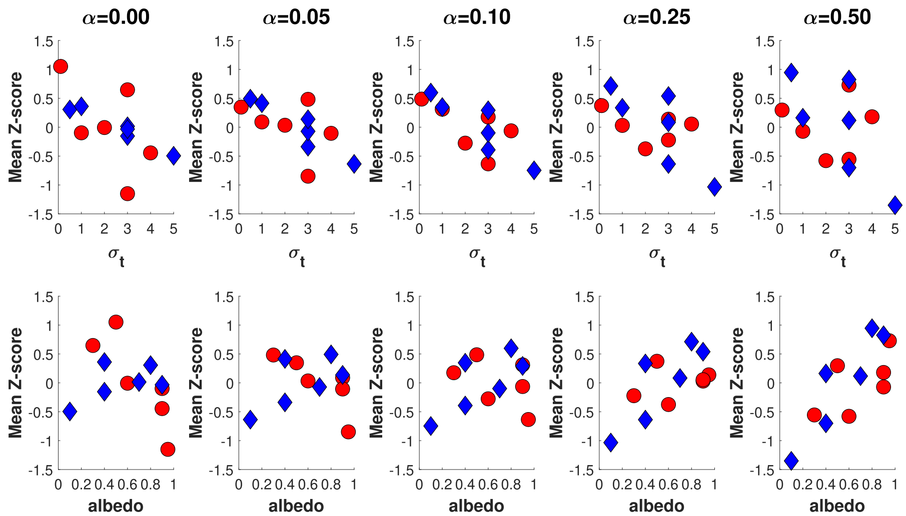

Fig. 16. Z-score as a function of the extinction coefficient (top row) and albedo (bottom). Sphere (red circles) and Lucy (blue diamonds). Linear correlations are apparent for Lucy.

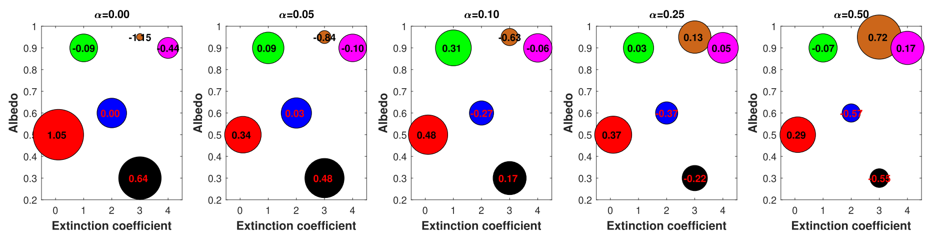

Fig. 17. The results for a sphere. Larger circle diameters represent a higher mean Z-score. Lower albedo and σ_t lead to a glossier look for smoother objects, while the trend changes as the roughness increases. Note that Z-scores are relative to the objects of the same roughness and circles of the same color are not directly comparable among the five plots.

4.2.3 Gloss, σ_t and Albedo. Till now the impact of subsurface scattering on gloss perception was discussed as a whole, single phenomenon. However, for modeling purposes in the future, it is of vital importance to identify how each particular physical attribute relates to the perceived gloss. Mean Z-score as a function of σ_t and albedo is shown in Figure 16, and the mean Z-scores in the σ_t-albedo space are shown in Figures 17 and 18. Interestingly, for Lucy, there is a negative linear correlation between Z-scores and σ_t, and a positive linear correlation between Z-scores and albedo (refer to Figure 19). As for the sphere, the albedo is negatively correlated with Z-scores when *alpha* is low, but it becomes positive for large *alphas* (refer to Figure 11). Figures 17 and 18 show that for both shapes the increase in *alpha* has a negative impact on low albedo materials and a positive impact on high albedo ones. The results for all other shapes are reported in the supplementary materials.

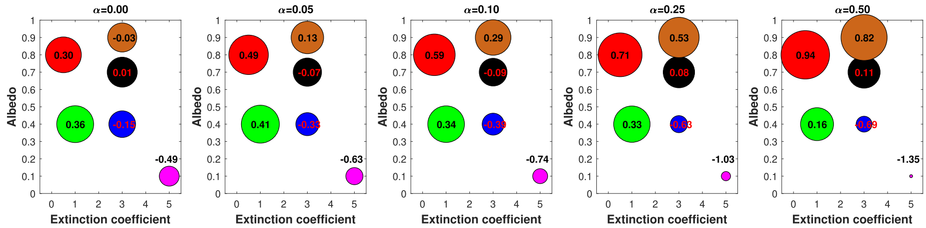

Fig. 18. The results for Lucy. High albedo and low extinction coefficient usually yield glossier stimuli.

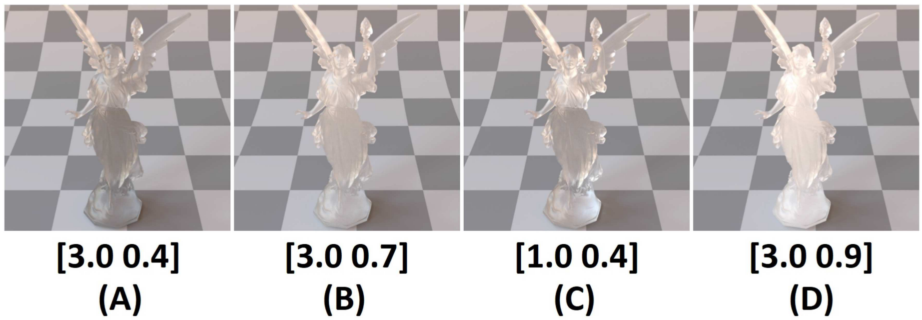

Fig. 19. The numbers in the brackets correspond to σ_t and albedo. *Alpha* = 0.25 for all objects. They can be ranked by glossiness from left to right, the rightmost one being the glossiest (difference between B and C is not significant though). We can observe that although A, B, and D have identical σ_t, higher albedo makes them look glossier, because it generates more highlights which apparently are mistaken for specular reflections. However, A and C have identical albedo, but differ in σ_t. Low σ_t of C generates more caustics, which are also mistaken for specular reflections.

4.3 Discussion

The object shapes come in different surface curvature and thickness (depth). The thickness of the objects is normalized to a unit sphere radius and is shown in Table 1 (columns 1–3). It is an important parameter, because the extinction coefficient is meaningful in terms of object size—the larger the distance light needs to travel within the medium, the larger the probability of absorption and scattering is. In other words, object depth directly impacts the appearance of the dielectric materials. This explains why the trends are similar between a sphere and a spiky sphere, as well as Lucy and low-resolution Lucy. Only subtle differences have been observed between a sphere and a spiky sphere, and between Lucy and low-resolution Lucy. However, an essentially different trend has been observed in cylinders, even though its thickness is nearly identical to the body of Lucy. This observation indicates that thickness does not account for all differences caused by shape and surface complexity—thus, curvature should also be considered.

Local surface curvature has been found on all points of the 3D object and an average value has been calculated. The curvature at a given point can have a positive or a negative sign. However, we are primarily interested in how rugged the overall surface is, and not in the directionality of the curvature, neither in convexity or concavity of the shape. Therefore, the average has been calculated among absolute values. The curvature measure is summarized in Table 1 (columns 4 and 5). Note that both Gaussian and mean curvatures are equal to 1 for a unit sphere, and Gaussian curvature is equal to 0 for a cylinder. Marlow and Anderson [38] demonstrate that

Table 1. The Depth of the Objects in **X**, **Y**, and **Z** Dimensions and Their Curvature

	X	Y	Z	GC	MC
Sphere	2	2	2	1	1
Spiky Sphere	2.09	2.10	2.10	742.81	22.48
Lucy	0.94	1.48	2.73	22691.61	58.44
Lowres. Lucy	0.88	1.48	2.68	89.11	7.61
Cylinder	0.45	0.45	1.90	0	2.48

A sphere and a spiky sphere are larger than the rest. Lucy is the tallest. Although dimensions for Lucy and low-resolution Lucy look substantially larger than that of a cylinder, this is due to the span of Lucy's wings. The approximate size of its body is 0.45 both in X and Y dimensions. The cylinder was designed after the torso of Lucy. Gaussian curvature (**GC**) and mean curvature (**MC**) are found locally for each point of the 3D object. The average of the absolute values is reported. Lucy is the shape with the highest curvature that is no surprise considering its level of fine details.

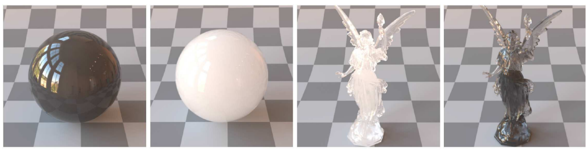

Fig. 20. Although objects A and B have identical shape and surface roughness, the lower albedo of subsurface scattering makes object A more mirror-like. Although spectral reflectance is identical, object A looks darker due to higher absorption inside the volume. Lucies in C and D have identical shape and surface roughness, but higher albedo of C generates more highlights. It is difficult to tell whether the highlights on C are specular reflections, caustics, or result of volume scattering, while specular reflections are easier to isolate on low albedo object D.

the weighted average of sharpness, contrast, and size of the highlights account for most of the variance in gloss judgements. The authors argue that these cues are constrained by the macro-, meso-, and microscale shape of the object. For instance, specular sharpness can vary as a function of curvature, as *"specular reflections will be sharpest in image regions that run parallel to local directions of high curvature, and will be most shallow (stretched) along directions of low curvature."* Their experiments have shown that higher curvature leads to higher specular sharpness and contrast, thus higher glossiness, albeit the correlation with specular coverage is subtle. However, their findings are based on fully opaque media. Sharpness and contrast will certainly be dependent on the light exiting the volume after subsurface light transport. The curvature of the surface can also influence the coverage area (size of the highlights) due to subsurface scattering, as it has been the case for high albedo Lucy in our experiment (image C in Figure 20). This indicates that their findings are not directly transferable to translucent materials. In the future work, cross-shape comparisons are needed (e.g., sphere with Lucy) to identify whether objects with higher curvature look glossier for translucent objects as well.

Interestingly, for low curvature objects, low σ_t materials (transparent) and materials with high σ_t and low albedo (dark opaque) are considered glossiest (refer to the first and fourth images from the left in Figure 3). We conducted an additional experiment with 15 smooth spherical objects and applied the nonclassical nonmetric MDS with raw user response frequency as a distance matrix. From the extracted features, we can see that

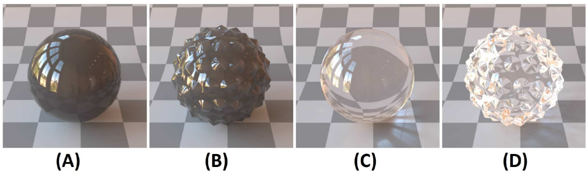

Fig. 21. Curvature influences glossiness cues—thus, the perceived relative glossiness of the objects.

transparent low σ_t and dark opaque materials were placed close to each other in 2D embedding (refer to Figure 37 in the supplementary materials). The same trend holds for higher dimensions. Marlow and Anderson [38] also see similarities between the two types of materials and propose that similar mechanisms might be used in both cases, as the clear image of the surrounding *"inside or behind the depth"* of the object body is visible in both cases—although one is the result of direct transmission, while the other is a mirror reflection image. The mirror reflections on dark opaque objects are intuitively associated with perceived gloss, but the link between the background image seen-through the transparent media and gloss certainly deserves further study.

Curvature could, however, explain the primary difference, as well as similarities in trends between a sphere and a spiky sphere (although we have not compared them directly). For low alpha, a low albedo dark opaque sphere (image A Figure 21) is among the glossiest, while that is not that case for a smooth spiky sphere made of the same material (image B Figure 21). This is because the high curvature of the spiky sphere does not permit a clear mirror reflection to be observed. However, the transparent object is the glossiest for both shapes (images C and D Figure 21). However, the image cues differ dramatically between the two. The transmission image is not visible for a transparent spiky sphere (image D Figure 21), but the curvature of spikes produces shiny highlights due to internal scattering (the resulting image is also affected by the limited dynamic range). Similarly, the lower curvature of low-resolution Lucy makes transparent one glossiest for all alphas, while that is not the case for Lucy, as its curvature does not permit clear transmission.

5 GENERAL DISCUSSION

The results of the two psychometric experiments have enabled us make the following observations:

- Subsurface scattering can impact apparent gloss. This impact depends on micro-scale surface roughness and macro-scale shape of the object.
- Subsurface scattering had larger impact on apparent gloss of smooth spherical objects than on that of rough spherical objects; for complex Lucy shape, the opposite was true—rough Lucy objects being more impacted than smoother ones; the impact of subsurface scattering on apparent gloss was subtle for cylindrical objects.
- For smooth spherical objects, apparent gloss is negatively correlated with albedo, but the correlation is positive for rough spherical objects. For Lucy, apparent gloss is negatively correlated with the extinction coefficient and positively correlated with albedo, regardless of roughness.
- Surface scattering has generally stronger effect on apparent gloss than subsurface scattering. However, in some particular instances, subsurface scattering could compensate for surface scattering effects, yielding equivalent gloss appearance on the objects with different surface roughness.

5.1 The Impact of Subsurface Scattering and Its Dependence on Roughness

The effect of subsurface scattering was statistically significant for numerous material pairs. This is a clear indication that subsurface scattering is a contributing factor to perceived gloss and should be considered in future studies on gloss perception. However, this impact differs among the object shapes. We hypothesize that this difference comes from different image cues present in objects of different shapes and surface roughness. For more opaque smooth spherical objects lower albedo led to a glossier appearance. As the lower curvature of a spherical object produces a distinct reflected image of the environment, we believe that this is a widely used cue by the HVS for glossiness perception. The darker the object, the more distinct the reflected mirror image is. Besides, the contrast between specular and non-specular areas is also large and the reflections stand out more. This phenomenon is demonstrated in Figure 20—objects A and B have an identical shape and surface roughness, but the subsurface scattering albedo of A is substantially lower, which makes it easier to observe the mirror reflection of the environment on it. This is consistent with the previous findings [49, 62]. As the sphere becomes rougher, the reflection of the environment, as well as specular reflections, disappear and the cues used for judgment of glossiness change. As rough objects look all Lambertian and non-glossy, the difference among them decreases. However, objects with higher albedo look lighter and shinier, which could potentially become a cue for glossiness [27, 52–54]. While the impact of *alpha* on gloss is monotonic, the impact of subsurface scattering is not. Qi *et al.* [54] have demonstrated the monotonic relationship between *alpha* and gloss, while they showed that the contribution of meso-scale roughness is non-monotonic. Further study is needed to explain why the impact is non-monotonic for spheres and why it starts increasing for *alpha* = 0.5. It is interesting that for smooth spheres, the materials with the lowest extinction coefficient looked glossiest. We have speculated above that the presence of the transmission image inside the object can be reminiscent of mirror reflection, while the association with familiar material (e.g., glass), as well as caustics could have also played the role.

5.2 Shape-dependence of the Effect

For Lucy-shaped objects, the opposite trend was observed. Usually, the albedo was positively correlated with gloss, the extinction coefficient was negatively correlated, and the overall impact was increasing with the roughness. If we inspect the Lucy-shaped images, then we will see that the surface geometry does not allow to observe a clear reflection image, neither clear specular reflections. Subjects seemingly rely on highlighted areas that result not only from the specular reflections, but from internal scattering and caustics as well. It is difficult to tell which highlight is a specular reflection, which one is caustic, and which ones are produced by subsurface scattering—especially in low-dynamic-range scenarios. Naturally, high albedo objects with lower extinction coefficient produce more highlights. Refer to images C and D in Figure 20. High albedo and limited dynamic range make it challenging to tell whether the highlights of image C were produced by specular reflections or subsurface scattering. The same task is a lot easier when the albedo is low (image D). The size of the highlights has been shown by Marlow and Anderson [38] to be positively correlated with perceived gloss. The curvature of the surface (as in the case of Lucy) can lead to large highlight areas due to high subsurface scattering. Interestingly, all smooth objects were considered equally shiny, while the differences between highlights start to prevail when the roughness is increased, producing a broader range of gloss perception.

These observations are consistent with Gigilashvili et al. [21]. They observed that the impact of translucency on gloss was different between spheres and complex female bust objects, qualitatively similar to Lucy. They interviewed the subjects and learned that the cues used for gloss estimation were different for different shapes, but they were also subject to individual interpretations. Further study is needed to investigate the reasons for the dramatic difference between sphere and Lucy results. Interestingly, the trends were similar between a sphere and a spiky sphere, as well as between Lucy and low-resolution Lucy. We believe this is correlated with the size of the objects. First, spheres and spiky spheres cover larger field-of-view, having a more apparent reflection of the environment than a low-resolution Lucy, which has simple surface geometry itself, but still occupies too little space of the field of view to reflect clear images of the environment. Second, translucency varies with the thickness of the object [14, 19] and the path light travels inside the volume is indeed more similar between a sphere and a spiky sphere than between a thick sphere and thin Lucy. However, these speculations need concrete experimental evidence. However, a cylinder is the least affected shape by subsurface scattering. The reason for this could be

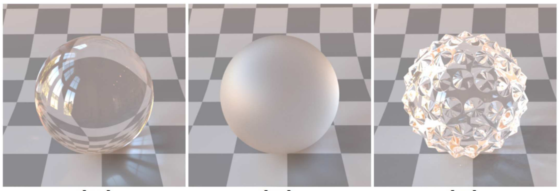

Fig. 22. The structure of the image A provides more cues on how to segment reflection and transmission components, while the task looks considerably more difficult for images B and C.

the fact that its curved surface enables a clear reflection image for all smooth ones, while the rough ones resemble in highlight coverage cues—in the end yielding little difference among the cylinders with the same *alpha*.

5.3 Surface versus Subsurface Scattering

We have observed that surface roughness usually has a stronger impact on material glossiness than subsurface scattering. However, we have also demonstrated notable examples when subsurface scattering effects compensated for surface roughness and smoother objects did not appear glossier. Interestingly, both surface roughness and subsurface scattering blur non-specular areas—both generating similar image-level measurements in these regions. If the surface is smooth and sharp specular reflections are visible, then the two cases can be effortlessly distinguished (because surface roughness, unlike subsurface scattering, blurs specular highlights too). However, estimating the contribution of subsurface scattering becomes increasingly difficult with the rougher surfaces (see B in Figure 22). It would be an interesting future direction to study, how adept the HVS is to estimate the contribution of the subsurface scattering when surface scattering is high, or when specular highlights are superimposed on the rendering of a rough object.

6 LIMITATIONS AND FUTURE WORK

This work has been the first attempt to explore how subsurface scattering contributes to apparent gloss. The materials addressed in this study represent a tiny subset of all possible materials that can exist around us. To keep the number of experimental stimuli within the manageable limits, we had to fix multiple intrinsic and extrinsic parameters, which also implies that our findings come with particular limitations, which need to be addressed in future works:

- We used isotropic phase function and wavelength-independent σ_t and albedo. Subsurface scattering in most real materials has large spectral and spatial variation. Materials with wavelength-dependent subsurface scattering (chromatic effects) and non-isotropic phase functions should be studied in the future. The phase function has been shown to be important for material appearance [24]. The authors provided two-dimensional perceptual embedding of the phase functions, where the dimensions modulate diffuse translucent and sharp, detailed, glass-like appearances, respectively. We hypothesize that the latter could be correlated with apparent gloss.
- While the index of refraction has been fixed to 1.5 in our experiments, we believe other indices of refraction also deserve attention in the future.
- We used Beckmann microfacet normal distribution to modulate surface scattering parameter. It is interesting to explore, whether our findings hold if the surface roughness is modeled with other distributions, such as Phong [30] or GGX [67]. We hypothesize that the impact will be negligible, as the clustering of

a large pool of parameters will converge to relatively similar appearances. However, this needs further study and experimental evidence.

- Although we plot Z-scores as a function of σ_t and albedo, the effects of the two parameters need to be studied separately and more in depth. The future experiments could include comparisons for each σ_t and albedo, separately. It is also important to explore the potential interaction between these two parameters. We believe that there is a significant interaction between the effects of the two parameters. For example, the impact of albedo can be large for high σ_t, but it becomes negligible when σ_t is very low. We believe a mixed effects statistical model is needed to describe the correlation between gloss and subsurface light transport, while σ_t, albedo and *alpha* can be treated as fixed effects, random effects, such as user physiological and display characteristics, should be also included.
- Illumination conditions have been fixed throughout the experiment. It has been shown before that illumination geometry affects both translucency [14, 73] and gloss [15, 48]. Therefore, the study should be extended to other illumination geometries.
- As a metric for clustering, Euclidean distance could be substituted with more perception-aware metrics, such as L^4-norm [50], the cubic root metric used by Gkioulekas *et al.* [24] or the appearance similarity metric proposed by Lagunas *et al.* [34]. Additionally, the perceptual accuracy could be improved if the comparisons were done in the CIELAB space instead of RGB [50, 61]. However, using RGB usually biases chromatic information [50, 61]. As our stimuli have been mostly achromatic, we believe the comparison in the RGB space has not introduced any significant bias in the clustering process.

Besides, addressing the research question from the perspective of image-based measurements has been beyond the scope of this work. However, we believe that future works should investigate how subsurface light transport affects image structure and statistics, which proposedly are glossiness cues. This could bring to light ***why*** and ***how*** subsurface scattering contributes to apparent gloss.

First, our results once again illustrate that no one-to-one correspondence between physical and perceptual properties exists and that our ability to segment specular reflections from image structure is limited [40]. This is why users might have mistaken caustics for specular reflections. The image-level intensities result from a combination of reflection and transmission. Unmixing those is an ill-posed problem and the HVS uses different constraints for this task, such as, apparent object shape [40]. While smooth spherical and cylindrical shapes facilitate separation of specular and non-specular components, the task becomes increasingly difficult for complex geometries. For instance, in Figure 22, it is easier to separate reflection and transmission components in image A than it is for images B and C. We hypothesize that additional factors that usually facilitate this segmentation, such as motion, binocular vision or surface texture [9, 56, 69] could decrease the impact of subsurface scattering on apparent gloss. This could explain why many users tied all physical objects in previous works when interaction was permitted [19, 21].

However, the users still saw a glossiness difference, even when segmenting specular and non-specular components should have been relatively simple—particularly, in the case of smooth spherical objects. We believe this happened because apparent gloss is not a function of apparent specular reflection only, but it also depends on extrinsic factors that are independent from specular reflections, such as lightness of the non-specular areas [49].

It remains an open question exactly which image cues and which psycho-visual mechanisms of gloss perception are affected by the subsurface scattering, and rigorous future work is needed to answer it. Similarly to Marlow and Anderson [38], psychophysical studies should be conducted in the future to measure how different image-level measurements, such as perceived coverage, sharpness and contrast of the highlights co-vary with the perceived glossiness of the materials of different shapes and light transport properties. This will help us understand the differences observed in this article, and the robustness of the state-of-the-art will also be tested in the context of light-transmissive media. Moreover, particular image statistics should be studied to quantify and model the impact of subsurface scattering on the gloss cues in the image space. Additional interviews with the subjects could potentially help with the identification of the most salient cues and interpreting the results. Particularly, eye tracking experiments in the controlled conditions could provide deeper insight into the actual image cues used for glossiness assessment. And last but not least, we believe that perceived gloss is *at least* two-dimensional—distinctness and contrast, as proposed by Pellacini et al. [49], being the major perceptual dimensions of gloss, even for translucent objects. However, the model quantifying these perceptual dimensions should include σ_t and albedo along with other physical parameters, to enable accurate placement of the translucent stimuli in the perceptual gloss space. We have observed in Experiment 1 that for high σ_t, when the light does not penetrate deep into the volume, the processes and findings are phenomenologically similar to Ward's model used by Pellacini et al. [49]. MDS similar to Reference [49] could reveal how σ_t and albedo contribute to distinctness and contrast, given that the stimuli are sampled densely enough in σ_t-albedo space. With that being said, we believe a separate embedding might be needed for each *alpha*, as the HVS might apply different internal perceptual functions to the stimuli with different roughnesses (i.e., with different gloss cues).

Our findings have practical implications for computer graphics, perception, as well as material appearance measurement and reproduction research. They show that material appearance modelling should be done on the shape we are particularly interested in and generalization of the findings based on one shape or surface roughness should be taken with extreme care. We also propose that future gloss perception research should include materials that permit subsurface light transport and the perceptual models of gloss should be updated so that they could account for potential contribution from subsurface scattering. Finally, gloss measurement protocols should accommodate translucent materials.

7 CONCLUSION

We have conducted psychophysical experiments to test whether subsurface scattering of light contributes to gloss perception and to characterize this impact qualitatively and quantitatively. The results support our hypothesis and provide ample evidence that gloss perception is impacted by subsurface scattering. The impact varies across shapes and surface roughness levels; this we believe is the result of different low- and high-level image cues being used (by the HVS) for different shapes to assess gloss. Our findings propose that modelling appearance should be taken with care and findings should not be generalized to other shapes and surface scattering models. Moreover, the state-of-the-art findings based on fully opaque materials might not be valid for transparent and translucent media. Understanding ***why*** subsurface light transport contributes to apparent gloss and ***how*** it is used by the HVS would be an important future direction. Eventually, a higher number of stimuli (ideally in HDR) will be needed to build a complete perceptual space of gloss. We believe the future work addressing gloss perception should not be limited to fully opaque materials and the perceptual models should account for subsurface scattering. Rigorous work is needed in the future to identify the exact mechanisms for predicting perceptual gloss from materials' surface and subsurface light transport properties.

REFERENCES

[1] ASTM. 2017. Standard Terminology of Appearance. ASTM International, West Conshohocken, PA. Retrieved from https://www.astm.org/Standards/E284.htm. https://doi.org/10.1520/E0284-17.
[2] Edward H. Adelson. 2001. On seeing stuff: The perception of materials by humans and machines. In *Human Vision and Electronic Imaging VI*, Vol. 4299. International Society for Optics and Photonics, 1–12.
[3] Barton L. Anderson and Juno Kim. 2009. Image statistics do not explain the perception of gloss and lightness. *J. Vision* 9, 11:10 (2009), 1–17.
[4] Jacob Beck and Slava Prazdny. 1981. Highlights and the perception of glossiness.*Percept. Psychophys.* 30, 4 (1981), 407–410.
[5] Alice C. Chadwick, George Cox, Hannah E. Smithson, and Robert W. Kentridge. 2018. Beyond scattering and absorption: Perceptual unmixing of translucent liquids. *J. Vision* 18, 11:18 (2018), 1–15.
[6] Nahian S. Chowdhury, Phillip J. Marlow, and Juno Kim. 2017. Translucency and the perception of shape. *J. Vision* 17, 3:17 (2017), 1–14.
[7] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences (2nd ed.).* Erlbaum, Hillsdale, NJ.
[8] Alireza Dastan. 2020. Gaussian and mean curvatures calculation on a triangulated 3D surface. Retrieved from https://www.mathworks.com/matlabcentral/fileexchange/61136-gaussian-and-mean-curvatures-calculation-on-a-triangulated-3d-surface.

[9] Katja Doerschner, Roland W. Fleming, Ozgur Yilmaz, Paul R. Schrater, Bruce Hartung, and Daniel Kersten. 2011. Visual motion and the perception of surface material. *Curr. Biol.* 21, 23 (2011), 2010–2016.
[10] Peter G. Engeldrum. 2000. *Psychometric Scaling: A Toolkit for Imaging Systems Development.* Imcotek.
[11] Christian Eugène. 2008. Measurement of "total visual appearance": A CIE challenge of soft metrology. In *Proceedings of the 12th IMEKO TC1 & TC7 Joint Symposium on Man, Science, and Measurement.* 61–65.
[12] Franz Faul. 2019. The influence of Fresnel effects on gloss perception. *J. Vision* 19, 13:1 (2019), 1–39.
[13] Roland W. Fleming. 2014. Visual perception of materials and their properties. *Vision Res.* 94 (2014), 62–75.
[14] Roland W. Fleming and Heinrich H. Bülthoff. 2005. Low-level image cues in the perception of translucent materials. *ACM Trans. Appl. Percept.* 2, 3 (2005), 346–382.
[15] Roland W. Fleming, Ron O. Dror, and Edward H. Adelson. 2003. Real-world illumination and the perception of surface reflectance properties. *J. Vision* 3 (2003), 347–368.
[16] Brendan J. Frey and Delbert Dueck. 2007. Clustering by passing messages between data points. *Science* 315, 5814 (2007), 972–976.
[17] Davit Gigilashvili, Fereshteh Mirjalili, and Jon Yngve Hardeberg. 2019. Illuminance impacts opacity perception of textile materials. In *Color and Imaging Conference.* Society for Imaging Science and Technology, 126–131.
[18] Davit Gigilashvili, Midori Tanaka, Marius Pedersen, and Jon Yngve Hardeberg. 2020. Image statistics as glossiness and translucency predictor in photographs of real-world objects. In *Proceedings of the 10th Colour and Visual Computing Symposium (CVCS'20).* CEUR Workshop Proceedings, 15 pages.
[19] Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen. 2018. Behavioral investigation of visual appearance assessment. In *Proceedings of the Color and Imaging Conference.* Society for Imaging Science and Technology, 294–299.
[20] Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg. 2019. Material appearance: Ordering and clustering. In *Proceedings of the IS&T International Symposium on Electronic Imaging.* Society for Imaging Science and Technology, 202:1–202:6.
[21] Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg. 2019. Perceived glossiness: Beyond surface properties. In *Proceedings of the Color and Imaging Conference.* Society for Imaging Science and Technology, 37–42.
[22] Davit Gigilashvili, Philipp Urban, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen. 2019. Impact of shape on apparent translucency differences. In *Proceedings of the Color and Imaging Conference.* Society for Imaging Science and Technology, 132–137.
[23] Ioannis Gkioulekas, Bruce Walter, Edward H. Adelson, Kavita Bala, and Todd Zickler. 2015. On the appearance of translucent edges. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 5528–5536.
[24] Ioannis Gkioulekas, Bei Xiao, Shuang Zhao, Edward H. Adelson, Todd Zickler, and Kavita Bala. 2013. Understanding the role of phase function in translucent appearance. *ACM Trans. Graph.* 32, 5 (2013), 1–19.
[25] Phil J. Green. 2003. A Colour Engineering Toolbox. Retrieved from http://www.color.org/resources/ColourEngineeringToolbox.zip.
[26] Dar'ya Guarnera, Giuseppe Claudio Guarnera, Matteo Toscani, Mashhuda Glencross, Baihua Li, Jon Yngve Hardeberg, and Karl R. Gegenfurtner. 2018. Perceptually validated analytical BRDFs parameters remapping. In *Proceedings of the ACM SIGGRAPH Talks.* 1–2.
[27] Yun-Xian Ho, Michael S. Landy, and Laurence T. Maloney. 2008. Conjoint measurement of gloss and surface texture. *Psychol. Sci.* 19, 2 (2008), 196–204.
[28] Sture Holm. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 2 (1979), 65–70.
[29] Richard S. Hunter. 1937. Methods of determining gloss. *NBS Res. Paper RP* 958 (1937), 19–39.
[30] Wenzel Jakob. 2010. Mitsuba renderer. Retrieved from http://www.mitsuba-renderer.org.
[31] Iona S. Kerrigan and Wendy J. Adams. 2013. Highlights, disparity, and perceived gloss with convex and concave surfaces. *J. Vision* 13, 1:9 (2013), 1–10.
[32] Juno Kim, Phillip Marlow, and Barton L. Anderson. 2011. The perception of gloss depends on highlight congruence with surface shading. *J. Vision* 11(9), 4 (2011), 1–19.
[33] Stanford University Computer Graphics Laboratory. 1994. The Stanford 3D Scanning Repository. Retrieved from http://graphics.stanford.edu/data/3Dscanrep/.
[34] Manuel Lagunas, Sandra Malpica, Ana Serrano, Elena Garces, Diego Gutierrez, and Belen Masia. 2019. A similarity measure for material appearance. Retrieved from https://arXiv:1905.01562.
[35] Michael S. Landy. 2007. A gloss on surface properties. *Nature* 447, 7141 (2007), 158–159.
[36] Frédéric B. Leloup, Gael Obein, Michael R. Pointer, and Peter Hanselaer. 2014. Toward the soft metrology of surface gloss: A review. *Color Res. Appl.* 39, 6 (2014), 559–570.
[37] Phillip Marlow, Juno Kim, and Barton L. Anderson. 2011. The role of brightness and orientation congruence in the perception of surface gloss. *J. Vision* 11(9), 16 (2011), 1–12.
[38] Phillip J. Marlow and Barton L. Anderson. 2013. Generative constraints on image cues for perceived gloss. *J. Vision* 13, 14:2 (2013), 1–23.
[39] Phillip J. Marlow, Juno Kim, and Barton L. Anderson. 2012. The perception and misperception of specular surface reflectance. *Curr. Biol.* 22, 20 (2012), 1909–1913.
[40] Phillip J. Marlow, Dejan Todorović, and Barton L. Anderson. 2015. Coupled computations of three-dimensional shape and material. *Curr. Biol.* 25, 6 (2015), R221–R222.

[41] Mark Meyer, Mathieu Desbrun, Peter Schröder, and Alan H. Barr. 2003. Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and Mathematics III.* Springer, 35–57.
[42] mfa Boston CAMEO. 2020. Paraffin Wax. Retrieved from http://cameo.mfa.org/wiki/Paraffin_wax.
[43] Isamu Motoyoshi. 2010. Highlight-shading relationship as a cue for the perception of translucent and transparent materials. *J. Vision* 10, 9:6 (2010), 1–11.
[44] Isamu Motoyoshi, Shin'ya Nishida, Lavanya Sharan, and Edward H. Adelson. 2007. Image statistics and the perception of surface qualities. *Nature* 447, 7141 (2007), 206–209.
[45] Takehiro Nagai, Yuki Ono, Yusuke Tani, Kowa Koida, Michiteru Kitazaki, and Shigeki Nakauchi. 2013. Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study. *i-Percept.* 4, 6 (2013), 407–428.
[46] Shin'ya Nishida, Isamu Motoyoshi, Lisa Nakano, Yuanzhen Li, Lavanya Sharan, and Edward Adelson. 2008. Do colored highlights look like highlights? *J. Vision* 8, 6 (2008), 339.
[47] Gaël Obein, Kenneth Knoblauch, and Françoise Viénot. 2004. Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *J. Vision* 4 (2004), 711–720.
[48] Maria Olkkonen and David H. Brainard. 2011. Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Percept.* 2, 9 (2011), 1014–1034.
[49] Fabio Pellacini, James A. Ferwerda, and Donald P. Greenberg. 2000. Toward a psychophysically based light reflection model for image synthesis. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques.* ACM Press/Addison-Wesley, 55–64.
[50] Thiago Pereira and Szymon Rusinkiewicz. 2012. Gamut mapping spatially varying reflectance with an improved BRDF similarity metric. In *Computer Graphics Forum*, Vol. 31. Wiley Online Library, 1557–1566.
[51] Zygmunt Pizlo. 2001. Perception viewed as an inverse problem. *Vision Res.* 41, 24 (2001), 3145–3161.
[52] Lin Qi, Mike J. Chantler, J. Paul Siebert, and Junyu Dong. 2012. How mesoscale and microscale roughness affect perceived gloss. In *Proceedings of the 3rd International Conference on Appearance.* Lulu Press, 48–51.
[53] Lin Qi, Mike J. Chantler, J. Paul Siebert, and Junyu Dong. 2014. Why do rough surfaces appear glossy? *J. Optic. Soc. Amer. A* 31, 5 (2014), 935–943.
[54] Lin Qi, Mike J. Chantler, J. Paul Siebert, and Junyu Dong. 2015. The joint effect of mesoscale and microscale roughness on perceived gloss. *Vision Res.* 115 (2015), 209–217.
[55] Ralph L. Rosnow and Robert Rosenthal. 2003. Effect sizes for experimenting psychologists.*Can. J. Experiment. Psychol.* 57, 3 (2003), 221.
[56] Yuichi Sakano and Hiroshi Ando. 2010. Effects of head motion and stereo viewing on perceived glossiness. *J. Vision* 10, 9:15 (2010), 1–14.
[57] Masataka Sawayama, Yoshinori Dobashi, Makoto Okabe, Kenchi Hosokawa, Takuya Koumura, Toni Saarela, Maria Olkkonen, and Shin'ya Nishida. 2019. Visual discrimination of optical material properties: A large-scale study. Retrieved from https://www.biorxiv.org/content/10.1101/800870v2.full.
[58] Alexandra C. Schmid, Pascal Barla, and Katja Doerschner. 2020. Material category determined by specular reflection structure mediates the processing of image features for perceived gloss. Retrieved from https://www.biorxiv.org/content/10.1101/2019.12.31.892083v1. https://doi.org/10.1101/2019.12.31.892083.
[59] Scientific Polymer Products, Inc.2020. Refractive Index of Polymers. Retrieved from https://scientificpolymer.com/technical-library/refractive-index-of-polymers-by-index/.
[60] Lavanya Sharan, Ruth Rosenholtz, and Edward H. Adelson. 2014. Accuracy and speed of material categorization in real-world images. *J. Vision* 14, 9:12 (2014), 1–24.
[61] Tiancheng Sun, Ana Serrano, Diego Gutierrez, and Belen Masia. 2017. Attribute-preserving gamut mapping of measured BRDFs. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 47–54.
[62] Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Gabriele Simone. 2017. Image contrast measure as a gloss material descriptor. In *Proceedings of the International Workshop on Computational Color Imaging*. Springer, 233–245.
[63] Louis L. Thurstone. 1927. A law of comparative judgment. *Psychol. Rev.* 34, 4 (1927), 273–286.
[64] Matteo Toscani, Dar'ya Guarnera, Giuseppe Claudio Guarnera, Jon Yngve Hardeberg, and Karl R. Gegenfurtner. 2020. Three perceptual dimensions for specular and diffuse reflection. *ACM Trans. Appl. Percept.* 17, 2 (2020), 1–26.
[65] Kristi Tsukida and Maya R. Gupta. 2011. *How to Analyze Paired Comparison Data.* Technical Report. University of Washington, Department of Electrical Engineering, Seattle, WA.
[66] Peter Vangorp, Jurgen Laurijssen, and Philip Dutré. 2007. The influence of shape on the perception of material reflectance. In *ACM Transactions on Graphics*, Vol. 26. ACM, 77:1–77:10.
[67] Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. 2007. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics Conference on Rendering Techniques.* Eurographics Association, 195–206.
[68] Gregory J. Ward. 1992. Measuring and modeling anisotropic reflection. In *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques.* 265–272.

[69] Gunnar Wendt, Franz Faul, Vebjørn Ekroll, and Rainer Mausfeld. 2010. Disparity, motion, and color information improve gloss constancy performance. *J. Vision* 10, 9:7 (2010), 1–17.
[70] Maarten W. A. Wijntjes and Sylvia C. Pont. 2010. Illusory gloss on Lambertian surfaces. *J. Vision* 10, 9:13 (2010), 1–12.
[71] Josh Wills, Sameer Agarwal, David Kriegman, and Serge Belongie. 2009. Toward a perceptual space for gloss. *ACM Trans. Graph.* 28, 4 (2009), 1–15.
[72] Bei Xiao and David H. Brainard. 2008. Surface gloss and color perception of 3D objects. *Visual Neurosci.* 25, 3 (2008), 371–385.
[73] Bei Xiao, Bruce Walter, Ioannis Gkioulekas, Todd Zickler, Edward Adelson, and Kavita Bala. 2014. Looking against the light: How perception of translucency depends on lighting direction. *J. Vision* 14, 3:17 (2014), 1–22.
[74] Bei Xiao, Shuang Zhao, Ioannis Gkioulekas, Wenyan Bi, and Kavita Bala. 2019. Effect of geometric sharpness on translucent material perception. Retrieved from https://www.biorxiv.org/content/10.1101/795294v1.

Received October 2020; revised February 2021; accepted March 2021

Supplementary Material for: The Role of Subsurface Scattering in Glossiness Perception

DAVIT GIGILASHVILI, Norwegian University of Science and Technology, Norway
WEIQI SHI and ZEYU WANG, Yale University, USA
MARIUS PEDERSEN and JON YNGVE HARDEBERG, Norwegian University of Science and Technology, Norway
HOLLY RUSHMEIER, Yale University, USA

A SUPPLEMENTARY MATERIALS

A.1 Stimuli

Figures 1–5 illustrate all stimuli used in the experiments.

1544-3558/2021/05-ART10
https://doi.org/10.1145/3458438

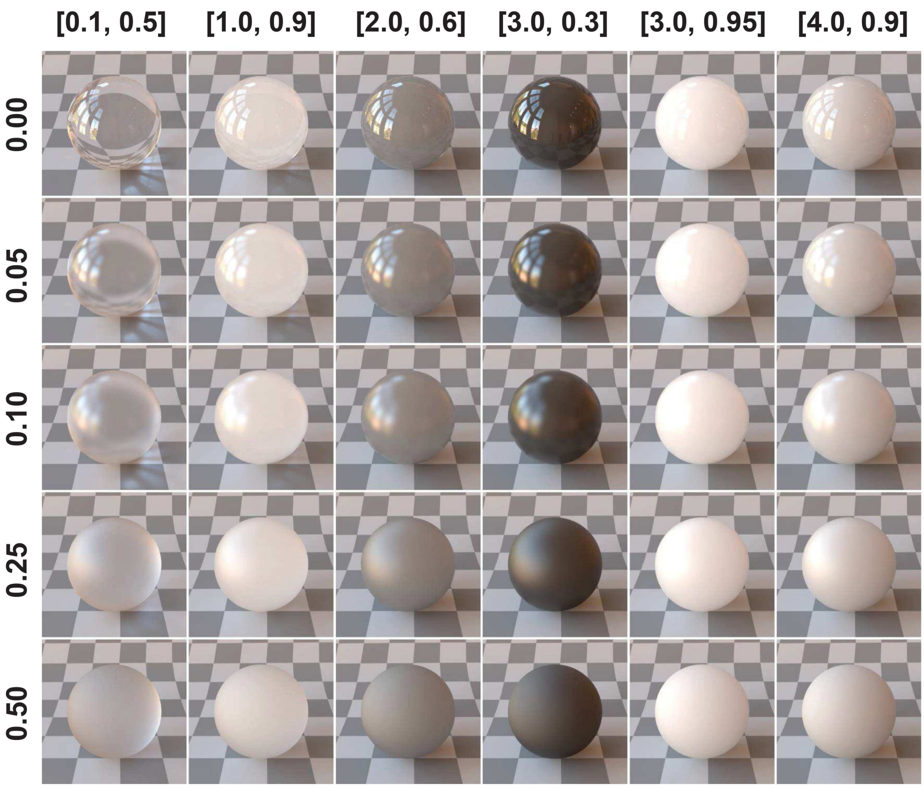

Fig. 1. All spherical objects used in the experiments. Columns correspond to σ_t and albedo, while rows correspond to different *alphas*, respectively.

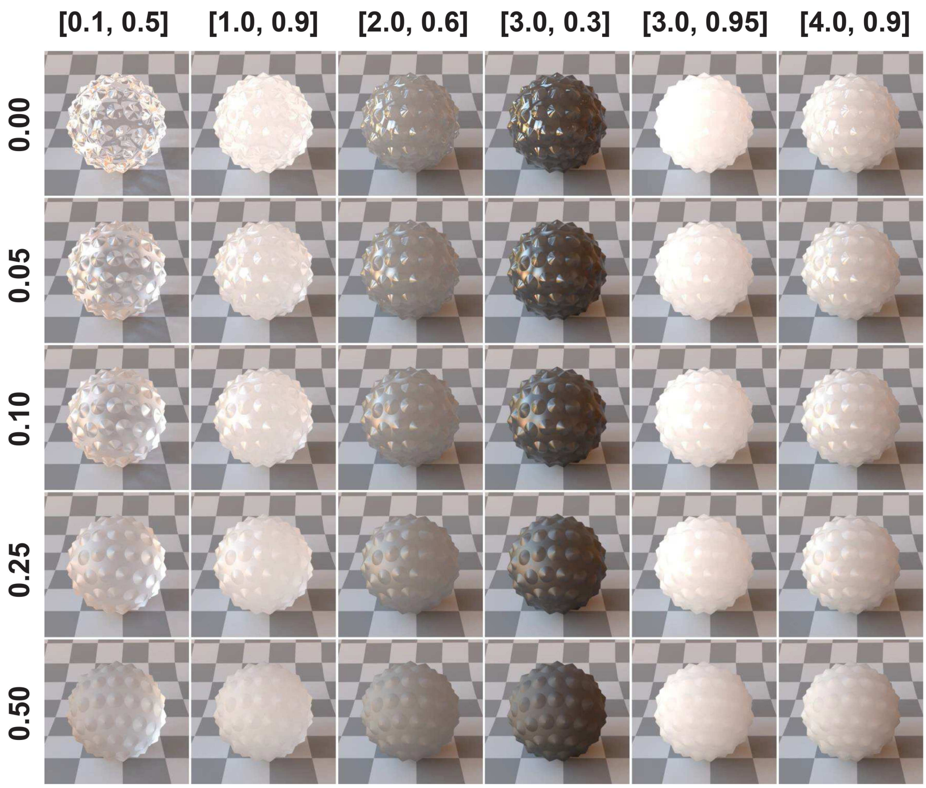

Fig. 2. All spiky sphere objects used in the experiments. Columns correspond to σ_t and albedo, while rows correspond to different *alphas*, respectively.

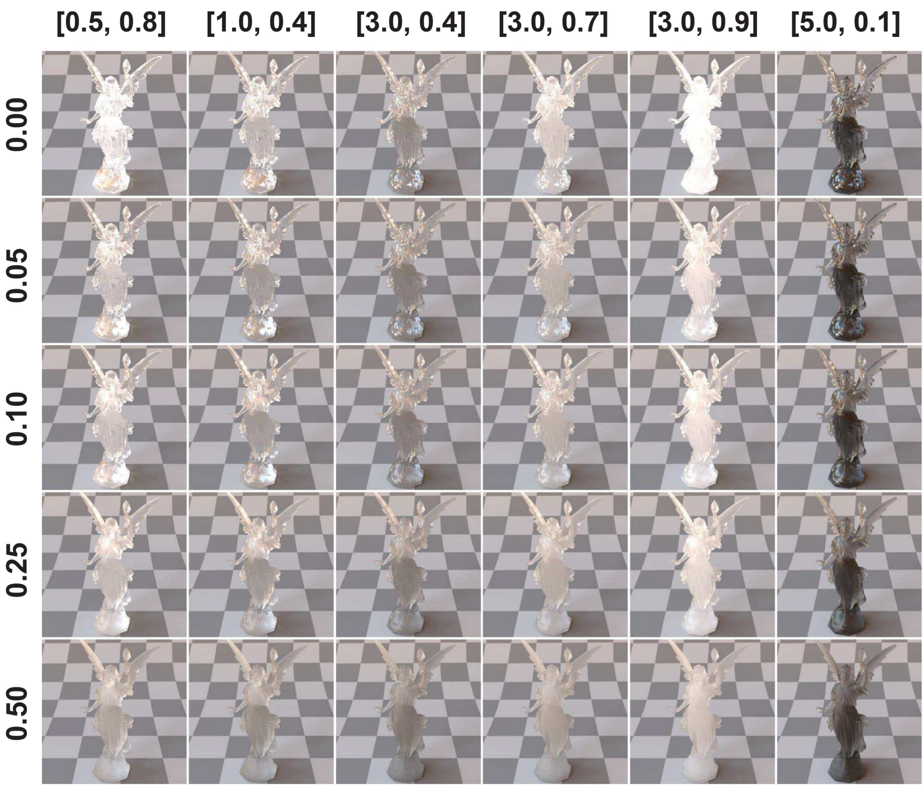

Fig. 3. All Stanford Lucy objects used in the experiments. Columns correspond to σ_t and albedo, while rows correspond to different *alphas*, respectively.

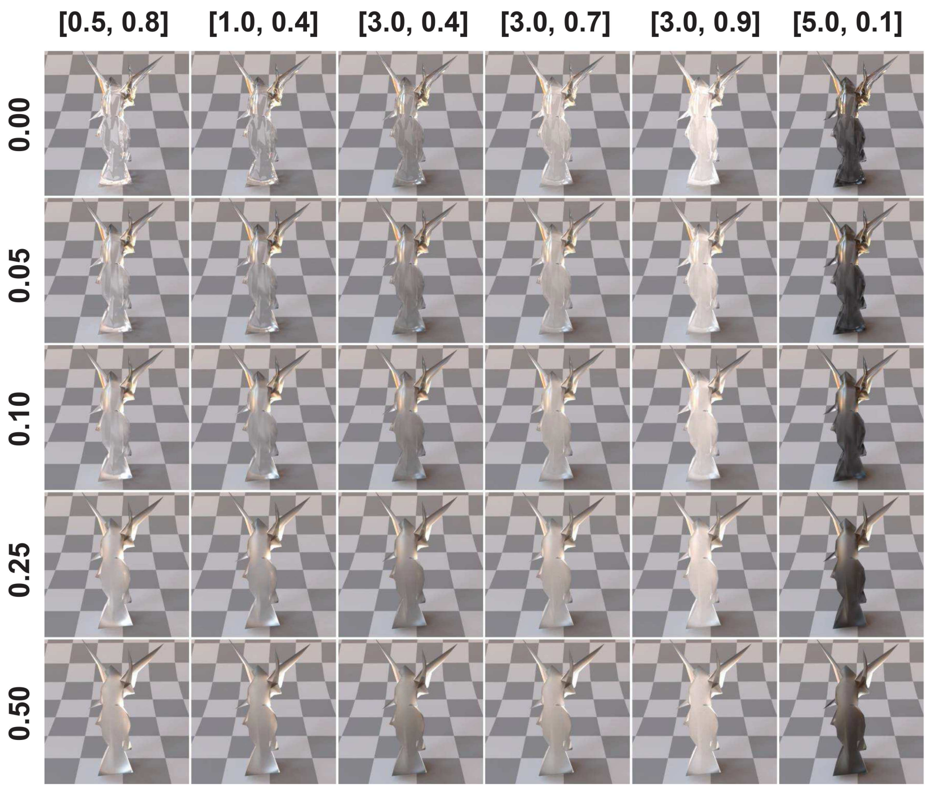

Fig. 4. All low-resolution Lucy objects used in the experiments. Columns correspond to σ_t and albedo, while rows correspond to different *alphas*, respectively.

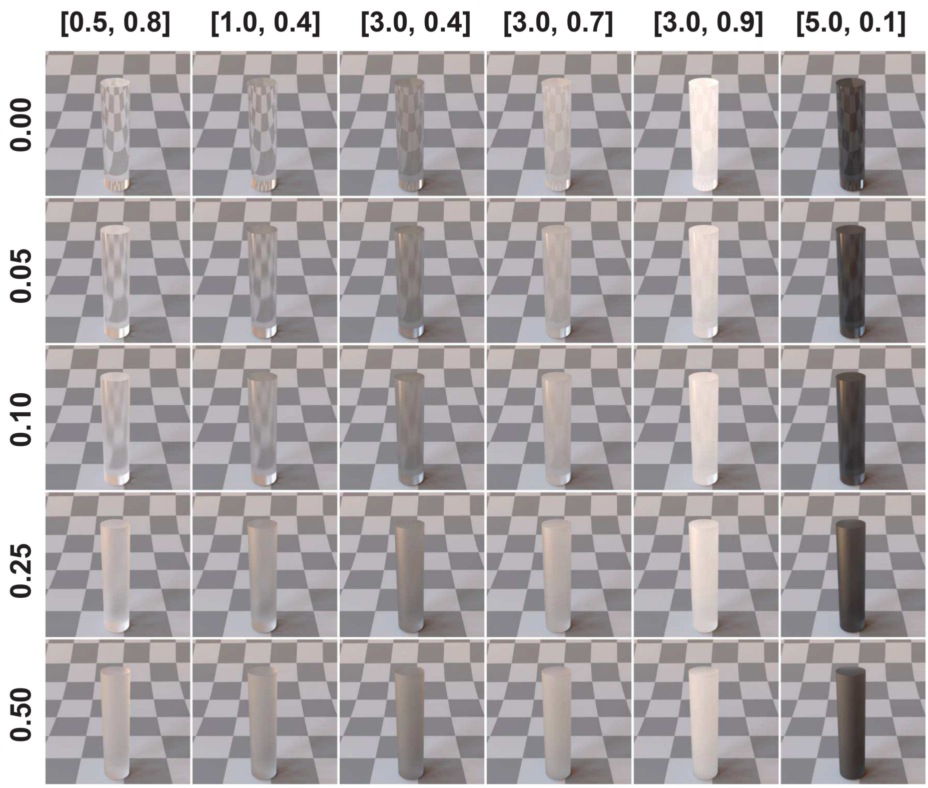

Fig. 5. All cylindrical objects used in the experiments. Columns correspond to σ_t and albedo, while rows correspond to different *alphas*, respectively.

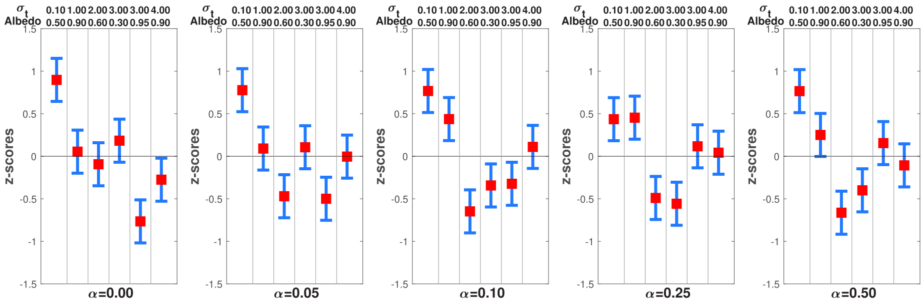

Fig. 6. Results for the spiky sphere. A red cube corresponds to the mean Z-score for a given object, while the error bar corresponds to 95% confidence interval. The difference among Z-scores diminishes with the increase in roughness.

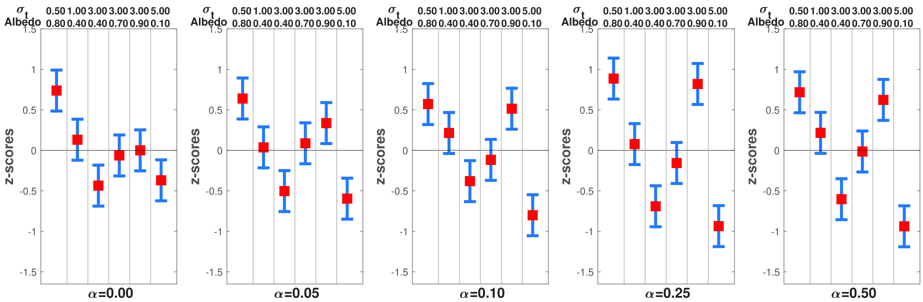

Fig. 7. Results for low-resolution Lucy. The difference among Z-scores climbs with the increase in roughness.

A.2 Z-scores

Z-score plots have been reported for spherical and the Lucy shapes in the main body of the manuscript. Below Z-scores are reported for three remaining shapes: spiky sphere, low-resolution Lucy, and a cylinder, in Figures 6–8, respectively. The overall trends between a sphere and spiky sphere, as well as between Lucy and low-resolution Lucy are largely similar. However, there are still some subtle differences. For instance, a low-albedo spherical object is among the glossiest for lower *alphas*. However, that is not true for a spiky sphere. The absence of the mirror-like reflection of the environment in spiky spheres due to surface curvature could be the explanation for this fact. There is a difference between Lucy and low-resolution Lucy as well. For low *alphas*, the transparent low σ_t low-resolution Lucy is the glossiest one, while that is not the case for the Lucy made of the identical material. The transparent low-resolution Lucy, similar to a sphere, permits seeing the background through the object, while that is not possible for Lucy due to high surface curvature. This can be an indication that background plays a role in gloss perception and this similarity between a sphere and the low-resolution Lucy should be scrutinized in the future studies. Finally, the impact of subsurface scattering on the glossiness of cylindrical objects is minimal in comparison with spherical and Lucy-shaped objects.

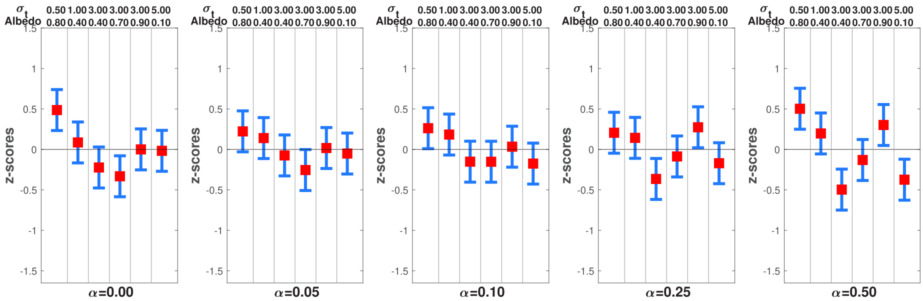

Fig. 8. Z-scores for a cylinder. Subsurface scattering has a subtle effect on the glossiness of cylinders and the majority of the objects with equal alpha looks equally glossy, except for very high alphas.

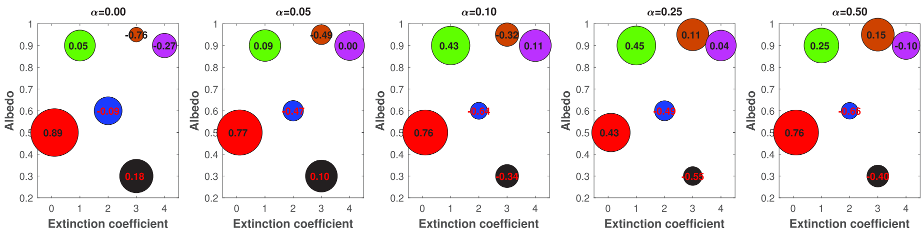

Fig. 9. Results for the spiky sphere. Larger circle diameters represent a higher mean Z-score. Similar to a sphere, lower albedo and σ_t lead to a glossier look for smoother objects, while the trend changes as the roughness increases. Note that Z-scores are relative to the objects of the same roughness and circles of the same color are not directly comparable among the five plots.

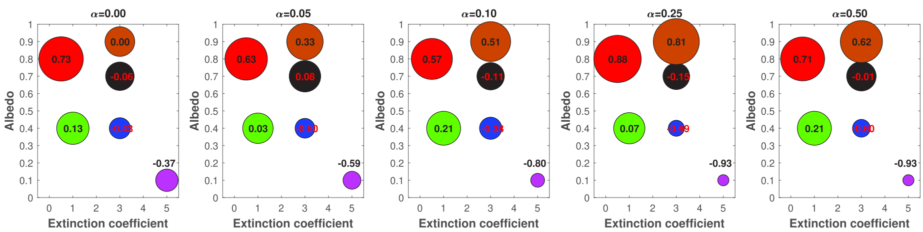

Fig. 10. Results for low-resolution Lucy. Similar to Lucy, high albedo usually yields glossier stimuli.

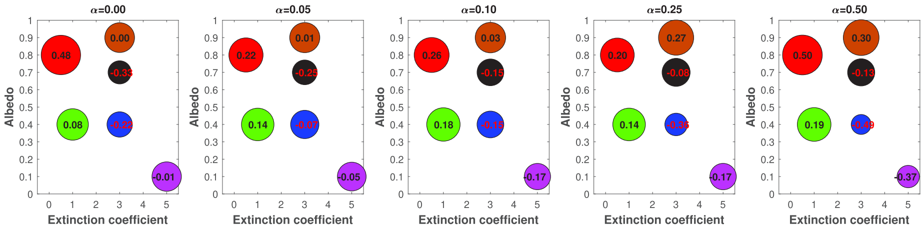

Fig. 11. Results for cylinder.

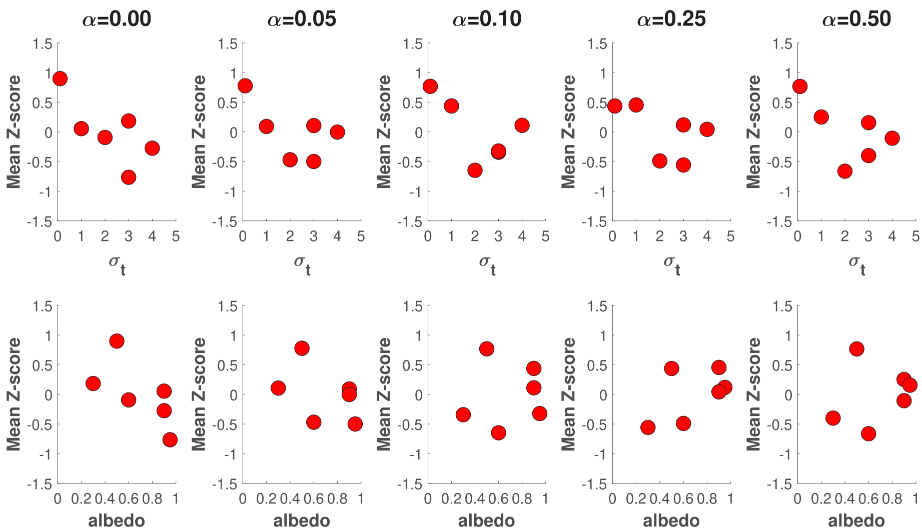

Fig. 12. Z-score as a function of extinction coefficient (top row) and albedo (bottom) for the spiky sphere.

A.3 Z-scores in σ_t-albedo Space

The variation of the Z-scores in σ_t-albedo space shows how these individual parameters affect gloss (Figures 9–11). The trends are similar between a sphere and a spiky sphere, as well as between the Lucy and low-resolution version of it. For spiky sphere, the larger circles are concentrated on the lower end of the σ_t axis, meaning that materials with lower σ_t usually look glossier. The tendency of black and brown circles in Figure 9 manifests how the negative correlation between albedo and Z-scores gradually changes into a positive one as the *alpha* increases. In the case of low-resolution Lucy, all large-diameter circles are concentrated in the high albedo part of the space (Figure 10), while the diameters do not differ largely in the case of cylinders (Figure 11), being less affected by σ_t-albedo variations.

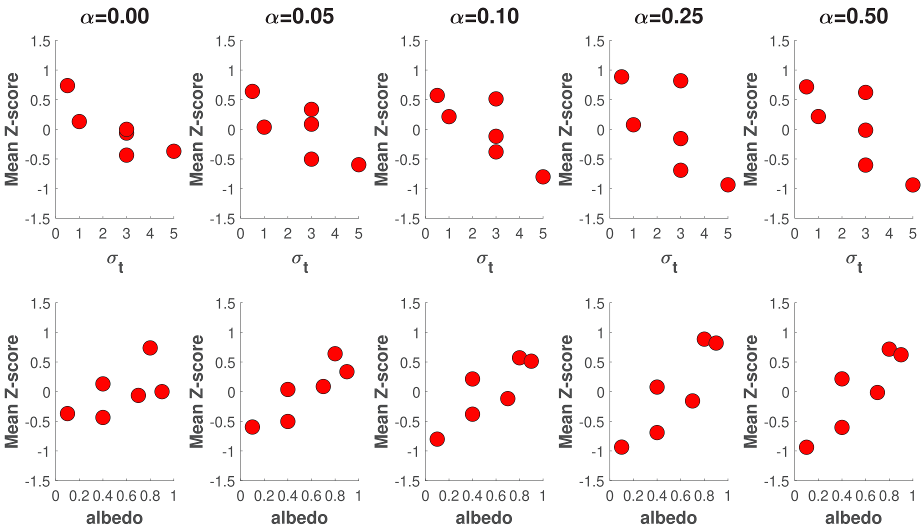

Fig. 13. Z-score as a function of extinction coefficient (top row) and albedo (bottom) for low-resolution Lucy.

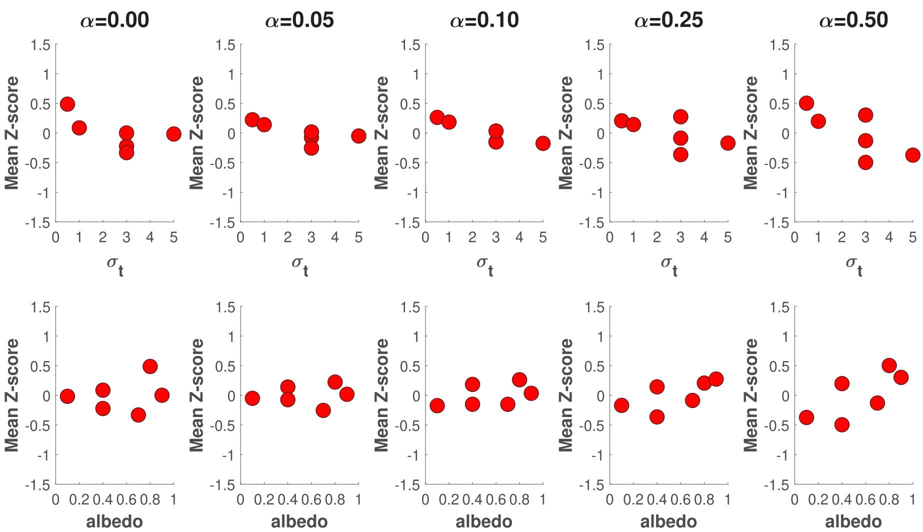

Fig. 14. Z-score as a function of extinction coefficient (top row) and albedo (bottom) for a cylinder.

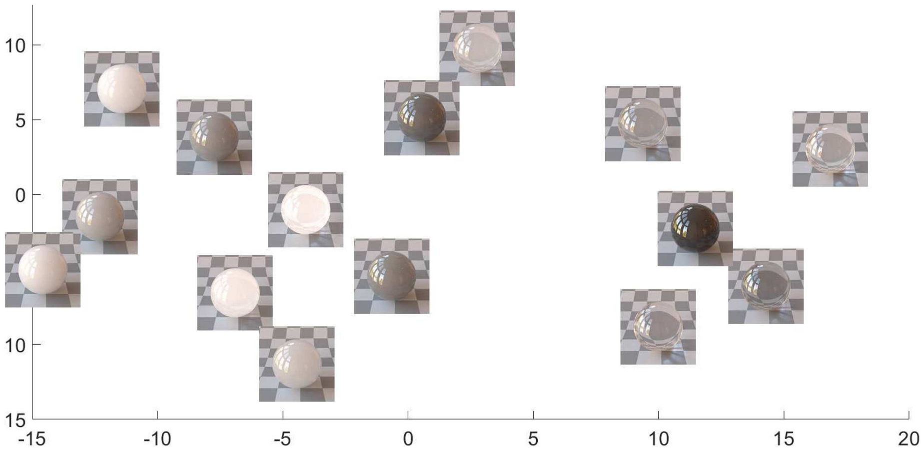

Fig. 15. Two-dimensional embedding of the 15 stimuli. It is apparent that transparent and dark opaque materials are located near each other. The same trend holds for higher-dimensional embeddings.

A.4 Z-scores as a Function of σ_t and Albedo

Z-scores are shown as a function of σ_t and albedo in Figures 12–14. In the case of a spiky sphere (Figure 12), the negative correlation can be seen between Z-scores and both physical parameters, only when *alpha* is low. In the plots of low-resolution Lucy (Figure 13), similar to Lucy, the negative correlation can be seen with the extinction coefficient, and the positive correlation is apparent with an albedo that becomes even stronger as the *alpha* increases. Being consistent with the prior reasoning, neither parameters affect the Z-scores of a cylinder.

A.5 Multidimensional Scaling

An additional experiment has been conducted using 15 smooth-surfaced spherical stimuli sampled in σ_t-albedo space. Interesting proximity between transparent low σ_t and dark opaque (high-σ_t and low-albedo) objects has been observed in two-dimensional, three-dimensional, as well as in the higher-dimensional embeddings.

Article E

Davit Gigilashvili, Philipp Urban, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg (n.d.). “The Impact of Optical and Geometrical Thickness on Perceived Translucency Differences.” In: *Under review in a journal*, 13 pages

The Impact of Optical and Geometrical Thickness on Perceived Translucency Differences

Davit Gigilashvili*, Philipp Urban*†, Jean-Baptiste Thomas*, Marius Pedersen*, Jon Yngve Hardeberg*
* Norwegian University of Science and Technology, Department of Computer Science; Gjøvik, Norway
† Fraunhofer Institute for Computer Graphics Research IGD; Darmstadt, Germany,

Abstract—In this work we study the perception of suprathreshold translucency differences for the purposes of expanding the knowledge about material appearance perception in imaging and computer graphics, as well as 3D printing applications. Translucency is one of the most considerable appearance attributes which significantly impacts the look of objects and materials. However, the knowledge about translucency perception remains limited. Even less is known about the perception of translucency differences between materials. We hypothesize that humans are more sensitive to small changes in absorption and scattering coefficients when optically thin materials are examined and when objects have geometrically thin parts. To test these hypotheses we generated images of objects with different shapes and subsurface scattering properties and conducted psychophysical experiments with these visual stimuli. The analysis of the experimental data supports these hypotheses. Additionally, based on post-experiment comments made by the observers, we argue that the results could be a demonstration of a fundamental difference between transparency and translucency perception mechanisms.

Index Terms—Material appearance, translucency, transparency

1 Introduction

TRANSLUCENCY is one of the major appearance attributes, significantly impacting the look of different objects and materials [2], [11]. The ASTM Standard Terminology of Appearance [1] defines the term as *"the property of a specimen by which it transmits light diffusely without permitting a clear view of objects beyond the specimen and not in contact with it."* Translucent appearance is usually the result of objects and materials permitting some degree of subsurface light transport. The visual stimulus that evokes the perception of translucency in the human visual system (HVS) is impacted by a multitude of intrinsic and extrinsic factors. Intrinsic factors include the optical material properties found in the *radiative transfer equation (RTE)*. Namely, wavelength-dependent absorption and scattering coefficients, wavelength-dependent scattering phase function, and wavelength-dependent index of refraction. The average distance travelled by a photon inside the material before it gets absorbed or scattered is called *mean free path*, which is defined as $1/(\sigma_a + \sigma_s)$, where σ_a and σ_s are wavelength-dependent coefficients of absorption and scattering, respectively. This means that when absorption and scattering coefficients are low, a photon in average travels a larger distance in a straight line within this material. A high absorption coefficient results in fewer photons escaping and exiting the material, while a higher scattering coefficient results in more photons redirected towards a different direction, i.e. less of the light structure is preserved and the scene behind the object becomes more blurry or completely visually occluded. Which direction a photon is redirected to after each scattering event is another important aspect. The distribution of this directionality is given by a scattering phase function.

Materials with large mean free path can be referred to as *optically thin*, while materials with short mean free path can be referred to as *optically thick*. Objects could look nearly transparent either because of a large mean free path permitted by low absorption and scattering coefficients, or because simply the object is geometrically thin, i.e. the distance that a photon needs to travel through the material is shorter, and consequently, the likelihood of absorption or scattering event is also lower. Therefore, the visual stimuli reaching the human retina are significantly affected by both - the geometrical thickness of the object and the optical thickness of the material it is made of.

Finally, some extrinsic factors also impact translucency assessment by human observers. Those are the conditions under which a given object is observed, such as illumination direction [38] and the color of the surface a translucent object is placed on [14].

Qualitative and quantitative understanding of translucency perception is an interesting topic in academia and industry alike - 3D printing being a vivid illustration of the latter. Accurate reproduction of spatially-varying color and translucency have been made possible by the recent advances in multi-material 3D printing [6]. Multi-material 3D printing enables mixing transparent printing materials with opaque colored materials, which considerably broadens the appearance gamut of the printers, i.e. the range of different possible looks of a 3D-printed object. On the other hand, as discussed above, object's shape and geometry can impact the appearance. This generates a need for a geometry-adaptive adjustment of the printing material mixing ratios when transferring translucency appearance from one shape to another. For proper communication and quality assurance in this process, quantification of visual translucency is needed, which could come in a form of a joint color and translucency space incorporating perceived color and translucency difference metrics. However, how translucency difference is perceived by the HVS, and how various factors contribute to that process, remain unanswered to date.

In this paper, we hypothesize that:

1) *Presence of thin parts in the object's shape increases perceived translucency differences.* It has been shown that objects with a complex shape possess a broader range of translucency cues, permit discrimination of more levels of translucency, and fail the constancy of perceived translucency faster [31], [38]. Besides, presence of thin parts affect the magnitude of perceived translucency, as the likelihood for a scattering or absorption event is lower than it is in structurally thicker parts of the same object. Xiao *et al.* [39] have recently shown that geometrically smooth and geometrically sharp objects made of an identical material differ in translucency appearance.
2) *Humans are more sensitive to translucency differences in optically thin materials.* The hypothesis is derived from the notion that transparency and translucency cues used by the HVS are essentially different and the background distortion, which is only present in see-through, optically thin materials [13], [32], is a stronger indicator of subsurface light transport differences than luminance contrast variations, which proposedly is a cue in non-see-through objects.

To analyze these hypotheses and to generate further research hypotheses on the topic, we have conducted two psychophysical experiments under controlled viewing conditions. The objective of the experiments was to identify the distance in absorption-scattering physical parameter space needed for visual detection of a suprathreshold translucency difference. The distance was measured psychophysically for different object shapes and was compared among five different regions in the absorption-scattering space.

The preliminary results for the first experiment were reported in Gigilashvili *et al.* [18]. The article provided the qualitative discussion around the *hypothesis 1* and it proposed the *hypothesis 2*, which inspired the second experiment. This manuscript extends the work by quantitative analysis of both hypotheses and by quantifying shape differences in terms of surface-to-medial-axis histograms.

Our major contributions in this paper are the following:

- We experimentally study how the object's shape, namely, structural (geometrical) thickness and presence of thin parts, impacts perception of translucency differences between materials. The impact is analyzed both quantitatively as well as qualitatively.
- We experimentally test the hypothesis that the HVS is more sensitive to variations in scattering and absorption coefficients in optically thin materials than in optically thick ones.
- We comment on the validity of *Alpha* differences (proposed by Urban *et al.* [36]) across variations in shape.
- We propose new hypotheses for translucency perception research.

2 Background

Fleming and Bülthoff [12] have noted that different models of transparency perception (such as Metelli's episcotister model of color fusion [26]) cannot explain perception of translucency in materials like cheese, milk and wax, as the background is not visible though the object and transparency perception cues (such as X-junctions [3]) are absent. They proposed that the HVS does not invert the optical process of light and matter interaction (*inverse optics* hypothesis [29]) to understand the intrinsic properties of a material, but instead it relies on simple image cues to assess translucency. The potential cues could be the occluded scene seen through the object, if the object is thin enough, either optically or geometrically, and also luminance statistics in particular image regions, when it is not possible to see the background through the material.

Motoyoshi [27] has shown that luminance contrast characteristics in non-specular regions of the object could potentially be a translucency cue for the HVS. Nagai *et al.* [28] have shown that perceived magnitude of translucency correlates well with the local luminance statistics, although the most informative region varies from image to image. Gkioulekas *et al.* [19] observed that edges contain the essential portion of the information needed for translucency assessment, while Sawayama *et al.* [31] proposed that rugged surface of the object facilitates discrimination of translucency. Both findings indicate that the parts where a photon needs to travel shortest distance contain the most of information about material translucence. Moreover, Xiao *et al.* [38] demonstrated that the Stanford Lucy shape [33] permits visual discrimination of more different degrees of translucency than a simple torus shape, proposedly attributed to its complex shape *"with thick and thin sections and features at multiple scales"*. Gigilashvili *et al.* [16] observed that thin parts increase perceived magnitude of translucency and could, in some cases, evoke similar magnitude of translucency as done by optically thinner but structurally (geometrically) thicker objects. Further factors that have been proposed to be affecting the perceived magnitude of translucency are illumination geometry [12], [15], [38] and scattering phase function [20].

Regardless these advances, multiple fundamental points remain yet to be clarified about perceptual translucency, such as perceptual dimensions of translucency, its relation with transparency and opacity, the extent of so called *translucency constancy* and the definition of the term across different contexts [17]. Urban *et al.* [36] have recently proposed *Alpha* - a nearly perceptually uniform measure of translucency, which links optical properties of a material with the magnitude of perceived translucency it evokes in humans. *Alpha* is software- and hardware independent, also adjustable according to the object's scale and suited for 3D printing applications. The authors used virtual homogeneous materials for the psychophysical experiments, in order to define the psychometric function. They used the method of constant stimuli, where the anchor pair was composed of optically thin materials with suprathreshold translucency difference shaped as the Stanford Happy Buddha [33]. This study inspired our work and we used similar virtual stimuli and experimental protocol.

Although the preliminary results for one of our two experiments (reported in [18]) revealed some indications that the presence of thin parts increase perceived translucency differences, the major observation was that human observers are more sensitive to suprathreshold translucency differences in optically thin materials. The entire experiment was based on an optically thin anchor pair, which was compared with both optically thin, as well as optically thick test pairs. This raised a concern about optically thin anchor pair as an universal and objective measure for suprathreshold translucency difference. Therefore, we replicated the experiment with an optically thick anchor pair. In depth analysis of both experiments, as well as a comparison between the two is reported in the subsequent sections.

3 Methodology

3.1 Objective

The primary objective of *Experiment 1* was to identify whether geometrically thin parts increase the magnitude of perceived translucency difference between two materials. *Experiment 2* was inspired by the results of *Experiment 1*. It was conducted to determine how sensitivity to translucency differences depends on the optical thickness of the materials.

3.2 Experimental Design

In order to determine suprathreshold translucency differences for each object's shape, the method of constant stimuli has been used [10]. This is a popular method for determining suprathreshold color differences [4], [35], and has been used for translucency as well [36], [37]. The anchor pair which consists of two Buddhas with a suprathreshold translucency difference and which remains fixed throughout the whole experiment is compared with a test pair. The test pair consists of two similarly-shaped objects - one control point (CP) and one respective test sample (see Section 3.3.1). The task of an observer is to identify which pair has larger difference in perceived translucency. The following instruction was given *"Please, select a pair, either a top, or a bottom one, with higher translucency difference"* without explicitly defining translucency. The observers had to use arrow keys of a standard keyboard to select the respective pair. The pairs were displayed atop each other on a neutral gray background. The two images within a pair were separated with a 3-pixel gap. The position of the pairs (whether the anchor is the top or the bottom one) and within each pair (left or right) was randomized. A sample comparison from the experiment is shown in Figure 1. The only difference between *Experiment 1* and *Experiment 2* was the anchor pair, as they were based on see-through and non-see-through anchor pairs, respectively. All other procedures were identical between the two experiments. 5 CPs $\times$ 4 directions $\times$ 5 samples per direction $\times$ 5 shapes, totalling to 500 comparisons per experiment.

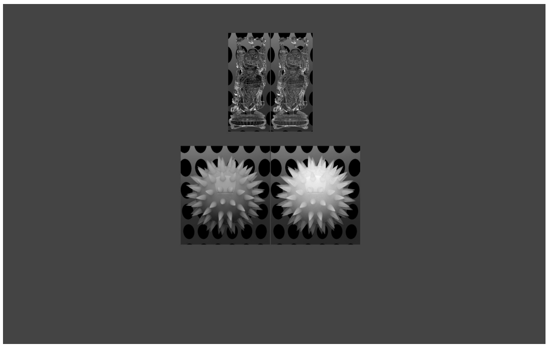

Fig. 1: An example of the comparison shown during the experiment. The top pair represents a see-through anchor pair with a suprathreshold translucency difference. The bottom pair is a test pair which consists of a control point (left) and test sample (right) materials. The task of the observer is to determine which pair has larger translucency difference.

3.3 Stimuli

We used a set of simple virtual materials, similar to those used by Urban *et al.* [36]. All stimuli were presented on a calibrated display using a *Virtual Viewing Booth* [36]. The CIE D65 diffuse light source is located on the ceiling, which is typical to what we encounter on a daily basis - both indoors and outdoors (refer to [36] for the full specification of the *Virtual Viewing Booth*). We used Monte-Carlo Bidirectional Path Tracer in the Mitsuba Physically-based Renderer [22] to solve the RTE. The minimum path depth was set to 20 and "Russian Roulette" termination was deployed afterwards.

In order to keep the degrees of freedom within the manageable range, identical surface roughness, scattering phase function and indices of refraction were used for all objects. In particular, we used perfectly smooth microfacet-scale surface roughness and isotropic phase function. The refractive index of the outer medium was set to 1 (vacuum), while that of materials was fixed to 1.3, which is characteristic for water and typically has small Fresnel reflection [36] (large portion of the light is refracted towards the subsurface).

3.3.1 Test pairs

We have varied only three parameters: absorption and scattering coefficients, which were assumed to be wavelength-independent, and shape. We selected five control points (CPs) in absorption-scattering physical parameter space, covering both optically thin and optically thick regions. The materials corresponding to the CPs are illustrated in Figure 2. All different absorption-scattering coefficient pairs used throughout the experiment are given in Table 1. For each CP, five sample points were selected on each of the four directions in absorption-scattering space. For each direction, we ensured that perceived translucency difference between a CP and at least one sample point was smaller than that of the anchor pair, and larger for at least one other sample point. In the selection process, we relied on *Alpha*-differences ($\Delta Alpha$ [36]) and visual inspection in

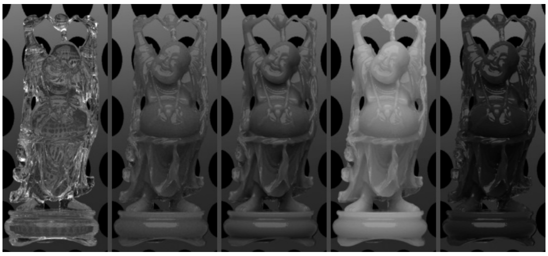

Fig. 2: Five CPs have been selected in the absorption-scattering space. The CPs 1-5 from left to right, respectively, illustrated on the example of Happy Buddha shape.

a trial-and-error manner. The exact procedure we used for sampling in different directions is given below:

- H 1-dimensional (horizontal) change - increasing scattering coefficient for CPs 1 and 2; decreasing scattering coefficient for CPs 3, 4 and 5; fixed absorption. Marked beige in Table 1.
- V 1-dimensional (vertical) change - increasing absorption for CPs 1 and 2; and decreasing absorption for CPs 3, 4, and 5; fixed scattering. Marked dark yellow in Table 1.
- D1 2-dimensional change (diagonal 1) - increasing absorption and scattering for CPs 1 and 2; decreasing absorption and scattering for CPs 3, 4, and 5. All points were located on a straight line defined as σ_a = σ_s for CPs 1-3; σ_a = σ_s - 116 for CP 4; and σ_a = σ_s + 116 for CP 5, where σ_s is scattering, and σ_a is absorption. Marked green in Table 1.
- D2 2-dimensional change (diagonal 2) - increasing absorption and decreasing scattering with points located on the following straight lines: σ_s = -σ_a + 9 for CP 1; σ_s = -σ_a + 155 for CP 2; σ_s = -σ_a + 301 for CP 3; and σ_s = -σ_a + 185 for CPs 4 and 5 (blue in Table 1).

These materials were shown in five different shapes: a Happy Buddha from the Stanford 3D Scanning Repository [33] (identical to the one used in [36]), a perfect sphere (with a radius of 5cm) and three additional spheres with 100 spikes of different length and thickness (illustrated in Figure 3). Adding spikes introduces geometrically thin areas on a compact spherical object that are easier for a photon to go through and that according to our hypothesis increases the magnitude of perceived translucency differences.

3.3.2 Anchor pairs

Both anchor pairs were composed of Buddha figures with suprathreshold translucency differences. The optically thin anchor pair (see-through) was identical to the one used in [36], using scattering coefficients of 0 cm^{-1} and 1.5 cm^{-1}.

The optically thick anchor pair (non-see through) was composed of Buddhas with similar absorption coefficient equal to 77.5 cm^{-1} and different scattering coefficients equal to 77.5 cm^{-1} and 142.48 cm^{-1}, respectively. In order to ensure that the perceptual difference in optically thick anchor pair was equivalent to that of the original anchor pair, we relied on the results from *Experiment 1*. Namely, one Buddha of the non-see-through anchor pair was identical to CP2, and the second one was T50-distance away from it (the distance to T50-point, where each of the two stimuli is selected by the 50% of the observers), as per *Experiment 1*. Our selection was also supported with nearly matching $\Delta Alpha$ in the two pairs. Both anchor pairs are shown in Figure 4.

3.4 Display and Viewing Conditions

Both experiments were conducted on the same color-calibrated display and under the same viewing conditions. The stimuli were displayed on a 24.1 inch EIZO ColorEdge CG246 LCD, which was calibrated to CIE D65 white point with gamma equal to 2.2. Konica Minolta CS-2000 spectroradiometer was used to measure the luminance of the monitor displaying a perfect diffuser patch. The maximum measured luminance was 196 $^{cd}/_{m^2}$ with 6542K color temperature. The monitor was warmed up for at least 30 minutes before each experiment.

The experiment took place in a completely dark room, where the display was the only light source. The distance between an observer and the display was roughly 60cm. 148×348 pixel images were used to display the anchor pairs, occupying 3.81° of the visual field horizontally and 8° vertically. The height of the test pair images was 348 pixels, while the width varied depending on the shape. The largest image with 348×348 pixels was used to display spheres with the longest spikes. All test pairs occupied 8° vertically and 6.20°, 6.48°, 8.10°, and 8.96° horizontally, for spheres with no, short, medium and long spikes, respectively.

3.5 Observers

27 observers including three co-authors of this article have participated in the experiments. 18 were male and 9 were female, representing 20 different nationalities. The median age was 31 years with standard deviation equal to 10.97. All of them passed a Snellen visual acuity test to make sure that they had normal or corrected-to-normal vision. As wavelength-independent absorption and scattering yield grayscale stimuli, color vision of the observers was not tested. 21 observers had a background in color science, imaging, vision, material appearance or related fields.

The comparisons were distributed over three sessions. Each comparison has been assessed by 20 observers. No comparisons were assessed by all observers, and not all observers were shown all comparisons. 23 participants assessed comparisons from both experiments, while 4 observers assessed comparisons from *Experiment 1* only.

3.6 Analysis

3.6.1 Probit Analysis and T50 Distances

We are interested to learn how far we need to move from the CP in the absorption-scattering space to notice the difference in perceived translucency. As discussed above, we hypothesize that this distance depends on the shape and it is shorter for the objects with thin parts. We also hypothesize that this distance is shorter for optically thin see-through materials.

First of all, we conducted a frequency analysis of the observer responses to observe how they change as the Euclidean distance in the absorption-scattering space increases

TABLE 1: The experimental stimuli differ with their locations in the absorption-scattering space. The table summarizes absorption and scattering coefficients used to render these stimuli. Each pair of columns represents one control point. Control point coordinates are marked red, while the test samples in four different directions are labeled with beige, dark yellow, green, and blue cells, respectively.

Control Point 1		Control Point 2		Control Point 3		Control Point 4		Control Point 5	
Scattering	Absorption	Scattering	Absorption	Scattering	Absorption	Scattering	Absorption	Scattering	Absorption
4.5	**4.5**	**77.5**	**77.5**	**150.5**	**150.5**	**150.5**	**34.5**	**34.5**	**150.5**
4.7	4.5	78	77.5	148	150.5	145	34.5	30	150.5
6	4.5	82	77.5	141	150.5	135	34.5	20	150.5
7	4.5	90	77.5	121	150.5	125	34.5	15	150.5
10	4.5	110	77.5	100	150.5	110	34.5	5	150.5
20	4.5	130	77.5	71	150.5	95	34.5	0	150.5
4.5	4.8	77.5	80	150.5	148	150.5	30	34.5	145
4.5	6	77.5	100	150.5	140	150.5	23	34.5	130
4.5	8	77.5	120	150.5	125	150.5	15	34.5	120
4.5	12	77.5	150	150.5	100	150.5	7	34.5	100
4.5	20	77.5	200	150.5	80	150.5	0	34.5	50
4.7	4.7	85	85	140	140	146	30	30	146
5.5	5.5	95	95	100	100	139	23	25	141
7	7	100	100	80	80	131	15	17	133
9	9	140	140	60	60	123	7	10	126
12	12	1000	1000	25	25	116	0	0	116
4.3	4.7	72	83	145	156	148	37	30	155
4	5	65	90	135	166	140	45	25	160
3	6	50	105	120	181	130	55	17	168
2	7	40	115	110	191	120	65	10	175
0	9	30	125	90	211	105	80	0	185

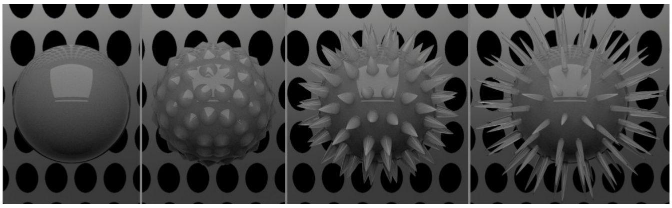

Fig. 3: We added spikes of different length and thickness to a sphere to generate three additional shapes. The spikes are geometrically thin and might contain significant information for assessing translucency differences. The material of the illustrated objects corresponds to that of CP 2.

Fig. 4: A see-through anchor pair (left) was used in Experiment 1, while Experiment 2 was based on the one that does not permit to see through (right).

between a CP and a test sample. Afterwards, we fit a Probit binomial model for each direction to estimate the Euclidean distance to the T50 point in the given direction. The distance to the T50 point a.k.a. the point of equal opportunity, is the distance between the sample and CPs at which 50% of the observers consider the test pair difference smaller than that of the anchor pair, while the other 50% consider it larger. In this case, we consider the difference in the test pair to be equal to suprathreshold translucency difference. The fitting was conducted using MATLAB's Probit link function in generalized linear regression (*glmfit()*).

The Probit model is suitable as the dependent variable is binary ("different from the anchor pair" or "not different from the anchor pair"). Although logistic regression could also have been used in this scenario, we opted for the Probit analysis in order to make our results comparable with that of Urban *et al.* [36]. The predictor variable is the Euclidean distance in the absorption-scattering space producing the estimate of the success rate. Success rate is the frequency of the test pair difference being considered larger than that of the anchor pair. We then invert the problem to get the Euclidean distance corresponding to 0.50 success rate. An example of the fitted curve is illustrated in Figure 5.

Afterwards, χ^2 goodness-of-fit test was conducted with α set to 0.05. Identification of the T50 point with high confidence has not always turned out possible. In some cases, the estimations failed a goodness-of-fit test (α>0.05), while

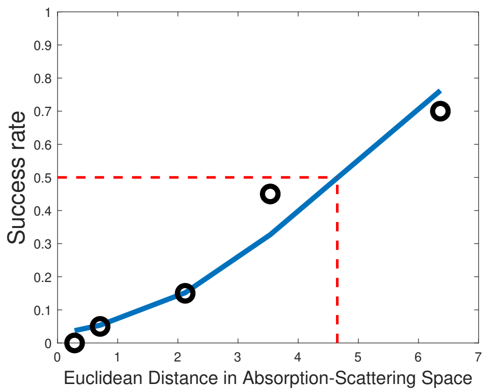

Fig. 5: The Probit fitting on the example of Buddha shape, CP1, the diagonal direction of increasing absorption and decreasing scattering. The black circles mark the test samples.

in some other cases, the estimated position was physically implausible (e.g. negative absorption or scattering). Finally, we analyzed how these distances to T50 points vary among different shapes, CPs and directions.

3.6.2 Quantifying Shape

Rank order analysis of the T50 distances among different shapes only enables qualitative assessment of the impact. We used a shape descriptor to quantify the geometrical characteristics of a shape and correlate it numerically with the T50 distances. We calculated surface-to-medial-axis distances [5], [30] and used its histogram statistics to characterize the shape. The medial axis is the topological skeleton of a shape. Surface points in thick and thin parts of the object are, respectively, further and closer to the medial axis.

4 Results

4.1 Results for each Control Point (CP)

The results for all individual CPs and types of anchor are illustrated in Figure 6.

For CP1, which is optically thin and see-through, the T50 point is determined in the vast majority of the cases. When non-see-through anchor pair was used, the T50 point was determined for all objects in all directions. The distances are generally short and they are shorter when these see-through test pairs are compared with a non-see-through anchor pair. The distance is usually largest for a spherical shape, while no clear difference is visible among other shapes.

For CP2, the distances are usually larger for a spherical shape, while no clear trend emerged for other shapes. Interestingly, T50 point is never determined for the more opaque direction ($\sigma_a = \sigma_s$) and the difference in both anchor pairs was usually judged larger than the difference between a CP and a test sample with higher absorption and scattering.

For CP3, the estimated T50 point coordinates for a spherical object have been negative, thus, physically implausible when the experiment was conducted on a see-through anchor pair. Also, T50 is never reached when scattering decreases and absorption goes up. This can be explained with the fact that observers considered all test samples already opaque and increasing absorption, simply affected their lightness, not their translucency cues. Interestingly, this trend changes and T50 points are reached more often when a non-see-through anchor pair is used.

For CP4, the distances to T50 points are rather large and generally consistent among all shapes. For a sphere with thin long spikes, T50 was determined only when non-see-through anchor-pair was used. Further observation is that decrease in absorption makes larger impact than decrease in scattering. This could be rooted in the fact that the magnitude of absorption of a CP is considerably smaller than the magnitude of its scattering.

For CP5, similarly to other CPs, T50 distances have been shorter and easier to determine when samples were judged against a non-see-through anchor pair. On many occasions, a scattering estimate has been negative, which can be attributed to the CP's proximity to the scattering axis. In both experiments, pairs of objects with thin parts have higher magnitude of perceived translucency difference than pairs of spherical objects. For this CP, the object with the medium-sized spikes has been the best one to detect translucency differences on.

4.2 General trends

On many occasions, T50 distances have been neither reached, nor estimated with high confidence. Refer to Table 2 - if the T50 point was reached for a given CP in a given direction, a respective cell is marked green, otherwise, it is marked red. The table shows that T50 point was not reached 38 times (number of red cells in the respective half of the table) out of possible 100 (5 shapes $\times$ 5 CPs $\times$ 4 directions) when see-through anchor pair was used and 28 times when the experiment was based on non-see-through anchor pair. As the test pairs have been identical in both experiments, this is an indication that observers are more sensitive to see-through cues of the transparent anchor pair when comparing it to non-see-through test materials. Additionally, the T50 point was reached for optically thin CP1 in all but one occasion in both experiments.

Most frequently, the T50 point is not reached for the spherical object and for the one with long thin spikes, while it is mostly reached for Buddha and the sphere with medium-sized spikes. A compact spherical object lacks thin parts and respective translucency cues, while Buddha per contra possesses finer details, i.e. broader range of translucency cues (as shown by [38]). These two observations are consistent with our hypotheses. However, the results for a sphere with the thinnest spikes is largely counter-intuitive. While the luminance gradient characteristic for translucency is visible on wider and thicker medium-sized spikes, it might be harder to detect on thinner spikes due to contrast sensitivity limitations - they simply occupy a smaller portion of the field of view (e.g. ref. to Figure 3). A second explanation for this result can be found in the remarks made by the observers - some of them noted that the spikes look so different from the spherical core that they thought they were made of different materials and decided to assess just the major body of the object. While these hypotheses deserve

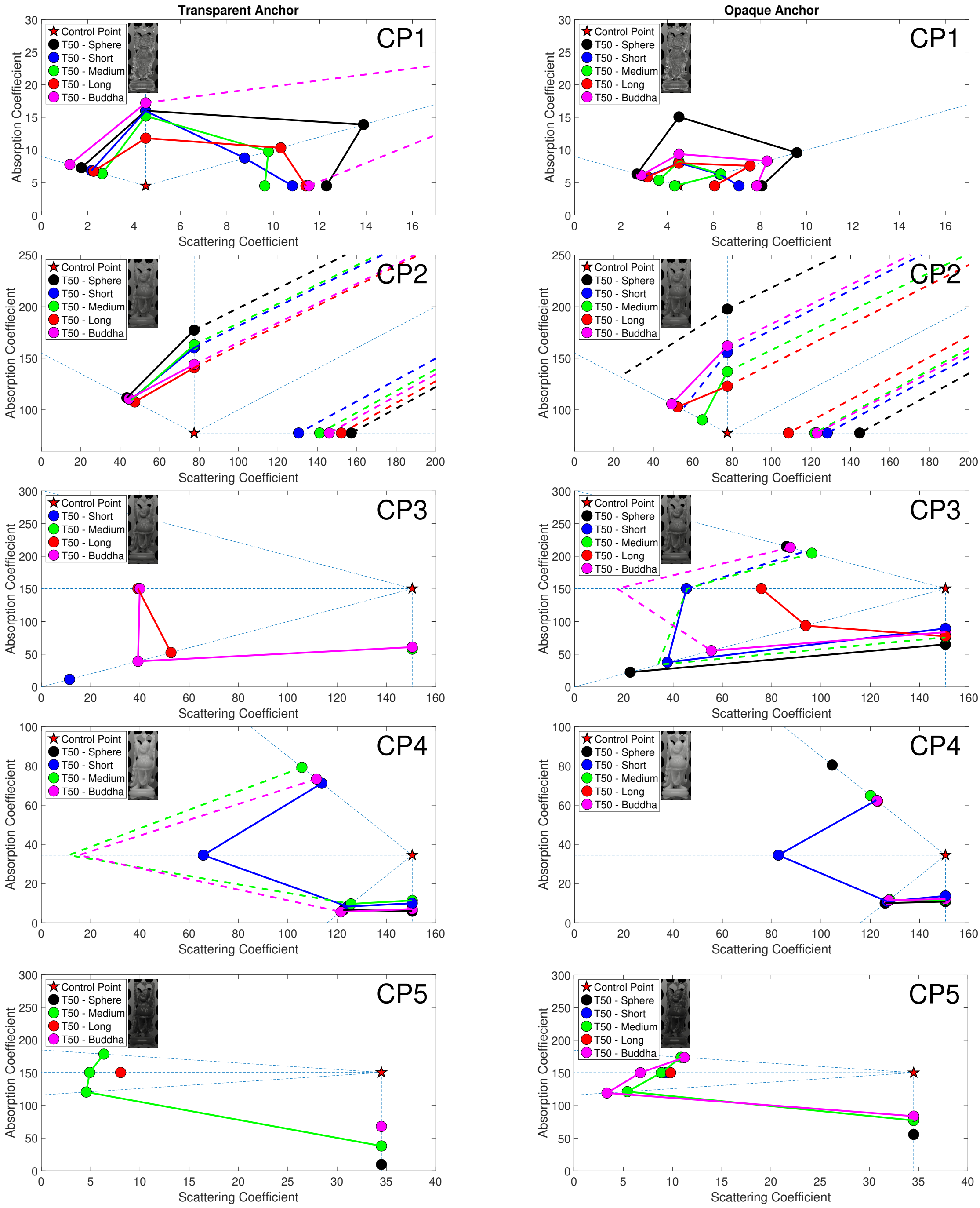

Fig. 6: The location of the T50 points for each CP and type of anchor (left column - transparent; right column - opaque). Only the T50 points that are physically plausible and passed goodness-of-fit test are shown and connected with solid lines. If the estimate failed a goodness-of-fit-test but it is within a plausible range, dashed-lines are directed towards its potential position. The Buddhas shown next to the legend illustrate the CP material.

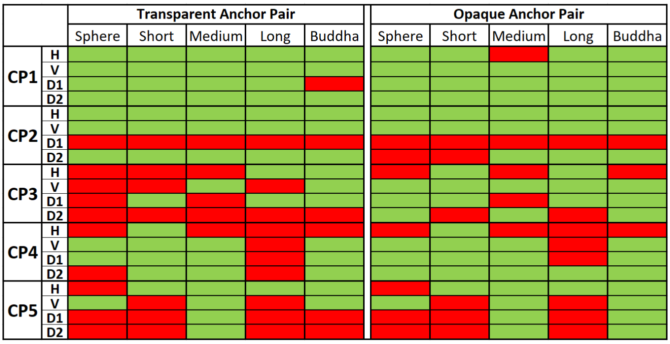

TABLE 2: Green cells mean that T50 point was reached and determined with high confidence for a given shape, CP and direction, red cells mean that it was not. The rows marked *H* and *V* correspond to one-dimensional change in horizontal (scattering) and vertical (absorption) directions, respectively; *D1* and *D2* correspond to the 2-dimensional change. The left and right halves of the diagram illustrate results from *Experiment 1* and *2*, respectively. The anchor pair with shine-through cues is referred to as *transparent* and its non-see-through counterpart as *opaque*, for simplicity's sake. The abundance of the red cells is noteworthy for spherical and long-spike shapes. Interestingly, for CP2, T50 point was never reached towards the optically thick direction (D1).

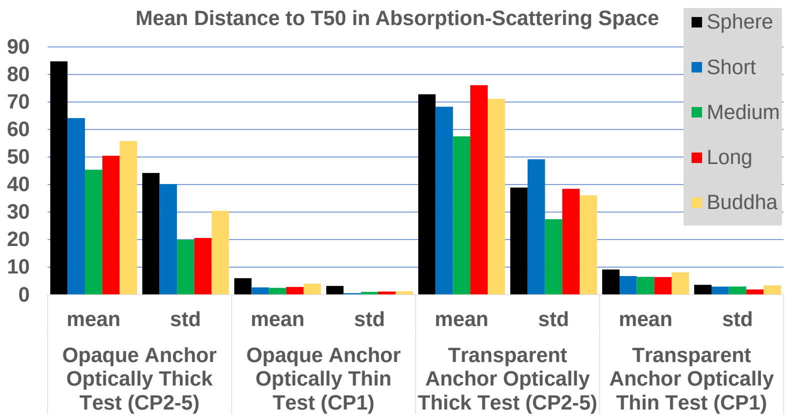

Fig. 7: Variation of the mean distance to T50 points in the absorption-scattering physical parameter space across different sets of comparisons. Only physically plausible estimates which had passed a goodness-of-fit test have been considered. The anchor pair with see-through cues is referred to as *transparent* and its non-see-through counterpart as *opaque*. It is apparent that the distance is shorter when optically thin test pairs are assessed. Also, using see-through anchor pair increases both the mean distance to the T50 point, as well as the standard deviation. For spherical objects the mean distance is usually larger than that of Buddhas and spheres with short- and medium-sized spikes.

further study, the relation between structural thickness and sensitivity to translucency differences is evidently neither straightforward, nor monotonic.

Figure 7 shows the mean distance needed to reach the T50 point for different shapes and materials. We can observe that the distance is considerably shorter for optically thin test materials and also when non-see-through anchor pair was used, being consistent with our hypotheses. Interestingly, the mean distance to the T50 point is larger for a sphere than it is for Buddha and spiky objects with short- and medium-sized spikes (although not with long spikes). Figure 8 illustrates average *Alpha* distances to T50 points. We can see that $\Delta Alpha$ largely accounts for the sensitivity difference between optically thin and optically thick materials. However, shape-specific adjustments are needed.

4.3 Histogram of surface-to-medial-axis distances

The histograms are shown in Figure 9. The summary statistics of the histograms and how they correlate with the mean T50-distances are given in Tables 3-4, respectively. Mean and median distance to the medial axis, as well as the percentiles, are significantly correlated with the T50 distances, except for the case, when the optically thin anchor pair is compared with the optically thick test pairs. However, standard deviation, skewness and kurtosis turned out poor correlates of the psychophysical data. The medial axis of a

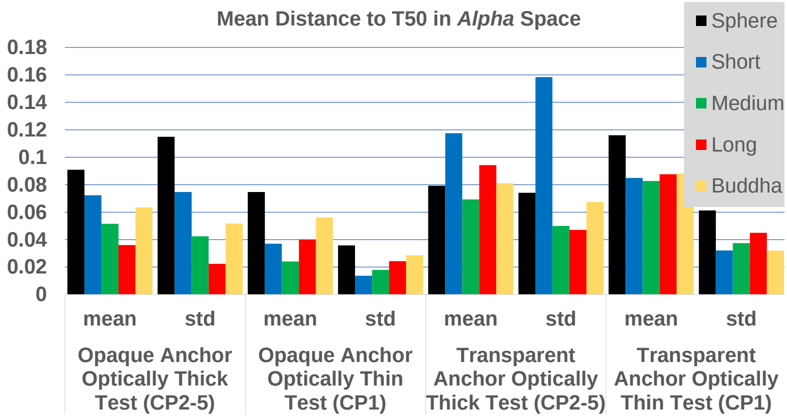

Fig. 8: *Alpha* distance to the T50 point. The figure illustrates that $\Delta Alpha$ largely accounts for inconsistency between optically thin and thick test samples which was apparent in absorption-scattering space. However, shape-dependent differences remain.

TABLE 3: The summary statistics of the surface-to-medial-axis histogram. The values are given in millimeters.

	Sphere	Short	Medium	Long	Buddha
Mean	50	7.37	3.88	2.65	4.84
Median	50	7.14	3.44	1.53	3.95
75th Percentile	50	8.67	5.29	4.20	7.14
99th Percentile	50	11.22	8.29	6.88	15.43
Std	0	1.77	1.88	1.86	3.52
Skewness	N/A	0.01	0.56	0.71	1.03
Kurtosis	N/A	3.12	2.51	2.12	3.71
Component ratio	N/A	0.19	0.42	0.81	0.84
Component mean ratio	N/A	0.72	0.45	0.26	0.30
Component mean difference	N/A	2.65	3.44	3.24	4.91

perfect sphere is its center and the distance to it is equal to its radius for all surface voxels. This makes it challenging to compare a perfect sphere with other shapes in terms of histogram statistics. The data for spheres has been only considered when a given metric was applicable to spheres.

While these metrics assume a single Gaussian distribution, there are in fact two distributions of thicknesses in spiky objects - the core sphere with larger distances and spikes with shorter distances. We fitted a two component Gaussian mixture model to the histogram and report the ratio between component proportions, as well as the difference and ratio between the means of the two Gaussians. Interestingly, neither of the three correlated with the observer data. We want to highlight that the accidental similarity in the magnitudes of geometrical thickness and T50 distances can produce high Pearson correlation, while the rank order correlation can still be not significant (e.g. see the difference between the means of the two Gaussians when optically thin test and optically thick anchor pairs are compared).

5 Discussion and Analysis

The analysis of the experimental results revealed several interesting trends, which are consistent with our hypotheses

TABLE 4: The Pearson and Spearman rank-order correlation between the histogram statistics and the mean distance to the T50 point. The Spearman coefficient is given in the parentheses. If *p-value*<0.10, the respective estimate is given in a boldface. The first six statistics assume single Gaussian distribution, while the last three fit a two-component Gaussian mixture model to the data.

	Thick Anchor Thick Test	Thick Anchor Thin Test	Thin Anchor Thick Test	Thin Anchor Thin Test
Mean	**0.92 (0.9)**	**0.91** (0.5)	0.27 (-0.1)	**0.82 (0.9)**
Median	**0.92 (0.9)**	**0.90** (0.5)	0.26 (-0.1)	**0.82 (0.9)**
75th Percentile	**0.92 (0.9)**	**0.92** (0.5)	0.27 (-0.1)	**0.83 (0.9)**
99th Percentile	**0.92** (0.8)	**0.96** (0.7)	0.30 (0)	**0.90 (1)**
Std	-0.69 (-0.7)	-0.51 (-0.3)	-0.10 (-0.3)	-0.33 (-0.3)
Skewness	-0.51 (-0.2)	0.73 (0.8)	0.25 (0.6)	0.55 (0.2)
Kurtosis	0.58 (0.6)	0.74 (0.4)	0.02 (-0.2)	**0.90 (1)**
Component ratio	-0.37 (-0.2)	0.68 (0.8)	0.58 (0.6)	0.42 (0.2)
Component mean ratio	0.62 (0.4)	-0.49 (-0.6)	-0.38 (-0.8)	-0.23 (0.4)
Component mean difference	-0.19 (-0.4)	**0.90** (0.4)	0.11 (0)	0.85 (0.4)

as well as with the state-of-the-art in translucency perception research. The most significant observations we have made are as follows:

- Presence of geometrically thin parts facilitate detection of translucency differences for some materials, but not for others. The correlation between structural thickness and sensitivity to translucency differences is possible to be characterized qualitatively, but quantitative modelling remains beyond reach.
- Human observers tend to be more sensitive towards changes in absorption and scattering properties when the object has see-through cues.
- In optically thick materials, increasing absorption

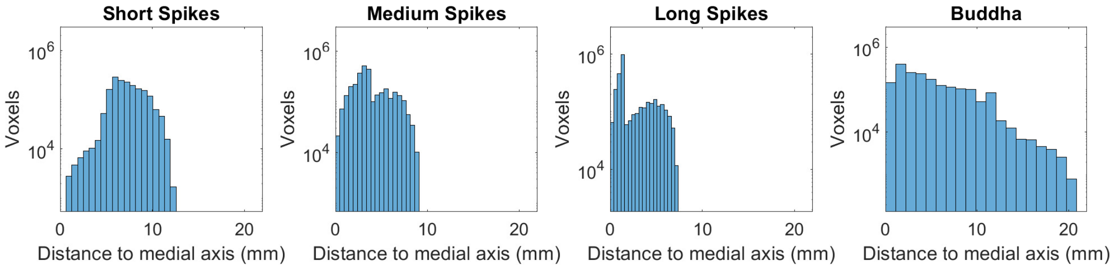

Fig. 9: The histogram of surface-to-medial-axis distances. As the physical dimensions of the objects vary, the histograms are not directly comparable in terms of the absolute number of voxels. However, the summary statistics of the histogram provide insight into the overall distribution of thin and thick parts. Thinner and longer spikes lead to a tail in the lower end of the histogram and produce stronger positive skew.

and scattering did not yield a suprathreshold translucency difference. It seems that we have not been able to increase the luminance contrast any further due to a diminishing return effect.

- Δ*Alpha*, the perceived translucency difference metric proposed by Urban *et al.* [36] needs to accommodate shape-specific adjustments.

5.1 Impact of Geometrical Thickness

The T50 distances that passed the goodness-of-fit test have oftentimes been larger for spherical and thicker objects than for those with thin parts. Moreover, on many occasions, the T50 distances have not been determined at all for a perfect sphere, while they had been found for other shapes.

We observed that mean and median surface-to-medial-axis distances have been better predictors of the experimental data than the measures quantifying asymmetry and distribution of thick and thin parts. However, the average thickness of the object does not adequately reflect presence or absence of the thin parts and can depend more on the scale of the object. Our observation is consistent with other research that postulates that the areas where a photon needs to travel shorter distance, such as edges [12], [19] and other fine details on the surface [28], [31], [38], contain the vital portion of the information about material translucence. Refer to Figure 10. Although the two optically thick materials are far away in the absorption-scattering space, the pixel-wise difference shows that they differ in thin parts, while remain relatively similar in thick areas.

While this trend is easier to characterize qualitatively, proper quantitative modeling of the impact of shape and geometry seems to range from very difficult to infeasible at this stage. As long as we do not know exactly which image regions the HVS relies on and how it weights the luminance information present in the image, we are not able to construct a shape descriptor that will correlate well with the perceived translucency differences. For instance, the thinnest spikes, against our expectations, turned out counter-productive in assessment of translucency differences. Moreover, a histogram of surface-to-medial-axis distances is viewpoint-blind and accounts for all spikes of the shape. However, we do not know, whether they are equally important [28], as for instance, the spikes located on the backside of the object might have a negligible role. However, this limitation is less important in real-life scenarios, as humans can interact with the objects and inspect them from all different geometries [16].

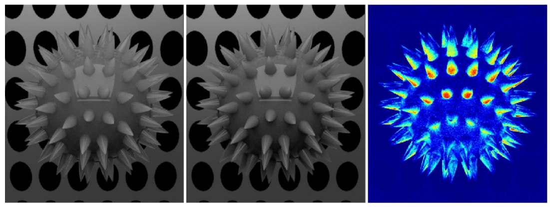

Fig. 10: σ_a and σ_s equal to 77.5 in the left image, and to 1000 in the middle one. The absolute pixel-wise differences between the two are shown in the rightmost pseudocolor image (blue - zero difference, red - maximum difference; scaled with a factor of 5 for visualization's sake). Regardless the considerable difference in optical properties, the thick parts remain relatively similar, while the difference is manifested in thin spiky regions.

Thin parts that are represented as bumps considerably affect the surface topography. They generate shadows and provide additional cues about the surface geometry of the object. Refer to Figure 11. Although the respective objects are made of the identical materials in both pairs, the difference becomes more apparent in the right pair, as the bumps of an optically thick material produce sharper shadows than its optically thinner counterpart. Similar image contrast resulting from a surface relief has already been shown by Xiao *et al.* [39] to be diagnostic for optical thickness of the material. Marlow *et al.* [25] have demonstrated that if surface and shading co-vary (as in the right image of the right pair in Figure 11), the material is perceived opaque, while if they do not, perception of translucency is evoked (left image in the right pair, Figure 11). However, a simple spherical shape with no concavities or bumps leaves less room for observing this kind of surface-shading co-variance and leaves more room for interpretation. This might have implication for material appearance research. If a sphere lacks translucency cues, it might be a bad idea to have it as a shape of choice in psychophysical experiments, being consistent with [20].

5.2 Impact of Optical Thickness

We have observed that when the magnitude of absorption and scattering is low, a smaller change is needed in absorption and scattering properties to notice a translucency difference. This observation is consistent with the findings by Urban *et al.* [36] and is accounted for by *Alpha* (see constant *Alpha* curves in Figure 3 of [36]). This fact is also consistent with Stevens' law [34] (similarly to [37]) and involves important implications for translucency perception research. Seemingly, the HVS is more sensitive to background contrast and background blur cues present in transparent materials [32] than it is to luminance contrast variation, which proposedly is a translucency cue for objects and materials that do not permit seeing through [12], [19], [23], [25], [28], [38]. See-through cues overtake and outweigh luminance contrast variation cues when both are present. This has the following implications:

- The perceptual mechanisms of assessing transparent and translucent materials might be fundamentally different. Therefore, they should be studied independently - not simply compared with one another. For instance, when the method of constant stimuli is used, the anchor pair should be selected with care and with full consideration of the corpus of the test pairs.
- This is important for material design both in 3D printing and computer graphics applications. In addition to the fact that this observation will affect material mixing ratios for perceived translucency matching, we hypothesize that the HVS is more sensitive to unintended artifacts when the material is optically thin.

The observers noted in the post-experiment interviews that they relied on background distortion in see-through images, while there was hardly any cue apart from lightness in non-see-through materials. Several observers noted that they always considered the difference in the see-through anchor pair to be larger, because *"the test pair was composed of two opaque materials, meaning that both objects in a pair had zero translucency and thus, were not different by translucency"*. To some the test pairs look like *"two fully opaque billiard balls with two different colors"*. This supports the proposal by Fleming and Bülthoff [12] that the HVS has poor ability to invert optics. It also seems that if the mean free path is very short, subsurface light transport is not noticeable at all. However, it remains an interesting open question why observers relied on lightness or brightness as a translucency cue. Lightness is natural to highly scattering media and it is characteristic to many translucent materials, such as snow and milk. Therefore, we cannot rule out that lightness itself is a cue for perceived translucency. Geometric information [25] can be important if the surface geometry is complex. For instance, observers might have compared lightness of the spikes, when they were present. However, for simple shapes, such as spheres, regional variation of the luminance distribution is minimal. This can explain why some observers use global mean luminance for assessing translucency, while others simply consider the material to be opaque.

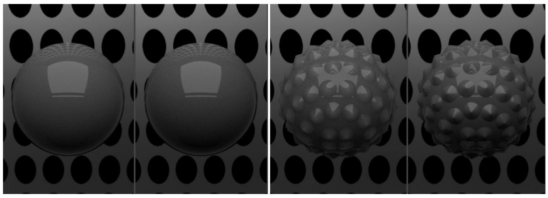

Fig. 11: σ_a and σ_s equal to 77.5 in the left images of the both pairs (CP2) and to 1000 in the right ones. Regardless the considerable distance in absorption-scattering space, spheres look nearly identical. The difference becomes more apparent for bumpy objects, as the bumps produce sharper shadows when the material is optically thicker.

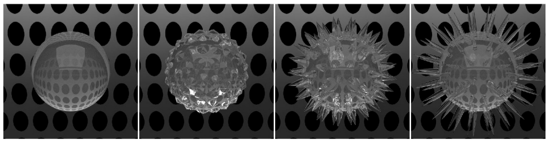

Fig. 12: The visibility of the background is strongly affected by object's shape. For instance, a bumpy object (second from the left) does not permit to see the background, even though it is made of a transparent material.

5.3 Interaction between Optical and Geometrical Thickness Effects

We hypothesize that there might be an interaction between geometrical thickness and optical thickness effects. This hypothesis is derived from the notion that background distortion is a strong cue for assessing translucency. However, visibility of the background does not only depend on the optical thickness of a material, but also on its surface geometry. For some object shapes, the background is not visible even if the material is optically very thin (see Figure 12).

5.4 Moving towards Optically Thick Direction

The special case of the lower visual sensitivity to absorption-scattering changes in optically thick materials is when both absorption and scattering are increased. In this direction the T50 point was never reached. As discussed above, luminance contrast is a cue used for distinguishing levels of translucency between the objects. While scattering and absorption coefficients are positively correlated with brightness and blackness, respectively [9], if both of them are increased equidistantly, there is a diminishing return effect in the resulting luminance contrast and the overall look remains roughly unchanged (refer to the left pair in Figure 11). This is consistent with the lightness reflectance measurements conducted by Urban *et al.* [36] (see Figure 3 in [36]). When the mean free path is short, a photon cannot traverse through a thick body of the material and the penetration depth into the volume is negligibly small. Therefore, shortening the mean free path further does not yield any perceptual difference. However, the mean free path becomes increasingly important as thinner and finer details are introduced in the shape.

5.5 Δ*Alpha* as a Measure of Perceived Translucency Difference

The visualization of the results in absorption-scattering and *Alpha* spaces shows that Δ*Alpha* accounts for sensitivity differences between optically thin and optically thick regions relatively well. On the other hand, Δ*Alpha* is not an adequate metric to reflect shape-specific effects on perceived translucency differences. The value of the ideal translucency difference metric should be identical for all shapes and all pairs of control-T50 points. *Alpha* is a shape-independent material property. However, the arrangement of *Alpha* in space depends on the shape. The psychometric function used in the definition of *Alpha* was measured on Buddha shapes [36]. Translucency cues that are present in the Buddhas might be absent from a sphere or other shapes - therefore, the psychometric function might be different. Δ*Alpha* should be adjusted so that it could accommodate shape-dependent effects on perceived translucency differences.

If we draw a parallel with colors, CIELAB is reasonable to quantify color for defined and fixed viewing conditions. However, for color differences various "parametric factors" need to be taken into consideration [21], [24] (such as distance between the patches, luminance level, presence of texture etc.). For instance, the CIEDE2000 color difference formula [24] can be fitted to parametric factors using *k-values*. We believe that a phenomenologically similar parametric factor for perceived translucency differences is object's shape, which Δ*Alpha* should be parametrized for. We hypothesize that the necessary features for parametrization can be extracted from shape descriptors in future work.

Interestingly, T50 distances in the *Alpha* space have not been identical even for Buddha shapes that *Alpha* was defined on. This can be an indication that besides parametric factors, inter-observer differences also exist.

5.6 Limitations

This study comes with multiple limitations that need to be addressed in the future. All experiments were conducted with the same scattering phase function, index of refraction, microfacet-level surface roughness and illumination geometry. These factors are known to be modulating perceived magnitude of translucency themselves [12], [13], [14], [20], [38] and we cannot rule out that our observations do not hold for other materials. We also assumed wavelength-independent absorption and scattering. Although it has been demonstrated earlier that translucency can be assessed without chromatic information [7], [12], in our daily lives we hardly ever encounter materials with wavelength-independent absorption and scattering; the HVS might not be properly trained to assess these kind of stimuli [7], [8].

Besides, our ability to model the correlation between sensitivity to perceived translucency differences and geometric shape are inherently limited by two factors: our knowledge on how the HVS selects the image regions to deduce translucency; and the availability of proper shape descriptors. For instance, we do not know whether the HVS relies on the information in all visible spikes, or if even a single spike is enough. Nagai *et al.* [28] have observed that different individuals rely on different image regions to assess translucency that might complicate definition of a robust shape descriptor for translucency differences.

Apart from that, we have used a Buddha shape as an anchor pair. Presence of thin parts in the anchor pair Buddhas might itself have increased sensitivity to anchor pair differences. Additionally, Gigilashvili *et al.* [16] observed that cross-shape translucency comparison is generally found challenging by observers. Hence, we believe that future studies should consider differently-shaped anchor pairs.

There might be some unintended noise in the data due to the interpretation of the object composition. For some materials the appearance of the sphere proper and its spikes is so different that a couple of observers thought that the spikes were made of a different material and thus, they decided to assess a core sphere only.

Finally, the remarks made by some observers about opacity of optically thick materials opens up a discussion where the conceptual boundary lies between *translucency* and *opacity* (see [17] for further discussion).

We believe future work should evolve towards three objectives: firstly, shape descriptors, which could correlate with detectability of perceived translucency differences, should be developed. We hypothesize that identification of image cues for perceived translucency could facilitate construction of such a descriptor. Subsequently, shape-related effects should be incorporated in translucency difference formulae. Finally, a perceptual translucency space should be constructed that could permit navigation in nearly perceptually uniform units instead of highly perceptually non-uniform absorption and scattering coefficients.

6 Conclusion

We have conducted psychophysical experiments to identify how object's shape and particularly, presence of thin parts, affect visual detection of perceived translucency differences. We have observed that presence of thin parts in many cases lead to easier detection of translucency differences. However, this impact is difficult to model quantitatively, as the limited knowledge about translucency perception mechanisms does not permit to design proper shape descriptors.

We also found that change in absorption and scattering make more visual impact if their absolute magnitude is smaller. This also reveals that see-through cues, when present, are emphasized over luminance contrast cues leading to the conclusion that perception of see-through transparent and non-see-through translucent materials should be studied independently. This has implications for material design in 3D printing and computer graphics, as well as for psychophysical methods when suprathreshold translucency difference is determined.

7 Acknowledgements

The work has been funded by the Measuring and Understanding Visual Appearance - MUVApp project of the Research Council of Norway (project #250293). The authors want to thank Tejas Tanksale for his contributions in the stimuli rendering process, Johann Reinhard for computing the distance between surface and medial-axis and all observers for their voluntary participation in the experiments.

References

[1] "ASTM E284-17 standard terminology of appearance." ASTM International, West Conshohocken, PA, 2017. [Online]. Available: https://doi.org/10.1520/E0284-17

[2] C. 175:2006, "A framework for the measurement of visual appearance," *International Commission on Illumination,*, p. 92 pages, 2006.

[3] J. Beck and R. Ivry, "On the role of figural organization perceptual transparency," *Perception & Psychophysics*, vol. 44, no. 6, pp. 585–594, 1988.

[4] R. S. Berns, D. H. Alman, L. Reniff, G. D. Snyder, and M. R. Balonon-Rosen, "Visual determination of suprathreshold color-difference tolerances using probit analysis," *Color Research & Application*, vol. 16, no. 5, pp. 297–316, 1991.

[5] H. Blum *et al.*, *A transformation for extracting new descriptors of shape*. MIT press Cambridge, 1967, vol. 4.

[6] A. Brunton, C. A. Arikan, T. M. Tanksale, and P. Urban, "3D printing spatially varying color and translucency," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 157:1–157:13, 2018.

[7] A. Chadwick, C. Heywood, H. Smithson, and R. Kentridge, "Translucence perception is not dependent on cortical areas critical for processing colour or texture," *Neuropsychologia*, vol. 128, pp. 209–214, 2019.

[8] A. C. Chadwick, G. Cox, H. E. Smithson, and R. W. Kentridge, "Beyond scattering and absorption: Perceptual unmixing of translucent liquids," *Journal of Vision*, vol. 18, no. 11:18, pp. 1–15, 2018.

[9] D. W. Cunningham, C. Wallraven, R. W. Fleming, and W. Straßer, "Perceptual reparameterization of material properties." in *Computational Aesthetics*, 2007, pp. 89–96.

[10] P. G. Engeldrum, *Psychometric scaling: a toolkit for imaging systems development*. Imcotek, 2000.

[11] C. Eugène, "Measurement of "total visual appearance": a CIE challenge of soft metrology," in *12th IMEKO TC1 TC7 Joint Symposium on Man, Science Measurement*, 2008, pp. 61–65.

[12] R. W. Fleming and H. H. Bülthoff, "Low-level image cues in the perception of translucent materials," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 3, pp. 346–382, 2005.

[13] R. W. Fleming, F. Jäkel, and L. T. Maloney, "Visual perception of thick transparent materials," *Psychological Science*, vol. 22, no. 6, pp. 812–820, 2011.

[14] D. Gigilashvili, L. Dubouchet, M. Pedersen, and J. Y. Hardeberg, "Caustics and translucency perception," in *Material Appearance 2020, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 2020, pp. 033:1–033:6.

[15] D. Gigilashvili, F. Mirjalili, and J. Y. Hardeberg, "Illuminance impacts opacity perception of textile materials," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2019, pp. 126–131.

[16] D. Gigilashvili, J.-B. Thomas, J. Y. Hardeberg, and M. Pedersen, "Behavioral investigation of visual appearance assessment," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2018, pp. 294–299.

[17] D. Gigilashvili, J. B. Thomas, J. Y. Hardeberg, and M. Pedersen, "On the nature of perceptual translucency," in *8th Annual Workshop on Material Appearance Modeling (MAM2020)*. Eurographics Digital Library, 2020.

[18] D. Gigilashvili, P. Urban, J.-B. Thomas, J. Y. Hardeberg, and M. Pedersen, "Impact of shape on apparent translucency differences," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2019, pp. 132–137.

[19] I. Gkioulekas, B. Walter, E. H. Adelson, K. Bala, and T. Zickler, "On the appearance of translucent edges," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5528–5536.

[20] I. Gkioulekas, B. Xiao, S. Zhao, E. H. Adelson, T. Zickler, and K. Bala, "Understanding the role of phase function in translucent appearance," *ACM Transactions on graphics (TOG)*, vol. 32, no. 5, pp. 1–19, 2013.

[21] S.-S. Guan and M. R. Luo, "Investigation of parametric effects using small colour differences," *Color Research & Application*, vol. 24, no. 5, pp. 331–343, 1999.

[22] W. Jakob, "Mitsuba renderer," 2010, http://www.mitsuba-renderer.org.

[23] J. J. Koenderink and A. J. van Doorn, "Shading in the case of translucent objects," in *Human Vision and Electronic Imaging VI*, vol. 4299. International Society for Optics and Photonics, 2001, pp. 312–320.

[24] M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," *Color Research & Application*, vol. 26, no. 5, pp. 340–350, 2001.

[25] P. J. Marlow, J. Kim, and B. L. Anderson, "Perception and misperception of surface opacity," *Proceedings of the National Academy of Sciences*, vol. 114, no. 52, pp. 13 840–13 845, 2017.

[26] F. Metelli, "The perception of transparency," *Scientific American*, vol. 230, no. 4, pp. 90–99, 1974.

[27] I. Motoyoshi, "Highlight–shading relationship as a cue for the perception of translucent and transparent materials," *Journal of Vision*, vol. 10, no. 9:6, pp. 1–11, 2010.

[28] T. Nagai, Y. Ono, Y. Tani, K. Koida, M. Kitazaki, and S. Nakauchi, "Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study," *i-Perception*, vol. 4, no. 6, pp. 407–428, 2013.

[29] Z. Pizlo, "Perception viewed as an inverse problem," *Vision Research*, vol. 41, no. 24, pp. 3145–3161, 2001.

[30] D. Rebain, B. Angles, J. Valentin, N. Vining, J. Peethambaran, S. Izadi, and A. Tagliasacchi, "LSMAT least squares medial axis transform," in *Computer Graphics Forum*. Wiley Online Library, 2019.

[31] M. Sawayama, Y. Dobashi, M. Okabe, K. Hosokawa, T. Koumura, T. Saarela, M. Olkkonen, and S. Nishida, "Visual discrimination of optical material properties: a large-scale study," *BioRxiv*, p. 35 pages, 2019.

[32] M. Singh and B. L. Anderson, "Perceptual assignment of opacity to translucent surfaces: The role of image blur," *Perception*, vol. 31, no. 5, pp. 531–552, 2002.

[33] "The Stanford 3D Scanning Repository," Stanford University Computer Graphics Laboratory, 1994, http://graphics.stanford.edu/data/3Dscanrep/.

[34] S. S. Stevens, "The psychophysics of sensory function," *American Scientist*, vol. 48, no. 2, pp. 226–253, 1960.

[35] P. Urban, M. Fedutina, and I. Lissner, "Analyzing small suprathreshold differences of lcd-generated colors," *JOSA A*, vol. 28, no. 7, pp. 1500–1512, 2011.

[36] P. Urban, T. M. Tanksale, A. Brunton, B. M. Vu, and S. Nakauchi, "Redefining A in RGBA: Towards a standard for graphical 3D printing," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 3, pp. 1–14, 2019.

[37] B. M. Vu, P. Urban, T. M. Tanksale, and S. Nakauchi, "Visual perception of 3D printed translucent objects," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2016, pp. 94–99.

[38] B. Xiao, B. Walter, I. Gkioulekas, T. Zickler, E. Adelson, and K. Bala, "Looking against the light: How perception of translucency depends on lighting direction," *Journal of Vision*, vol. 14, no. 3:17, pp. 1–22, 2014.

[39] B. Xiao, S. Zhao, I. Gkioulekas, W. Bi, and K. Bala, "Effect of geometric sharpness on translucent material perception," *Journal of Vision*, vol. 20:7, no. 10, pp. 1–17, 2020.

Article F

Davit Gigilashvili, Lucas Dubouchet, Marius Pedersen, and Jon Yngve Hardeberg (2020). “Caustics and Translucency Perception.” In: *Material Appearance 2020, IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 033:1–033:6

https://doi.org/10.2352/ISSN.2470-1173.2020.5.MAAP-033

Caustics and Translucency Perception

Davit Gigilashvili, Lucas Dubouchet, Jon Yngve Hardeberg, Marius Pedersen;
Department of Computer Science, Norwegian University of Science and Technology; Gjøvik, Norway;

Abstract

Caustics projected onto the surface carry very interesting information regarding the material they are cast by. It has been observed in previous studies that caustics could be a widely used cue for translucency assessment by human subjects. We hypothesize that changing the reflectance properties of the surface an object is placed on, and removal of the caustic pattern might impact perceived translucency of the material. We conducted psychophysical experiments to investigate the correlation among caustics, environment colors and translucency perception, and found very interesting indications that materials appear less translucent under the conditions where caustics are absent.

Introduction

According to Lynch [1], caustic is "three dimensional envelope of imperfectly focused rays" or "two-dimensional pattern formed when a caustic falls on a surface." According to Wand and Straßer, *"caustics occur if light is reflected (or refracted) at one or more specular surfaces, focused into ray bundles of a certain structure, and then received as patterns of light on a diffuse surface."* [2] As many translucent objects cast caustic patterns onto other surfaces, and particularly, onto the surfaces they are located on, we encounter this phenomenon on a daily basis - a glass of water projecting caustic pattern onto the table can be one of the simplest examples among many.

It has been identified in the previous study [3] that caustics could be a significant cue for assessment of material subsurface light transport properties. The observers were asked to order objects from the *Plastique* [4] artwork collection. While many observers used translucency as a primary attribute for ordering, the caustic pattern cast through the object onto the white paper was widely used to assess translucency of the material. This phenomenon is illustrated in Figure 1, where caustic is visible under translucent objects, while it is missing around the opaque one.

In some cases, caustics can be the only cue for translucency perception. For instance, refer to Figure 2. While various cues provide information regarding light transmission properties of the spheres, caustics below sphere E is the only indicator that the object is not opaque. Moreover, indications have been found in [5] that as the human visual system has proposedly limited ability to invert optics [6], and as many caustic patterns have high luminance similar to specularities, internal and external caustics and the glittering effect of the caustic highlights might be mistaken for specular highlights and thus, increase perceived glossiness. This phenomenon is illustrated in Figure 3.

Little is known about the mechanisms of translucency perception, and factors contributing to that. Fleming and Bülthoff [6] proposed that translucency perception is a result of interpretation of simple image cues without inverting the underlying optics. Gkioulekas *et al.* [9] studied the role of phase function in translucent appearance, while Xiao *et al.* [10] extended the study to interactions among phase function, illumination directionality, and apparent translucency.

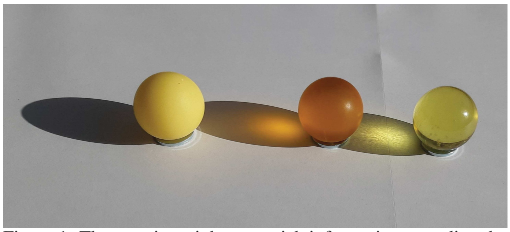

Figure 1: The caustics might carry rich information regarding the material properties. Even without looking at the objects themselves, just by looking at the shadow and caustic pattern, we can deduce the color of the object, as well as some information about its light transmission properties. Illustration taken from [7].

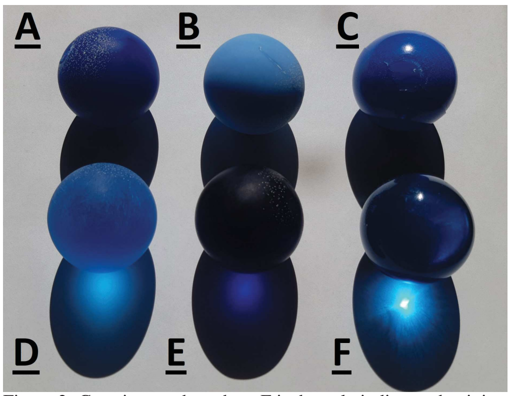

Figure 2: Caustics cast by sphere E is the only indicator that it is a translucent and non-opaque material. Illustration taken from [8].

Rendering caustics in computer graphics is a computationally costly process. Although caustics might have negligible role in some contexts, they play an important role in photorealism of some scenes [2, 11]. Kán and Kaufmann [12] have shown that caustics increase perception of realism in augmented reality, although it does not have the vital importance. However, Papadopoulos and Papaioannou [13] illustrate that caustics play a very significant role in realistic appearance of underwater scenes.

While, on the one hand, some studies highlight importance of caustics in realistic appearance, and on the other hand, the studies about translucency perception focus on the translucent object itself, to the best of our knowledge, no work has been done up-to date to investigate the importance of the external caustics as a

Figure 3: The caustics might contribute to glossiness perception. While the surface properties of all nine objects are identical, many subjects consider translucent ones more glossy, as caustics and back-reflected light are either mistaken for specular highlights, or increase total luminance and "shininess" of the object.

cue for translucency perception. We have conducted a study to identify whether presence of caustics and the reflection properties of the surface they are projected onto play any role in perceived translucency. The study revealed interesting trends that definitely deserve further follow-up in the future.

The paper is organized as follows: in the next section we present the experimental setup and stimuli generation process. Afterwards, the results are presented and discussed. Finally, we outline the directions for the future work.

Experimental Setup

We conducted psychophysical experiments in order to observe whether presence or absence of caustics could impact perceived translucency of a given material.

Stimuli

We rendered 30 images using Mitsuba Physically-based renderer [14]. We used bidirectional path tracer to render glass objects in 5 different shapes: sphere, cube, Stanford bunny, elephant, and wineglass - all of them placed in the Cornell box. Each object was rendered with 6 different degrees of transparency-translucency. While intrinsic material properties remained the same, translucency was manipulated using the *alpha* parameter, which "specifies the roughness of the unresolved surface micro-geometry using the root mean square (RMS) slope of the micro-facets" [14]. In other words, we manipulate light transmission properties by changing surface scattering, while volume scattering properties remain the same. The material property was loaded from the .mtl material library. For each of the color channels, ambient component was set to 0, while diffuse and specular components were set to 0.6 and 0.9, respectively. The refractive index was set to 1.5. The *alpha* values were equidistantly sampled between 0 and 1.

The shapes are illustrated in Figure 4. The impact of the *alpha* value on the material appearance of the object, is illustrated in Figure 5. It is worth mentioning that while smooth surfaces look more transparent, rough surfaces start looking translucent never reaching full opacity (although opacity does not necessarily imply complete absence of transmission as observed in [8, 7]) and even the roughest object has some degree of light transmission property. In this case, we expect relative judgement of translucency rather than an absolute one.

Afterwards we had to render identical objects but without caustics in order to compare perceived translucency between the two setups. We considered six different ways of removing caustics:

1. Using a rendering technique that does not produce caustics (with caustics "off"). However, the results would have been physically inaccurate.

2. Manually editing images in the graphics editor. This methodology will end in physically inaccurate and unrealistic results.

3. Rendering a fully opaque object of the identical shape, cropping the translucent object, and placing into the render in place of the opaque object. This result is also physically inaccurate and unrealistic.

4. Varying refractive index that is directly correlated with the caustics phenomenon. This is an interesting direction that we think of addressing in the future, but at this stage, we focused on single material property for all test samples avoiding an additional degree of freedom.

5. Occluding caustics with other objects. The methodology is promising, but considering that we had to accommodate occlusion of objects with various shapes and sizes, judgment of complex scenes that vary among images might have caused confusion among subjects, and also might be challenging to interpret due to the unintended side-effects occluding objects bring into the scene.

6. Making the floor most of the caustics are projected onto fully absorbing black. Although we cannot remove internal caustics this way, and some other cues are also impacted (e.g. lightness) in addition to caustics, the result is physically accurate, the scene structure remains the same (unlike the occlusion option), and understanding the impact from the surface color itself might have an application in the real world. Therefore, we opted for the latter approach.

Experimental Conditions

We hypothesize that introduction of the black floor makes objects look less translucent. The example of the effects of the floor color is shown in Figure 6. In order to test the hypothesis, we conducted an online user study (also referred to as "psychophysical experiment") using QuickEval [15] web-based tool. We used category judgement psychometric scaling protocol, where observers had to assign objects to one of the six categories varying from the most translucent to the least translucent, i.e. most opaquish one. To facilitate decision-making for the observers, we took two measures:

1. Placed an additional spherical object in all test images, in order to enable subjects judge material consistently across different shapes.

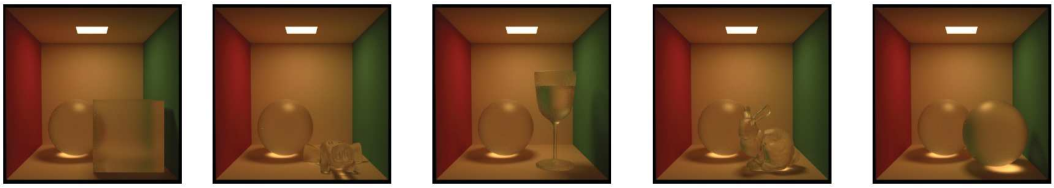

Figure 4: Five different shapes have been used in the study. The illustrated images are rendered with *alpha* value equal to 0.2.

Figure 5: The impact of the *alpha* value on the material appearance illustrated with the example of the elephant shape. *Alpha* is equal to 0, 0.2, 0.4, 0.6, and 0.8, from left to right, respectively.

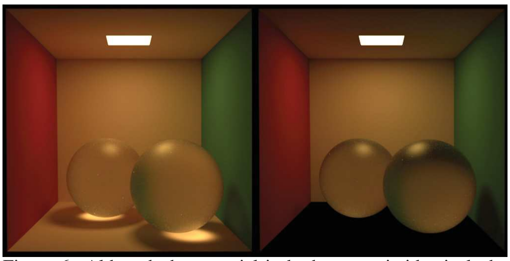

Figure 6: Although the material in both scenes is identical, the floor color changes its appearance.

2. Illustrated the reference spheres with the two extremes of *alpha* value under two different conditions, in order to facilitate scaling between the extremes. Having access to the extremes of the dataset ensures that observers make relative judgements, as they are not expected to perform absolute judgement based on a very small subset of the transparency-translucency-opacity spectrum. A sample scene from the experiment is shown in Figure 7.

In total 50 observers participated in the study. 13 observers were asked to rank objects by translucency, without providing a definition or interpretation of translucency. 37 observers were given more detailed instructions, as follows: *"Assess translucency of the material in the left image on 1-6 scale using a drop-down menu. 1 - most translucent, 6 - least translucent, i.e. closer to opacity. Sample materials of maximum (left column) and minimum (right column) translucency are illustrated on the right hand side of the panel."* However, both groups demonstrated identical trends, and thus, below we will only report aggregated results.

Results

If we assume that observers have used an equally spaced scale for their judgments, we can compare mean observer scores (category) for each shape and alpha value between the two setups. The mean observer scores and their 95% confidence intervals are shown on Figure 8. The lower category values correspond to more apparent transparency-translucency, while higher values correspond to more opacity. As we see in the figure, the mean observer-assigned category is larger (i.e. more opaque and less translucent) in the presence of the black floor for all shapes and all levels of alpha. For the vast majority of the cases, there is no overlap between the 95% confidence intervals that makes us conclude that the difference is significant and deserves further attention. The most apparent exceptions where confidence intervals overlap are smoothly-surfaced objects. This might be explained with the fact that these objects are transparent and highly specular (glossy), standing out from the rest of the stimuli, making match between the two setups easier for the observers. This might be an indication that floor color and caustics removal have larger impact on translucency perception rather than on transparency perception, further supporting our proposal in [16] that translucency and transparency cues differ significantly.

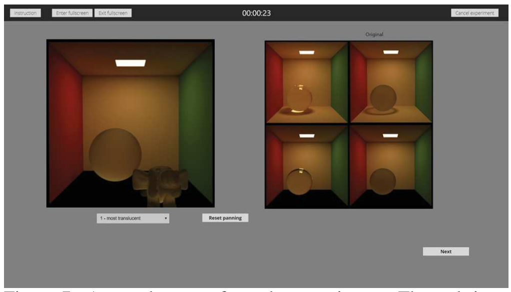

Figure 7: A sample scene from the experiment. The task is to select a category from the dropdown menu for the test material shown on the left side. The reference spheres with *alpha* equal to 0 and 1, are displayed for facilitating the judgement.

Considering the above-mentioned observation, it is likely that the assumption about an evenly spaced scale does not hold. Therefore, we applied Torgerson's categorical judgement model [17] (as cited in [18]), finding z-scores and corresponding scale

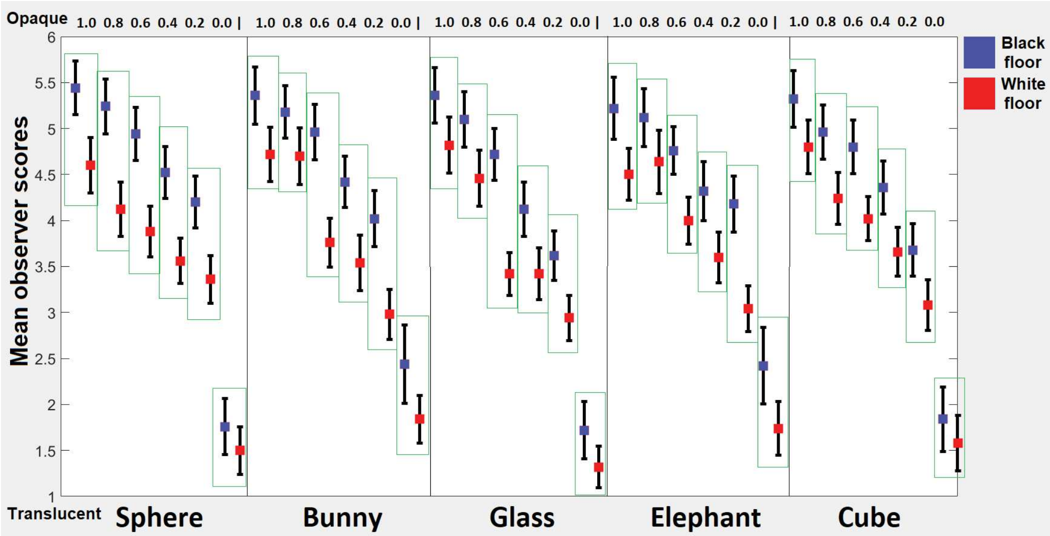

Figure 8: Vertical axis corresponds to mean observer scores, while the results are grouped by shape horizontally (bottom horizontal axis). The top horizontal axis corresponds to alpha values for a given shape. Squares correspond to mean observer scores for a given shape and alpha value. The blue squares signify materials shown on a black floor, while red squares correspond to materials shown on a regular Cornell box floor. The whiskers extend to the 95% confidence interval for mean observer scores. For clarity's sake, the results for each black-floor / white-floor pair of a given object are separated with a green rectangular frame from the results for other objects.

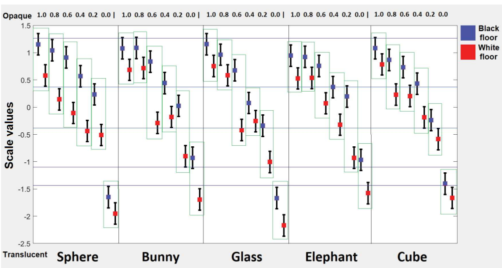

Figure 9: Vertical axis corresponds to scale values derived from Torgerson's categorical judgement model. Blue horizontal lines mark category boundaries. Note that equal variance is assumed for all samples.

values and category boundaries. The results are illustrated in Figure 9. The figure shows that the trend is identical to the one observed for mean observer scores in Figure 8. Interestingly, despite no overlap between the 95% confidence intervals, some materials fall in the same category both in case of white – and black-floor scenarios. This could indicate that 6 categories are not enough to adequately quantify translucency levels within this dataset and denser sampling of the potential categories is needed across the transparency-translucency-opacity spectrum.

Furthermore, it is interesting to figure out, whether the impact is identical for all shapes. For illustration's sake, the latter results are sorted by alpha value in Figure 10. Although the separation between two floor setups stands out, the confidence intervals overlap among all or most shapes for a given alpha value and floor color. Even if the shape could potentially impact the results, presence of the spherical object in all scenes might have compensated that effect. This should be considered in the future and the material should be shown only in one particular shape at a time. In some cases, e.g. Bunny with alpha=0.6, the impact of the floor change is very apparent. One of the explanations for this fact is the sequence the images were shown to the observers. When the material or shape is identical between two consecutive trials, the toggling effect impacts observer responses, and the assumption that all observations are independent does not hold anymore (refer to [19] about toggling and change blindness).

Finally, we plotted mean observer scores as a function of alpha surface roughness for each shape and floor color (Figures 11). As we can observe in the figure, the mean observer scores are always lower when floor is "white", i.e. caustics are visible. Perceived translucency decreases monotonously with the increase of alpha. The correlation between alpha and mean observer scores look linear and Pearson's linear correlation coefficient equals to 0.92 if all points are included, and increases up-to 0.98, if transparent smooth objects are excluded. The apparent drop in mean observer values when alpha equals to 0, is further indication that transparency and translucency judgments might differ by nature.

Discussion

We see clear indications that objects shown on the black surface look less translucent to human observers, even though the observers had a reference where they could observe appearance change between the two setups. The reference could potentially help them match the identical material between the two conditions, but we observe that the difference is significant even with

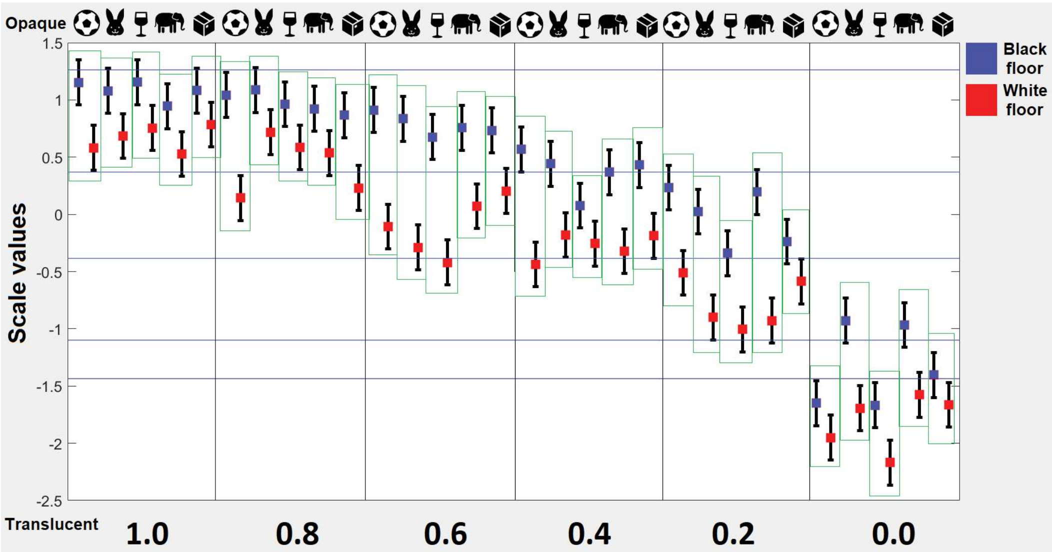

Figure 10: The results reported in Figure 9 but grouped by alpha values. The icons in the top horizontal axis correspond to spherical, Bunny, wineglass, elephant, and cube shapes, respectively. We observe no significant difference among shapes for a given alpha and floor.

this factor. It is worth mentioning that changing the floor color does not only remove caustics, but also affects other cues, like lightness - objects becoming darker, as no light is reflected back from the floor. While the cues used for translucency perception by the human visual system is not well understood, removal of caustics might not be the only explanation for the observed trends in the experiment. It is important to isolate caustics in future studies. However, the primary challenge is that either rendered images will be physically inaccurate that we never encounter in real lives, or physically accurate techniques to remove caustics from the scene will also affect cues other than caustics. As this work is the first step towards this direction, the proper trade-off needs to be found and implemented in the future for in depth analysis of the question. In addition to this, changing floor color removes just that portion of caustics which is projected on the floor, while caustics projected onto other surfaces, yet less apparent, still remain visible (e.g. refer to Figure 4, middle image – the caustics cast by the wineglass are visible on the green wall). Whether this cue was used by subjects remains unanswered within the scope of this study. Besides, we have observed the linear correlation between surface roughness and perceived translucency. The role of surface scattering in translucent appearance and its relation with subsurface scattering is an interesting question to be addressed in the future.

Finally, the study comes with one limitation that is also worth discussing. If we refer to Figure 5, the objects become darker as roughness increases. This can be intuitive at first glance due to facet shadowing and masking. However, model's failure to take interreflections between facets into account leads to energy loss, and might be contributing factor in rougher objects' dark appearance. This is a common problem microfacet-based models usually suffer from at some extent [17]. No single straightforward approach exists to solve this problem and several compensation techniques have been proposed to mitigate its effect [18, 19]. However, different techniques might lead to perceptually different results and assessment of their physical accuracy is beyond the scope of this paper. Besides, we want to highlight that we do not ask subjects for absolute translucency assessment. We show the extremes, the brightest and darkest objects, and ask observers to locate the test objects on a scale relative to these two. In this case, we assume that the impact of the energy loss might not have the critical importance. Although this question can be addressed in follow-up studies.

Conclusion and Future Work

We hypothesize that placing objects on a black floor that in itself leads to disappearance of the caustic pattern projected onto it, makes objects look less translucent. We have conducted psychophysical experiments to test this hypothesis. Considering above-presented results, we have clear indications that introduction of the black floor decreases perceived degree of translucency for a given material. However, whether this phenomenon can be attributed to absence of caustics only, or whether other cues affected by the floor color contributed to apparent translucency as well, needs further investigation.

While we discussed fully absorbing black floor and a binary case, between caustics and no caustics scenarios, floors with different reflection properties should be studied in the future in order to observe, whether sharpness or shininess of the caustics matter and at what extent. If we refer to Figure 1 again, we can see that sharpness and shininess of the caustics can give us a hint about translucency or transparency of the orange and yellow objects. For this purpose, real as well as synthetic stimuli can be used, where the floor will be colored in different levels of gray and have

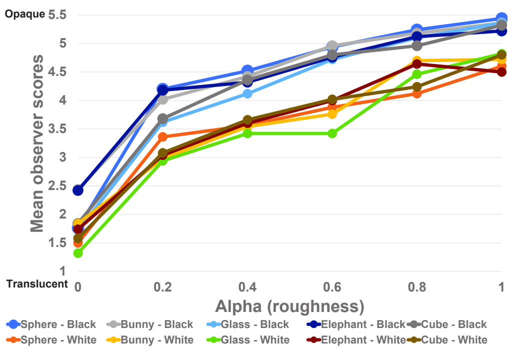

Figure 11: Mean observer score as a function of alpha with visible linear correlation between the two. Visible caustics always lead to lower mean scores. The line is fit for clarity's sake.

different roughness as well. Furthermore, we believe that the impact of shape deserves further attention in the future, in order to understand how shape impacts perceived translucency and to isolate what is the role of caustics cast by a particular shape.

Finally, considering the richness of the information embedded in caustics, we believe caustics could facilitate measurements. In case the straightforward correlation between material properties, shape and 2-dimensional caustic pattern is found, caustics can be used in image-based measurements of material properties. On the other hand, the potential existence of "caustic metamers", i.e. two different materials producing identical caustics, might limit the method. To the best of our knowledge, this methodology has not been studied yet and its limits are yet to be understood.

Acknowledgements

The work has been funded by the MUVApp (Measuring and Understanding Visual Appearance) project of the Research Council of Norway (project number 250293).

References

[1] David K Lynch, William Charles Livingston, and William Livingston, *Color and light in nature*, Cambridge University Press, 2001.

[2] Michael Wand and Wolfgang Straßer, "Real-time caustics," in *Computer Graphics Forum*. Wiley Online Library, 2003, vol. 22, pp. 611–620.

[3] Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg, "Material appearance: Ordering and clustering," in *2019 IS&T International Symposium on Electronic Imaging*. Society for Imaging Science and Technology, 2019, pp. 202:1–202:6.

[4] Jean-Baptiste Thomas, Aurore Deniel, and Jon Y Hardeberg, "The plastique collection: A set of resin objects for material appearance research," *XIV Conferenza del Colore, Florence, Italy*, p. 12 pages, 2018.

[5] Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg, "Perceived glossiness: Beyond surface properties," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2019, pp. 37–42.

[6] Roland W Fleming and Heinrich H Bülthoff, "Low-level image cues in the perception of translucent materials," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 3, pp. 346–382, 2005.

[7] Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen, "Behavioral investigation of visual appearance assessment," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2018, pp. 294–299.

[8] Davit Gigilashvili, Fereshteh Mirjalili, and Jon Yngve Hardeberg, "Illuminance impacts opacity perception of textile materials," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2019, pp. 126–131.

[9] I. Gkioulekas, B. Xiao, S. Zhao, E. H. Adelson, T. Zickler, and K. Bala, "Understanding the role of phase function in translucent appearance," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 5:147, pp. 1–19, 2013.

[10] Bei Xiao, Bruce Walter, Ioannis Gkioulekas, Todd Zickler, Edward Adelson, and Kavita Bala, "Looking against the light: How perception of translucency depends on lighting direction," *Journal of Vision*, vol. 14, no. 3:17, pp. 1–22, 2014.

[11] Jong Seo Kim, Kang Soo You, and Hoon Sung Kwak, "Caustics effects with photo-realistic rendering on movie ('Cars')," in *5th ACIS International Conference on Software Engineering Research, Management & Applications (SERA 2007)*. IEEE, 2007, pp. 274–280.

[12] Peter Kán and Hannes Kaufmann, "High-quality reflections, refractions, and caustics in augmented reality and their contribution to visual coherence," in *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2012, pp. 99–108.

[13] Charilaos Papadopoulos and Georgios Papaioannou, "Realistic real-time underwater caustics and godrays," in *Proc. GraphiCon*. Citeseer, 2009, vol. 9, pp. 89–95.

[14] Wenzel Jakob, "Mitsuba renderer," 2010, http://www.mitsuba-renderer.org.

[15] Khai Van Ngo, Jehans Jr. Storvik, Christopher André Dokkeberg, Ivar Farup, and Marius Pedersen, "Quickeval: a web application for psychometric scaling experiments," in *Image Quality and System Performance XII*. International Society for Optics and Photonics, 2015, vol. 9396, p. 93960O.

[16] Davit Gigilashvili, Philipp Urban, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen, "Impact of shape on apparent translucency differences," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2019, number 1, pp. 132–137.

[17] Warren S Torgerson, *Theory and methods of scaling.*, Wiley, 1958.

[18] Ján Morovič, *Color gamut mapping*, John Wiley & Sons, 2008.

[19] Steven Le Moan and Marius Pedersen, "Evidence of change blindness in subjective image fidelity assessment," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3155–3159.

Article G

Davit Gigilashvili, Marius Pedersen, and Jon Yngve Hardeberg (2018). "Blurring impairs translucency perception." In: *Color and Imaging Conference*. Society for Imaging Science and Technology, pp. 377–382

Blurring Impairs Translucency Perception

Davit Gigilashvili, Marius Pedersen, Jon Yngve Hardeberg; Norwegian University of Science and Technology; Gjøvik, Norway

Abstract

Translucency and factors impacting its perception is not yet fully understood. Various studies have examined the correlation between physical material properties and perceived translucency. Furthermore, the concept of translucency constancy has been introduced. However, to the best of our knowledge, no study has been conducted to identify how image quality impacts perceived translucency. In this study, we address to one particular image quality attribute - blurriness. We quantified blur with objective image quality metric and conducted psychometric scaling experiments to identify how blurring impacts the perceived degree of translucency. The analysis of the results show some indications that blur impairs translucency perception.

Introduction & Background

Translucency is among the least studied appearance attributes [1]. No single agreed definition of translucency exists. According to Eugene [1], "translucency occurs between the extremes of complete transparency and complete opacity... If it is possible to see only a "blurred" image through the material (due to some diffusion effect), then it has a certain degree of transparency and we can speak about translucency"; while Gerbino [2], defines distinction between transparency and translucency as "transparent substances, unlike translucent ones, transmit light without diffusing it."

The most extensive survey of the image cues affecting translucency perception has been carried out by Fleming and Bülthoff [3]. They review a broad range of the factors affecting the perceived translucency, like specular highlights, color, object scale, image contrast and illumination direction.

Furthermore, the study by Xiao *et al.* [4] concludes that perceived degree of translucency depends strongly on the illumination direction, phase function used in rendering (i.e. a probability distribution over directions, which describes the angular distribution of scattered light), and object geometric properties. In the same paper, they introduce the concept of translucency constancy, i.e. an ability of human beings "to estimate translucency in a consistent way across different shapes and lighting conditions."

It has been illustrated that the background and the pattern seen through the translucent material can also have dramatic impact on translucency perception [5].

In computer vision, translucency perception is often considered within the broader problem of material identification [6, 7, 8, 9]. Even though image quality has been considered an important factor for identification tasks in other fields, e.g. biometrics [10, 11], references to image quality as one of the factors impacting material identification and object appearance, is limited. Motoyoshi [12] argues that blurring non-specular regions, while keeping the specular highlights intact, increases the perceived degree of translucency, but blurring the whole image is not mentioned in the paper.

An interesting study has been conducted by Sharan *et al.*[9], where the authors demonstrated that blurring impairs material categorization. They tried to study the role of surface properties, like color, texture and gloss, in material categorization. They created a database using images from Flickr image sharing website[1] and sorted the images into nine material categories. The authors conducted psychophysical experiments, where observers had to categorize the materials. Afterwards, they introduced different degradation in the images, which they believed removed or decreased the role of the different surface properties (e.g. they used grayscale images to remove the role of color), conducted another material categorization experiment with the degraded images and compared the categorization accuracy with that of the original experiment. In order to remove high spatial information and impair texture recognition, the authors blurred the images and demonstrated that blurring the images decreased categorization accuracy from original 91% to 75.5%.

Sharan *et al.*[9] did not explicitly refer to image quality, but the authors obviously degraded the quality of the images when they blurred them. Blurring the images impairs not only texture recognition, but also makes surface geometry, shadings and highlights more ambiguous, because the luminance histogram is shrunk and the high contrast areas get smoother, as demonstrated by Motoyoshi[12]. The fact that blurring the images decreases material categorization accuracy on the one hand, and degrades the cues that are demonstrated to be correlated with translucency perception [3, 4, 12], on the other hand, we found it interesting to investigate further, how image quality, in terms of blurriness, impacts perceived degree of translucency.

The key research question is the following:"can blur of the image impact perceived degree of translucency?"

However, dilemma was whether to blur the whole image, or just the object. Therefore, as mentioned above, two different experiments were held with different stimuli: one with the images with the whole scene and context, and another one with the translucent objects cropped and displayed on neutral gray backgrounds. Hence, another research question arouse:"do blurred objects seen in the blurred scene demonstrate higher degree of translucency constancy than blurred objects seen in isolation?"

There are two major points that motivated us study the correlation between blur and translucency perception: first of all, in broader perspective, we are interested how different translucency perception is between the people with impaired and normal vision. Secondly, we want to identify how image quality impacts the perception of appearance attributes - translucency, in this case, and whether there is any threshold, when the quality becomes not acceptable when addressing the images of the translucent objects. In contrast with full scene images, isolated object images look unnatural. However, in non-blurred versions of them, we still have

[1] https://www.flickr.com

enough cues to consider the objects translucent. In the broader perspective, we want to identify, what are those cues, when they vanish and when a translucent object becomes a non-translucent blob.

To the best of our knowledge, no study has been conducted to examine the impact the blurriness of the image has on perceived degree of translucency of the materials. The aim of the study is to identify, and if possible, quantify, the impact blurriness of the image has on perceived translucency. The subsequent chapters are organized as follows: in Research Methodology & Experimental Setup chapter, we will discuss the approach applied to the problem. Afterwards, we will illustrate and discuss the results in Results & Discussions and finally, draw the conclusions from the latter and define directions for the future work.

Research Methodology & Experimental Setup

Design of the Experiments

The psychometric experiments were conducted using QuickEval web-based tool[13] . The experiments were held in two parts: one for the whole-scene images, and another one for the isolated objects. As mentioned above, full-scene experiments included the original images and blurred versions of them, while in isolated-object experiments, the translucent objects were cropped from the original images (and blurred versions of them), and placed on the neutral gray background, in order to remove scene and contextual information. Both experiments were pairwise comparisons [14], where the observers were shown two images and were given the following instruction: "Select the object with higher degree of translucency, i.e. transmitting higher amount of light." Additional oral instructions given, if needed. The experiment was conducted in forced-choice regime, where the observer necessarily had to select either object of the pair. The same pair was displayed twice in a flipped order. No reference image was displayed separately. Three different versions of each image were used in each of the experiments: original, moderately blurred and highly blurred. This totals to nine images within each experiment. All the images were compared against each other - hence, considering that the pairs were shown in a flipped order as well, 72 comparisons and about 10 minutes were needed for each of the experiments.

The observers could recognize the objects shown in isolated object experiments on the gray background were simply cropped from the full-scene images that they had already seen in full-scene image experiments. In order to discard the effect of this issue, we used different triplets of the images for full-scene and isolated object experiments. All the observers completed the experiments in the following order: 1. Full Scene Image-based experiment. 2. Isolated Object Image-based experiment.

Stimuli

We used the Flickr Material Database created by Sharal *et al.*[9] As the focus of this research is translucency, we used the images from the single category "Glass". Six different images were selected in total - three for full-scene image experiments, and three for isolated object image experiments. All images were RGB color images, provided in JPEG format and with resolution of 512×384 pixels. In order to avoid confusion among observers, only images with a single translucent object were selected.

The images were randomly selected from the database (with the constraint of including just single translucent object). "Moderate" and "High" Gaussian blur was applied to each of the images, with standard deviation of 5 and 25 respectively (with default kernel size of MATLAB *imgaussfilt* function [15]). The examples of the blurred images are illustrated on Figures 1 - 6. We understand that the number of images is low. However, we focused on the number of observers, rather than the number of images, as according to Sharma [16] *"given the amount of time necessary to perform these experiments,it is often more desirable to have a larger number of observers."*

Display

The experiments were conducted in controlled conditions. The images were displayed on EIZO CG246 display, with 1920x1200 resolution and 59 Hz refresh rate. The display was calibrated according to the following parameters: Gamut: sRGB; Gamma: 2.20; Brightness: 80 cd/m^2; Black point: 0.19 cd/m^2; White Point: 6502K, with the following x,y coordinates: (0.3127,0.3293); Contrast Ratio: 412:1;

The experiment was held under dim ambient illumination. The illumination was 27 lux in front of the keyboard and the color temperature of the ambient illumination was 4450K. The distance to the screen was approximately 50 centimeters.

Observers

20 observers, 12 males and 8 females, with normal, or corrected-to-normal vision voluntarily participated in the experiment. Average age of the observers was 28.1 years. The observers had technical background, but were naïve to translucency studies.

Analysis of the Collected Data

Collected subjective evaluation data was analyzed in the following way: first of all, Z-scores and their 95% confidence intervals [14] were used to illustrate the responses of the observers. Furthermore, binomial sign tests were conducted to examine the significance of the difference between the observations [17, 18]. The raw data, as well as the p-values out of the binomial sign tests are reported below.

On the other hand, objective metric was used to quantify the degradation of the image quality. Namely, Structural Similarity (**SSIM**) - Full Reference image quality metric [19] - where the original image was considered a reference, with SSIM score of 1, while the SSIM score was found for two blurred images. 1 is considered best score (full similarity), while 0 is the worst case (no similarity). SSIM is one of the metrics used to measure Gaussian blur degradations [20]. It's worth mentioning that metrics, like BRISQUE, or blur-specific [21, 22], CPBD [23] and JNBM [24] failed to adequately quantify very high amount of blur.

Finally, Pearson's Linear Correlation Coefficients were found between the objective image quality assessment metric and the mean z-scores obtained for each of the psychometric scaling experiments.

Results & Discussion

Image Quality

SSIM metric reflects the changes in the image quality and the score has a decreasing tendency as the image is blurred. Please, refer to the Figure 7.

Figure 1. "Glass" full-scene image: Original (left), moderately blurred (middle), and highly blurred (right)

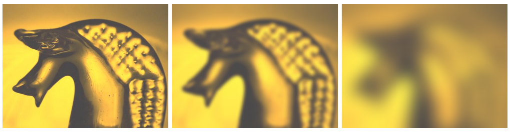

Figure 2. "Horse" full-scene image: Original (left), moderately blurred (middle), and highly blurred (right)

Figure 3. "Pot" full-scene image: Original (left), moderately blurred (middle), and highly blurred (right)

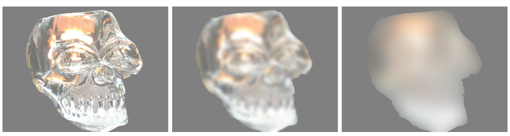

Figure 4. "Scull" isolated object image: Original (left), moderately blurred (middle), and highly blurred (right)

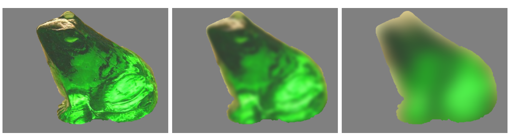

Figure 5. "Frog" isolated object image: Original (left), moderately blurred (middle), and highly blurred (right)

Figure 6. "Horse" isolated object image: Original (left), moderately blurred (middle), and highly blurred (right)

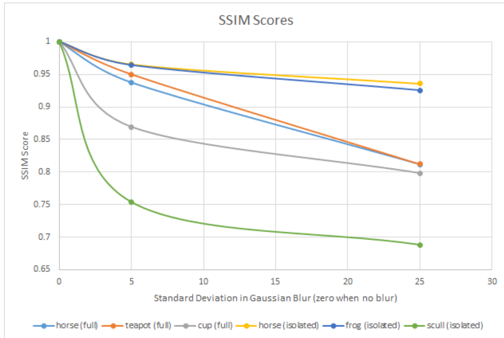

Figure 7. SSIM score as a function of the amount of blur

Psychometric Scaling Experiments

Z-scores of the psychometric scaling experiments are illustrated on Figure 8 and Figure 9. Figure 8 summarizes the z-scores of the three full-scene images with three different degrees of blurriness. As we can see on the figure, mean z-score for the undistorted image is always higher than that of its blurred versions. For all three images, there is no overlap of the confidence intervals between more blurred and less blurred versions of a particular scene (although there is a substantial overlap between 95% confidence intervals for different images (Cup, Horse, and Teapot)). Considering this clear separation, we can conclude that perceived degree of translucency decreases for a given object when the image is blurred. This is logical and intuitive for the images with high amount of blur, as high blur removes all the cues necessary for translucency perception (highlights, shades, background that is seen through, surface geometrical properties) and transforms the translucent object into a nearly homogeneous patch. On the other hand, when blur is moderate, translucency perception is impaired less dramatically in comparison with the original. Therefore, we can conclude that translucency perception impairment is correlated with the amount of degradation introduced.

Figure 9 illustrates z-scores for three isolated object images with three different degrees of blurriness, when the objects are seen in isolation on the neutral gray background. The trend remains the same as in case of the full-scene images: mean z-scores, i.e. perceived degree of translucency decreases, as the blurriness increases. However, in contrast with the full-scene images, the gap between mean z-scores, as well as between the confidence intervals of the different versions of the same image is less than that of full-scene images.

This is opposite to our expectation that access to the full-scene context might lead to higher translucency constancy. Whether impact of the full-scene information is statistically significant, needs further examination with larger dataset. As the images used for the two experiments are different, they are not directly comparable. The reason for the difference can be content of the image and characteristics of the objects, rather than the lack of access to the full context information. However, one of the explanations for this indication is that cropped objects, in contrast with the objects in blurred full-scene images, still stand out from the homogeneous background, considering that the edges are clear,

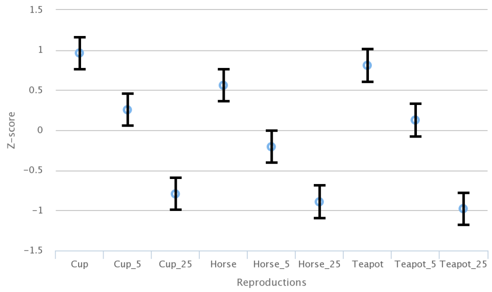

Figure 8. *Translucency z-scores for each of the examined full-scene images. The number after image name indicates the standard deviation of the Gaussian blur. The error bars and the blue circles show 95% confidence interval and the mean z-scores respectively. Same variance assumed for all the samples.*

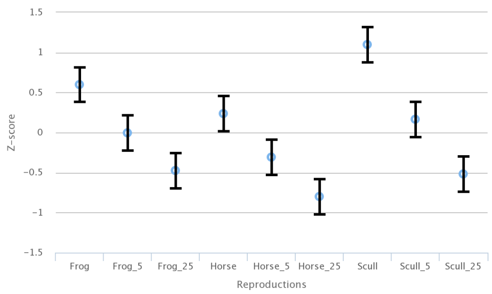

Figure 9. *Z-scores for each of the examined Isolated Object images. Same variance assumed for all the samples.*

evoking a perception of the object as a hole transmitting light. Besides, the highlights, texture other translucency cues might be more apparent when observing on the homogeneous background. This can be a topic for further investigation.

Pearson's Linear Correlation

The fact that image quality distortion is correlated with impairment of translucency perception, means that image quality assessment is important while working on quantification of perceived translucency and image quality metrics could be used to predict the extent to which translucency constancy could hold. In order to further examine this hypothesis, Pearson's Linear Correlation coefficients between SSIM values and mean z-scores were found. For Full-scene images, the correlation coefficient was significantly high - equal to 0.91. On the other hand, SSIM values and z-scores for Isolated Object images demonstrated little correlation, as the coefficient was equal to 0.36.

This can be explained with the fact, that the large area covered with neutral gray background in the isolated object images, leads to high structural similarity even for highly blurred images, while as we have already seen, blur significantly decreases mean z-scores for those kind of images (refer to Figure 10). Therefore, we found SSIM values from cropped objects only, disregarding gray background in the SSIM pooling step. However, correlation coefficient between new SSIM values and mean z-scores increased insignificantly - up to 0.42. The reason for this could be the content of the images: as shown on Figure 10, the scull has very complex shape and fine details that lead to higher structural dissimilarity when blurred. Hence, need for further investigation with larger and more diverse dataset, as well as for the application specific image quality metric arises and should be considered in the future work.

Figure 10. *The Quality Maps for moderately (left) and highly (right) blurred "Scull" (top) and "Glass" (bottom) images. Lighter areas mean higher similarity, darker areas mean less similarity with the original.*

In this particular case, we could have found correlation between z-scores and the amount of blur introduced (standard deviation of the Gaussian blur function) avoiding objective image quality metrics. In this case, high correlation has been demonstrated even for isolated images. However, in real-life situations information about the distortion might not be available. This is the primary reason, why it is very important to have objective metrics that quantify the amount of degradation and correlate well with the perceived degree of translucency.

Sign Tests

In order to further substantiate the credibility of our findings, we studied the raw data and conducted binomial sign tests on them. The raw data can be found in Tables 1 and 2, for full-scene - and isolated object images, respectively. The number in the cell signifies the number of the observers, which considered the object of the corresponding row more translucent than the object of the corresponding column. The names of the objects without numbers represent the original images, while the names with the numbers signify the blurred images with the number signifying the standard deviation of the Gaussian blur. The number of responses for each pair sums up to 40, as there were 20 observers and each pair was shown twice, in a flipped order. In order to compensate the problem of multiple comparisons, we applied Bonferroni[25] correction to our data.

As we observe for full-scene images, objects with high amount of blur are mostly significantly less translucent, except for the cases, when compared against other highly blurred images. Moderately blurred images are also significantly less translucent than the original ones. Refer to Table 1. The results are color-coded: if object in the corresponding row is significantly more translucent than the object in the corresponding column, the cell is green; if it is significantly less translucent, the cell is red; while

Table 1. The raw data of the observer responses for full-scene images. Raw p-values obtained from the binomial sign tests are given in the parentheses. Green cell: object in the corresponding row is significantly more translucent than the object in the corresponding column; Red cell: object in the corresponding row is significantly less translucent than the object in the corresponding column; White cell: no statistically significant difference.

	Cup	Cup_5	Cup_25	Horse	Horse_5	Horse_25	Teapot	Teapot_5	Teapot_25
Cup		34(8.36e-06)	34(8.36e-06)	29 (0.0064)	35(1.38e-06)	36(1.86e-07)	17 (0.4295)	35(1.38e-06)	35(1.38e-06)
Cup_5	6 (8.36e-06)		35(1.38e-06)	16 (0.2681)	26 (0.0806)	35(1.38e-06)	14 (0.0806)	18 (0.6358)	36(1.86e-07)
Cup_25	6 (8.36e-06)	5 (1.38e-06)		7 (4.23e-05)	11 (0.0064)	18 (0.6358)	7 (4.23e-05)	7 (4.23e-05)	16 (0.2681)
Horse	11 (0.0064)	24 (0.2681)	33(4.23e-05)		34(8.36e-06)	35(1.38e-06)	11 (0.0064)	33(4.23e-05)	36(1.86e-07)
Horse_5	5 (1.38e-06)	14 (0.0806)	29 (0.0064)	6 (8.36e-06)		33(4.23e-05)	8 (0.0001)	11 (0.0064)	33(4.23e-05)
Horse_25	4 (1.86e-07)	5 (1.38e-06)	22 (0.6358)	5 (1.38e-06)	7 (4.23e-05)		6 (8.36e-06)	5 (1.38e-06)	22 (0.6358)
Teapot	23 (0.4295)	26 (0.0806)	33(4.23e-05)	29 (0.0064)	32 (0.0001)	34(8.36e-06)		33(4.23e-05)	36(1.86e-07)
Teapot_5	5 (1.38e-06)	22 (0.6358)	33(4.23e-05)	7 (4.23e-05)	29 (0.0064)	35(1.38e-06)	7 (4.23e-05)		37(1.95e-08)
Teapot_25	5 (1.38e-06)	4 (1.86e-07)	24 (0.2681)	4 (1.86e-07)	7 (4.23e-05)	18 (0.6358)	4 (1.86e-07)	3 (1.95e-08)	

Table 2. The raw data of the observer responses for isolated object images. Raw p-values obtained from the binomial sign tests are given in the parentheses. Green cell: object in the corresponding row is significantly more translucent than the object in the corresponding column; Red cell: object in the corresponding row is significantly less translucent than the object in the corresponding column; White cell: no statistically significant difference.

	Frog	Frog_5	Frog_25	Horse	Horse_5	Horse_25	Scull	Scull_5	Scull_25
Frog		36(1.86e-07)	35(1.38e-06)	19 (0.8746)	34(8.36e-06)	35(1.38e-06)	6 (8.36e-06)	29 (0.0064)	30 (0.0022)
Frog_5	4 (1.86E-07)		35(1.38e-06)	14 (0.0806)	25 (0.1538)	33(4.23e-05)	5 (1.38e-06)	17 (0.4295)	29 (0.0064)
Frog_25	5 (1.38e-06)	5 (1.38e-06)		12 (0.0165)	18 (0.6358)	30 (0.0022)	5 (1.38e-06)	10 (0.0022)	25 (0.1538)
Horse	21 (0.8746)	26 (0.0806)	28 (0.0165)		30 (0.0022)	29 (0.0064)	10 (0.0022)	21 (0.8746)	25 (0.1538)
Horse_5	6 (8.36e-06)	15 (0.1538)	22 (0.6358)	10 (0.0022)		29 (0.0064)	5 (1.38e-06)	18 (0.6358)	22 (0.6358)
Horse_25	5 (1.38e-06)	7 (4.23e-05)	10 (0.0022)	11 (0.0064)	11 (0.0064)		4 (1.86e-07)	9 (0.0006)	13 (0.0384)
Scull	34(8.36e-06)	35(1.38e-06)	35(1.38e-06)	30 (0.0022)	35(1.38e-06)	36(1.86e-07)		34(8.36e-06)	36(1.86e-07)
Scull_5	11 (0.0064)	23 (0.4295)	30 (0.0022)	19 (0.8746)	22 (0.6358)	31 (0.0006)	6 (8.36e-06)		36(1.86e-07)
Scull_25	10 (0.0022)	11 (0.0064)	15 (0.1538)	15 (0.1538)	18 (0.6358)	27 (0.0384)	4 (1.86e-07)	4 (1.86e-07)	

white cells signify no statistically significant difference. The rows of the original images (Cup, Horse, and Teapot) are composed of 16 green, 8 white, and 0 red cells. The number of green cells decreases down to 8 for moderately blurred image rows, while there are 7 red, and 9 white cells. Finally, the rows corresponding highly blurred images are composed of just 17 red and 7 white cells. There is a clear trend that less blurry versions are considered more translucent by the observers.

On the other hand, difference is not significant in many cases when judging cropped objects. The original image rows are composed of 11 green cells, and 7 out of them is accounted for the "Scull" image that is considered the most translucent one among the nine images.

Besides, the amount of blur does not make significant difference between the versions of the Horse image. It is very interesting that this object at some extent demonstrates translucency constancy. One of the reasons for this could be the dark texture that can be perceived as being inside the object and that is present even on the blurred image.

Furthermore, highly blurred version of the Scull is significantly different only from other less blurry images of the Scull. Considering this, we could hypothesize that the impact of blur on the perception of very translucent objects is limited. However, the observers might be biased with their knowledge about the original Scull image - relying on that information regardless the appearance of the actual blurred version. Another reason could be that the Scull is achromatic with a lot of specularities that as has been demonstrated by another study [26] might also significantly impact translucency perception.

It is also worth mentioning that many differences might become more significant, if the experiment is conducted with higher number of observes. Example of this is illustrated on Table 3.

Table 3. P-values decrease, when the number of observers increases, but the portion of the observers with similar response remains the same.

Number of observers with similar responses	P-values
15 out of 20	0.04138947
30 out of 40	0.002221434
45 out of 60	0.000134514
60 out of 80	8.58E-06
75 out of 100	9.58E-07

Conclusion and Future Work

To summarize, we have introduced different amount of Gaussian blur to the Flickr Material Database images. Afterwards, the blurriness were quantified by objective image quality assessment metric and psychometric scaling experiments were conducted to determine, how introduction of blur impairs perception of translucency.

The data analysis has shown that for given images, blur significantly impairs translucency perception and the degree of impairment is correlated with the amount of image degradation.

We have also demonstrated that for full-scene images, SSIM objective image quality assessment has significant correlation with the perceived degree of translucency, while introduction of homogeneous background in isolated images, decreases this correlation. As examined image quality metrics, like BRISQUE, CPBD, and JNBM failed to adequately quantify high amount of blur, needs for more application specific metric arise.

Furthermore, there are some indications that the effect of blur is more dramatic when full scene is blurred. We hypothesize that cropped blurred images with sharp edges are unnatural and might evoke the perception of the object as a hole transmitting the light. Besides, the translucency cues might be more apparent on the homogeneous background. This can be a topic of the further study comparing appearance of identical objects in those two setups.

In order to model the impact of blur on translucency perception and identify the limits of translucency constancy, larger number of images, as well as smaller steps in blurriness variation are needed in the future study. More diverse database will also help figure out the fundamental reasons why blur impairs translucency perception and what are the cues people use for translucency assessment.

Finally, we were limited just to a single type of image distortion in this paper. In further study, we will examine how distortions other than blur, e.g noise, or compression artifacts, impact translucency perception and translucency constancy.

References

[1] Christian Eugene, "Measurement of "total visual appearance": a cie challenge of soft metrology," in *12th IMEKO TC1 TC7 Joint Symposium on Man, Science Measurement*, September 03.-05.2008 Annecy, France, pp. 61–65.

[2] Walter Gerbino, Casimir I Stultiens, Jim M Troost, and Charles M de Weert, "Transparent layer constancy," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 16, no. 1, pp. 3, 1990.

[3] Roland W Fleming and Heinrich H Bülthoff, "Low-level image cues in the perception of translucent materials," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 3, pp. 346–382, 2005.

[4] Bei Xiao, Bruce Walter, Ioannis Gkioulekas, Todd Zickler, Edward Adelson, and Kavita Bala, "Looking against the light: How perception of translucency depends on lighting direction," *Journal of Vision*, vol. 14, no. 3, pp. 17–17, 2014.

[5] Juno Kim and Phillip J Marlow, "Turning the world upside down to understand perceived transparency," *i-Perception*, vol. 7, no. 5, 2016.

[6] Lavanya Sharan, Ce Liu, Ruth Rosenholtz, and Edward H Adelson, "Recognizing materials using perceptually inspired features," *International journal of computer vision*, vol. 103, no. 3, pp. 348–371, 2013.

[7] Gabriel Schwartz and Ko Nishino, "Material recognition from local appearance in global context," *arXiv preprint arXiv:1611.09394*, 2016.

[8] Sean Bell, Paul Upchurch, Noah Snavely, and Kavita Bala, "Material recognition in the wild with the materials in context database," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7-12 June, 2015, Boston, Massachusetts, USA, pp. 3479–3487.

[9] Lavanya Sharan, Ruth Rosenholtz, and Edward H Adelson, "Accuracy and speed of material categorization in real-world images," *Journal of Vision*, vol. 14, no. 9, pp. 12–12, 2014.

[10] Javier Galbally, Sébastien Marcel, and Julian Fierrez, "Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 710–724, 2014.

[11] Elham Tabassi and Patrick Grother, "Fingerprint image quality," in *Encyclopedia of Biometrics*, pp. 482–490. Springer, 2009.

[12] Isamu Motoyoshi, "Highlight–shading relationship as a cue for the perception of translucent and transparent materials," *Journal of Vision*, vol. 10, no. 9, pp. 6–6, 2010.

[13] Khai Van Ngo, Jehans Jr. Storvik, Christopher André Dokkeberg, Ivar Farup, and Marius Pedersen, "Quickeval: a web application for psychometric scaling experiments," in *Image Quality and System Performance XII*. International Society for Optics and Photonics, 2015, vol. 9396, p. 93960O.

[14] Peter G Engeldrum, *Psychometric scaling: a toolkit for imaging systems development*, Imcotek, 2000.

[15] https://se.mathworks.com/help/images/ref/imgaussfilt.html, *Accessed: 26/02/2018*.

[16] Gaurav Sharma and Raja Bala, *Digital color imaging handbook*, CRC press, 2002.

[17] Jean Dickinson Gibbons and Subhabrata Chakraborti, "Nonparametric statistical inference," in *International Encyclopedia of Statistical Science*, pp. 977–979. Springer, 2011.

[18] Myles Hollander, Douglas A Wolfe, and Eric Chicken, *Nonparametric statistical methods*, John Wiley & Sons, 2013.

[19] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[20] Alain Hore and Djemel Ziou, "Image quality metrics: Psnr vs. ssim," in *Pattern recognition (icpr), 2010 20th international conference on*. IEEE, 2010, pp. 2366–2369.

[21] A Mittal, AK Moorthy, and AC Bovik, "Referenceless image spatial quality evaluation engine," in *45th Asilomar Conference on Signals, Systems and Computers*, 2011, vol. 38, pp. 53–54.

[22] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[23] Niranjan D Narvekar and Lina J Karam, "A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection," in *Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on*. IEEE, 2009, pp. 87–91.

[24] Rony Ferzli and Lina J Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb)," *IEEE transactions on image processing*, vol. 18, no. 4, pp. 717–728, 2009.

[25] C Bonferroni, "Teoria statistica delle classi e calcolo delle probabilita," *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, vol. 8, pp. 3–62, 1936.

[26] D Gigilashvili, JB Thomas, M Pedersen, and JY Hardeberg, "Behavioral investigation of visual appearance assessment," in *26th Color and Imaging Conference*, 2018.

Article H

Davit Gigilashvili, Midori Tanaka, Marius Pedersen, and Jon Yngve Hardeberg (2020). “Image Statistics as Glossiness and Translucency Predictor in Photographs of Real-world Objects.” In: *10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*. Vol. 2688. CEUR Workshop Proceedings, pp. 1–15

Image Statistics as Glossiness and Translucency Predictor in Photographs of Real-world Objects

Davit Gigilashvili[1], Midori Tanaka[1,2], Marius Pedersen[1], and Jon Yngve Hardeberg[1]

[1] The Norwegian Colour and Visual Computing Laboratory, Norwegian University of Science and Technology, Gjøvik, Norway
davit.gigilashvili@ntnu.no, marius.pedersen@ntnu.no, jon.hardeberg@ntnu.no

[2] Chiba University, Department of Imaging Sciences; Chiba, Japan
midori@chiba-u.jp

Abstract. We interpret our surrounding based on the visual stimuli, and perceive objects and materials around us to have various attributes, like color, glossiness, and translucency. We analyze the three-dimensional world based on the two-dimensional images detected by our retina. The state-of-the-art works conclude that the human visual system has a poor ability to fully understand and invert the complex optical nature of light and matter interaction. Some authors rather propose that the human brain calculates image statistics to perceive appearance, demonstrating correlation between perceptual attributes and various statistical metrics. However, the illustrated examples are usually unrealistic nearly-perfect stimuli, making real-life robustness of the findings questionable. In this study, we analyzed image statistics of photos of real world objects, and assessed the performance of statistical image metrics proposedly used by the human visual system. We identified very interesting trends, as well as limitations.

Keywords: Material appearance · image statistics · gloss · translucency

1 Introduction

Appearance is a complex psychovisual phenomenon that implies attributing particular characteristics to surrounding objects based on the interpretation of the visual data. CIE 175:2006 [23] (as quoted in [5]) defines appearance as *"the visual sensation through which an object is perceived to have attributes as size, shape, colour, texture, gloss, transparency, opacity, etc."* The CIE identifies color, gloss, translucency and texture as four major appearance attributes [23]. Appearance measurement has been developed towards hard metrology, i.e. instrumental measurements [12, 18, 30], and soft metrology relying on psychophysics [8, 22, 27].

While we orient ourselves in the 3-dimensional world we still interpret the environment based on the 2D retinal images. And we do pretty well: humans with normal vision can easily distinguish glossy and matte or translucent and opaque objects; furthermore, we are good at identifying materials, easily distinguishing ceramics from wax or human flesh from plastic dummies. Although multisensory information, like tactile or auditory, facilitate this process, the crucial amount of information is extracted from the above-mentioned 2D retinal images. Fleming and Bülthoff [6] have proposed that the human visual system (HVS) has poor optics inversion abilities, and that it relies on simple image cues to interpret material properties. Motoyoshi [21] tried to correlate image statistics with material properties, and found indications that skewness, or a similar measure of luminance histogram asymmetry, might be used by the HVS to judge surface properties. The finding is further supported and manifested by Landy [15]. Marlow and Adelson [16] demonstrated that sharpness, contrast and coverage area of the highlights are correlated with perceived level of glossiness. Qi et al. [24] tried to find correlation between perceived glossiness and various statistics of specular highlights, like spread, size, number, strength, and percentage coverage, and found a statistically significant correlation between the percentage of the highlight coverage and perceived glossiness.

Image statistics have been used for studying perceived translucency as well. Motoyoshi [20] manipulated images of various materials and concluded that *"spatial and contrast relationship between specular highlights and non-specular shading patterns is a robust cue for perceived translucency of three-dimensional objects"*. On the other hand, it has been also shown that image statistics alone do not entirely explain the complex nature of appearance perception and they are usually subject to multiple photo-geometric constraints [2, 13, 14, 16].

Although the above-mentioned findings are interesting, they are oftentimes based either on the synthetic stimuli, rendered in constrained and unrealistic environments, or few photographs taken in limited conditions. The studies using large photograph databases have no access to the physical ground truth of the material (e.g. [25, 31]), while wherever the ground truth is available, the number is stimuli is low (e.g. [21]).

The novelty of this study is using a photograph dataset with full access to the ground truth physical stimuli. We had a particular motivation for using photographs in this study. The vast majority of the authors using computer generated stimuli do not account for imperfections and artifacts present in the real world. As computer vision emerges, with autonomous vehicles among the most prominent applications, in-the-wild performance of particular metrics becomes vitally important for material identification. Therefore, we decided to extract image statistics not from the synthetic stimuli, but from photographs of real world objects coming with unintended artifacts, and to study the robustness of image statistics, as predictors for actual material appearance. We photographed objects with varying degree of gloss and translucency and described them with statistical metrics known to be correlated with them.

The paper is organized as follows: in the next section, we present the acquisition setup and methodology. The results are presented and discussed in Sections 3 and 4, respectively. Finally, we draw conclusions and summarize the potential directions for the the future work.

2 Methodology

2.1 Stimuli

We photographed spherical resin objects from the *Plastique* collection [28]. The objects have been created by an independent artist with an intention to be used in material appearance research. The resin substrate material is colored with different combinations of blue, yellow and white colorants, followed by different levels of surface processing (polishing). The objects come in three levels of surface coarseness that affects apparent gloss of the materials. We photographed 30 spheres in total with 3 different levels of surface roughness, 3 hues, and various levels of translucency (Fig. 1). It is worth mentioning that the objects have several visible artifacts, like scratches on the surface and bubbles in the volume, that make them good targets for testing the robustness of image-based metrics. The close-ups of some of the objects are shown in Fig. 2. Renderings of spherical

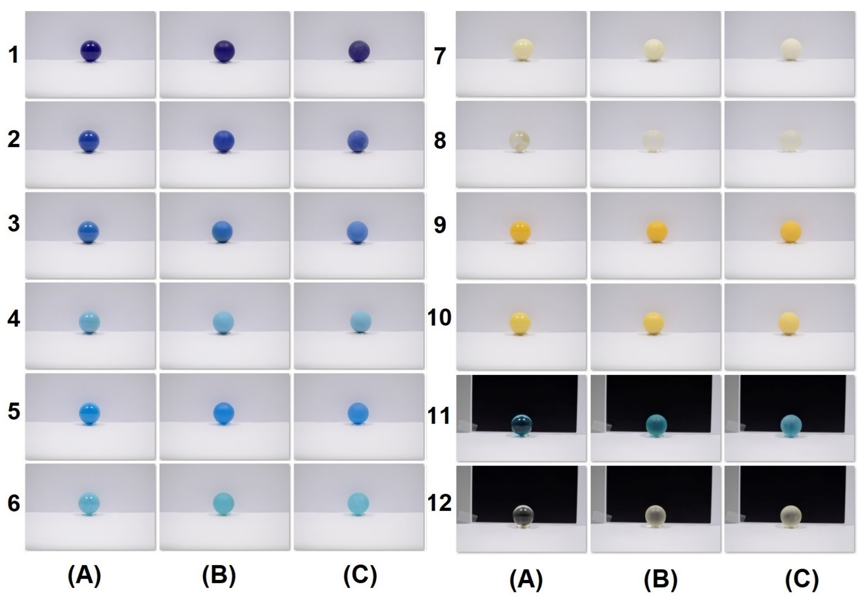

Fig. 1. The resin objects used as targets. Column A - objects with smooth surface; Column B - rougher objects; Column C - the roughest among the three. Objects in the same row are made of the identical material and differ only in surface processing.

objects are very commonly used in computer graphics for studying appearance (e.g. [22, 29]), and a simple curved shape of a sphere ensures apparent specular reflections, as well as distinctness of image gloss, that are very widely used cues for glossiness assessment by the HVS [8, 9, 22].

2.2 Image Acquisition

The objects were photographed in a GretagMacbeth Spectralight III viewing booth under diffuse D50 illumination with around 4900K color temperature. The illuminant is placed on the ceiling of the viewing booth, placing all objects under top-lit geometry - the most commonly encountered illumination geometry, both outdoors under sunlight, as well as in an office environment. The light intensity on the bottom of the viewing booth was 1858 lux, as it was directly exposed to the light, while it was 900 lux on the background. The acquisition setup is shown in Fig. 3.

The objects were placed on a white matte paper. Metal rings were used to fix the position of the spherical objects. In order to avoid possible bias from highly specular metal rings, they were covered with a white tape sticker. The immediate background of the object was white for opaque objects, while translucent objects were photographed twice, with black and white backgrounds. A Nikon D3200 camera was used with ISO 100, shutter speed 1/250 sec., F-stop 2.2, and 50mm focal distance. The object was located around 50 centimeters away from the camera. The camera was characterized using a MacBeth ColorChecker. The estimates of CIE XYZ values were obtained by a regression-based method using manufacturer-provided and camera-acquired color coordinates of the color checker patches. The color correction matrix was found by the least squares approximation. The spheres were segmented from the images of 3008×2000 pixels.

We are aware of the limitations related to the acquisition pipeline. Although the camera response function (CRF) has not been measured or estimated, the non-linearity of the CRF that is typical to consumer cameras might have affected the results. It is especially worth highlighting that the limited dynamic range of the acquisition system and clipping of the high luminance information could have impacted the recorded luminance histogram and its statistical moments.

2.3 Analysis of the Data

Only manually segmented images were studied and the background is not included in the statistics. It has been proposed that chromatic information has negligible impact on gloss perception [9, 22, 27]. In depth analysis of this is beyond the scope of this work. We assume that the vital portion of the information needed for glossiness estimation is embedded in luminance, and therefore, analyze the luminance channel Y from CIE XYZ. We found luminance histograms for each of the segmented objects and calculated the first four moments of it. Finally, the following statistical measures have been considered for the analysis: skewness and kurtosis of the luminance histogram, coverage of the highlights,

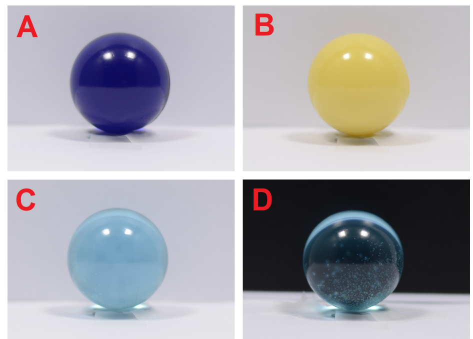

Fig. 2. The difference in contrast as well as in the reflected image is apparent between the dark blue and yellow smooth-surfaced objects (A and B). The object shown in illustrations C and D is the same, but its appearance differs due to the change in the background color. Some artifacts and bubbles are visible in image D.

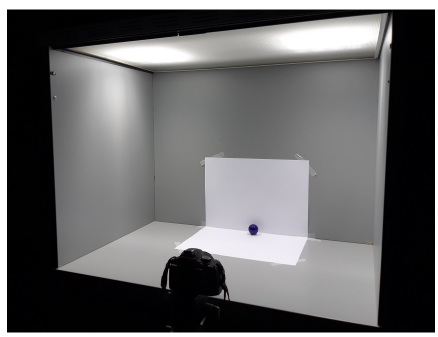

Fig. 3. The setup used for image acquisition.

mean luminance of the object, and standard deviation of the luminance distribution. The coverage was defined as the percent of the total surface covered by the areas which were larger than 20 pixels and had luminance value above 0.9 (luminance is normalized to 0-1 range, 1 corresponding to the largest luminance recorded by the acquisition system. We do not report cm/m^2 measurements). A correlation between gloss and the size of the highlights has been reported in the literature [16, 17]. Finally, we used these five statistical metrics for clustering the objects.

3 Results

The images of the 30 objects are shown in Figure 1. Objects shown in rows 11 and 12 are the same as the ones in rows 6 and 8, respectively, but photographed

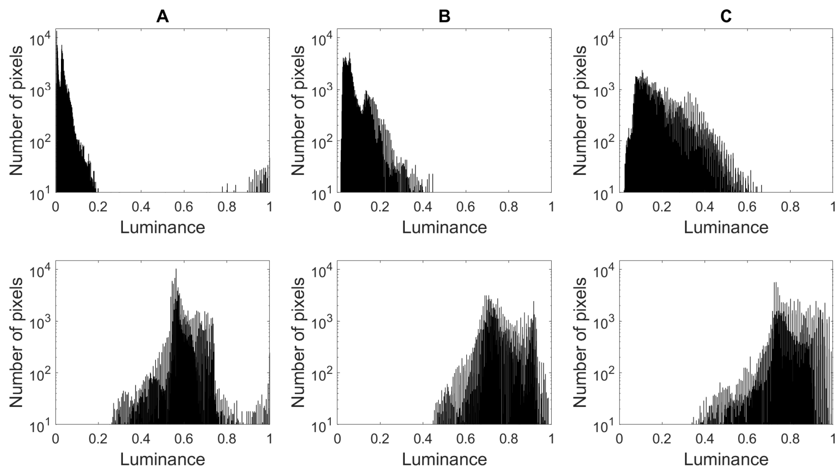

Fig. 4. Luminance histograms for segmented images. Top row - dark blue objects (row 1 in Fig. 1); bottom row - white objects (row 7 in Fig. 1). Column A - smoothest objects; Column C - roughest objects. The histograms show that the smoothest objects are positively skewed. As the mean luminance is lower, the skewness is stronger for the dark blue one. The histogram of the roughest white object is negatively skewed.

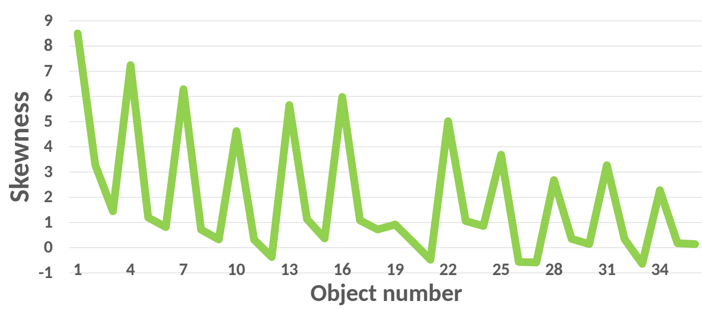

Fig. 5. Skewness of the luminance histogram for each segmented object (as to be counted from left to right, and top to bottom, in Fig. 1. Clear regularity of the triplets is visible in the pattern.

with the black background. Two major histogram asymmetry metrics have been studied: skewness and kurtosis. How the luminance histogram varies among different colors and levels of surface roughness is illustrated in Fig. 4. The results for skewness are shown in Fig. 5-6. As we see from the plots, the luminance histogram of the objects with smoother surface, i.e. higher gloss (difference in perceptual glossiness is apparent among the three levels of surface coarseness, although not quantified psychophysically), has always a positive skew, and the skewness is higher than that of the rougher, i.e. less glossy objects. Skewness difference between the two other surface levels is visible, but not large. A clear regular pattern for the triplets is visible in Fig. 5. If we refer to rows 1-4 in Fig. 1, the objects vary from darkest blue in row one, to lightest blue in row four. As we increase lightness of the object, the skewness of the luminance histogram decreases. Row 7 stands out on the plot with its low histogram skewness. This can be explained with the fact that the object is white, close to the illumination color. As the specular reflections on the surface are also whitish, they cause less skew in the luminance distribution, than for the low luminance bluish objects, where the tail of the distribution was high luminance specular highlights.

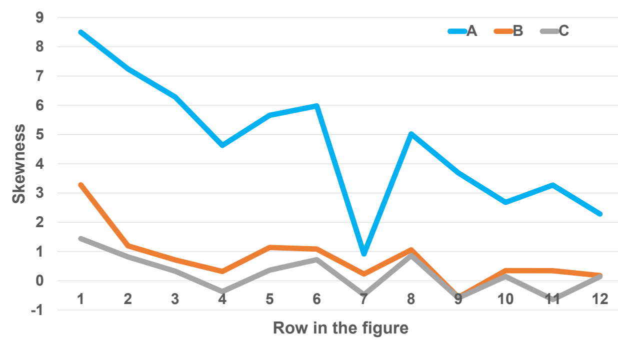

Fig. 6. Skewness of the luminance histogram for each segmented object. Each curve corresponds to each level of surface coarseness, as shown in columns in Fig. 1. Numbers in horizontal axis correspond to the rows in Fig. 1.

While skewness measures asymmetry towards particular direction, either positive, or negative skew, kurtosis measures general ”tailedness” of the distribution in both directions. Kurtosis for glossiest class of the objects is highest, and generally follows the same pattern, as it is for the skewness (refer to Fig. 7). However, the distinction between the two other classes is negligible with this measure.

The surface coverage by specular highlights was equal to zero for all rougher objects (columns B and C in Figure 1). The only exception was row 7, where the whitish color of the object biased our calculations and led to unreasonably many false positives. On the other hand, the coverage did not differ significantly among the smooth objects (column A), and the specular highlights covered around 0.8% of the total visible area of the sphere.

Mean luminance for each object is summarized in Fig. 8. Studying mean luminance can be interesting for two reasons: first of all, overall shininess of the object, as observed in [9], can evoke gloss perception in itself; secondly, it has been demonstrated [16, 22] that contrast between specular and diffuse areas, has significant impact on perceived gloss. Considering that specular highlights are white and nearly identical on all objects, we assume that mean luminance of the object is inversely correlated with the contrast gloss - i.e. higher the mean

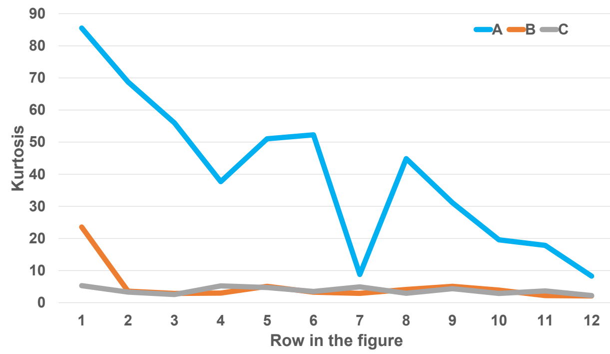

Fig. 7. Kurtosis of the luminance histogram for each segmented object. Each curve corresponds to each level of surface coarseness, as shown in columns in Fig. 1. Numbers in horizontal axis correspond to the rows in Fig. 1.

luminance of the entire object, lesser is the contrast between specular and diffuse areas. The objects with smoothest surface have less mean luminance than objects made of the identical material but with rougher surfaces. This can be explained with the fact that the substrate white material is exposed to the surface due to scratches, artifacts and irregularities presented on the rough surface.

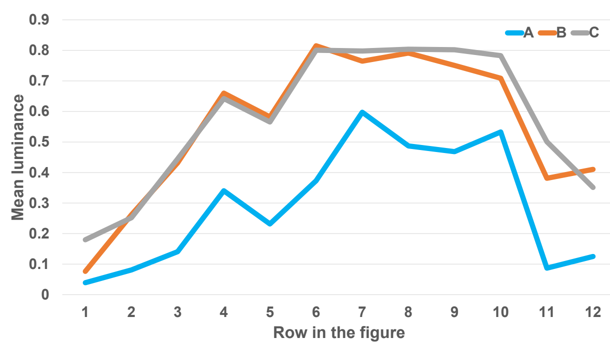

Fig. 8. Mean luminance for each segmented object. Each curve corresponds to each level of surface coarseness, as shown in columns in Fig. 1. Numbers in horizontal axis correspond to the rows in Fig. 1.

In contrast with the findings by Wiebel *et al.* [31], standard deviation of the luminance distribution is a poor predictor for surface coarseness class in our study (refer to Fig. 9). However, it significantly rises when the background of the object is changed from white to black. This might be a good indication that impact of the background on the luminance variance is a result of volume scattering - thus, used as a predictor of translucency for see-through objects.

In addition to this, we hypothesize that complexity of the scene might impact the statistics. We photographed objects in one additional condition placing a checkerboard-covered cube close to the object (the setup is shown in Fig. 10). The general trend is that rougher the surface is, the smaller the impact of the cube on statistical measures. This can be explained with the fact that the smooth surface has a good distinctness-of-image reflection, and the image reflected from the surface significantly impacts the statistics, while rough surfaces diffuse the light and no pattern is visible on the surface reflections. This trend deserves further attention.

Afterwards, we compared the metrics for the identical objects between white and black background photographing conditions (the results for rows 6 and 8 are compared with the results for rows 11 and 12, respectively). Interestingly, skewness and kurtosis decrease when the background is changed to black. To some extent, this can be accounted for many white-colored artifacts of the object which are visible on the black background only. As expected, mean luminance is decreased for black background due to absorption of the energy by the background, and thus, less back-reflections. Finding the ratio of the luminance measured on white and black backgrounds is an established technique for transmittance measurement of the flat objects (e.g. [10]). This observation holds at some extent for spherical objects as well. Also, as already discussed above, standard deviation changes significantly due to change in the complexity of the background. It has been demonstrated [9] that translucency, when objects are placed on a white background, can make objects look glossier. Here we observe that white background leads to more skewed luminance histograms that itself is proposedly related to gloss. Therefore, there might be a gloss-translucency cross-attribute interaction that is described by changes in image statistics. However, this needs further experimental evidence.

Finally, we conducted clustering to validate our hypothesis that the five statistical measures are good predictors for object class (smooth, moderately rough, and highly rough surfaces). We used *k-means* clustering with 3 clusters. Falsely detected highlight coverages for objects in the seventh row were manually set to 0. The cluster was defined as the centroid being the mean of all points in that particular cluster. Maximum number of iterations was set to 1000. Cluster centroids were initialized using *k-means++ algorithm* [3]. All objects with rougher surfaces (columns B and C) ended up in the same cluster. A small separate cluster was objects 1A and 2A, i.e. dark blue objects with low mean luminance, with the highest positive skew in luminance histogram. Four smooth-surfaced objects 7A, 10A, 11A, and 12A were clustered together with rough objects. While all other smooth objects were grouped together in a separate cluster. Clustering gives us an indication that five variables, the five statistical descriptors we use, might be enough to separate very smooth and glossy objects from rougher and less glossy objects. However, they fail describing intra-group differences.

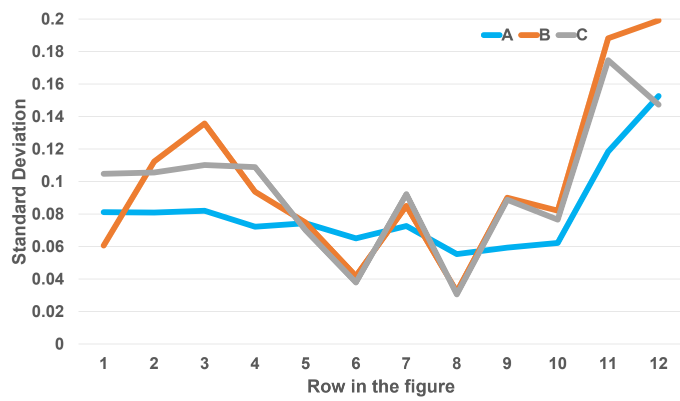

Fig. 9. Standard deviation of the luminance histogram for each segmented object. Each curve corresponds to each level of surface coarseness, as shown in columns in Fig. 1. Numbers in horizontal axis correspond to the rows in Fig. 1.

4 Discussion

We have observed that as the surface becomes rougher, skewness and kurtosis of the luminance histogram decrease, and the distribution becomes less tailed. While glossy objects look solid opaque, like billiard balls, rougher surfaces look milkier, and at some extent evoke illusion of subsurface scattering via surface scattering only that is not surprising considering that the HVS has poor optics inversion ability [6]. This can be an indication in support for Motoyoshi's proposal [20] that blurring non-specular regions, i.e. squeezing the tails towards the center, reflected in decreased skewness and kurtosis, can enhance translucency perception. On the other hand, it has been observed earlier [7] that translu-

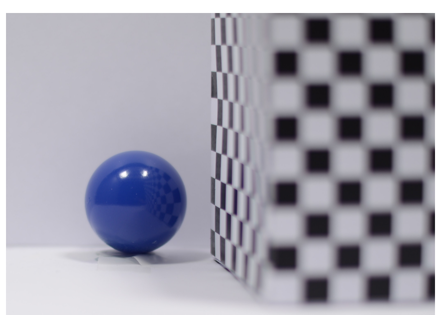

Fig. 10. A cube next to the object is covered with a checkerboard texture that is reflected on the surface of the sphere affecting its image statistics and appearance.

cency perception declines with blurring the entire object or image, i.e. when we decrease variance and histogram asymmetry. However, this proposal certainly needs validation psychophysically.

Although specular highlights are small and very simple in texture (just saturated blobs), covering less than 1% of the total visible area of the sphere, they strongly skew the luminance histogram, and evoke strong perception of gloss. Interestingly, van Assen *et al.* [4] have studied photographs with various patterns of highlights (disk, square, window etc.), and found that simpler specular highlights evoke stronger gloss perception than more complex ones. However, the role of the highlights should not be exaggerated, as the perception of gloss is a complex cognitive process and neither specular reflections are the only source of the highlights, nor all highlights evoke perception of glossiness. To demonstrate this, we have superimposed specular highlights of a smooth surface to a rougher surface of the identical materials (Fig. 11). In one case, the target rough object has relatively homogeneous texture, while in the other case, there are very apparent scratches and visual artifacts that help observers deduce the surface composition of the object. While glossiness for the former object looks reasonably realistic, the latter object does not look glossy as the highlights start looking more like artifacts. Presence of roughness cues limit perception of glossiness, although the statistical metrics are similar to that of glossy objects. This once again demonstrates photo-geometric constraints limiting the usage of image statistics as an appearance predictor. Interestingly, the HVS can still be tricked in particular scenarios when additional cues are missing (the manifestation of this phenomenon is the viral *glossy legs* illusion [19]).

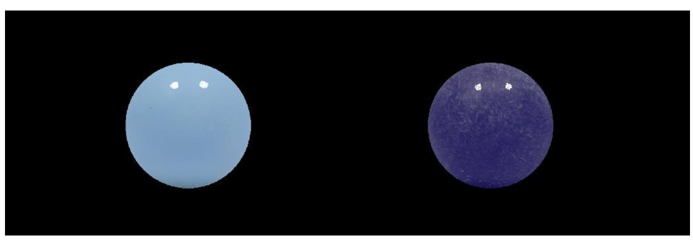

Fig. 11. While both highlights are artificial, the left object looks glossier due to the lack of artifacts, while the scratches help us know the right object is not smooth, i.e. not glossy.

Hunter [11] names contrast gloss, i.e. contrast between specular and diffuse areas, among one of the types, or dimensions of gloss. Pellacini *et al.* [22] have demonstrated that darker objects look glossier than lighter ones, and this effect has also been observed in other studies [9, 16, 29]. Although we did not have a direct measure for contrast in this work, considering that highlights were nearly identical among objects, we assumed that mean luminance of the entire object is inversely correlated with the contrast gloss. It has been demonstrated that up-to some threshold rough and light surfaces might look glossy [8, 24]. Moreover, it has been proposed that luminance information associated with shininess might significantly increase perceived glossiness [9]. Although mean luminance alone cannot be a good predictor for apparent gloss of the materials, it might carry rich information regarding contrast and distinctness-of-image (another dimension of gloss according to Hunter [11, 12]), and could be eventually included in the perceptual gloss model.

Standard deviation turned out a poor predictor of surface roughness class in our study. This interestingly contradicts with Wiebel *et al.* [31], who studied natural images, observed a strong positive correlation between standard deviation of luminace histogram and gloss, and found it a better predictor for gloss than skewness. Although we have not conducted perceptual measurements of our stimuli, we can draw some qualitative parallels. The inconsistency can be explained with the type of objects depicted in authors' natural images. If we examine the images illustrated in [31], we notice that images considered glossy consist of large segments of contrasting luminances, i.e. photographed complex shaped objects yield high number of pixels with highlights and also high number of dark pixels with shading - leading to large standard deviation. Unlike theirs, the highlights covered less than 1% of our stimuli, while the luminance gradient on the rest of the sphere was relatively homogeneous. This led to strong skew but was not enough for yielding high standard deviation in the luminance histogram.

Distinctness-of-the-reflected-image, the mirror image of the surrounding we can see on very glossy surfaces is another cue for glossiness. The background and surrounding vary dramatically in dynamic scenes, and hardly ever are as simple in real life, as studied in the laboratory conditions. Image statistics are neither static, nor consistent among different conditions. We observed that even a minor change in the environment (Fig. 10) can affect image statistics that makes its possible use by the HVS and even by machines, questionable. On the other hand, appearance is also dynamic; even though the HVS has ability to perceive some appearance attributes consistently across different conditions, i.e. demonstrates some constancy (e.g. color constancy), the constancy is valid up-to certain extent only, and completely fails in many conditions. While the vast majority of the studies trying to explain appearance with image statistics rely on a few images in very limited conditions, it remains an open question how appearance and image statistics co-vary. We have shown above that particular image statistics are promising and deserve further attention, but for more solid conclusions, psychometric measurements are needed. Understanding how image statistics correlate with perceived appearance can be beneficial in two ways:

- It can unveil further mechanisms that are used by the HVS to interpret the surrounding.
- It can have commercially significant applications in computer vision. Many image statistical metrics are extremely efficient computationally, and might be used for material identification and quality assurance. Moreover, generality across different conditions might not be the mandatory requirement for image statistics. Many computer vision applications are limited to very spe-

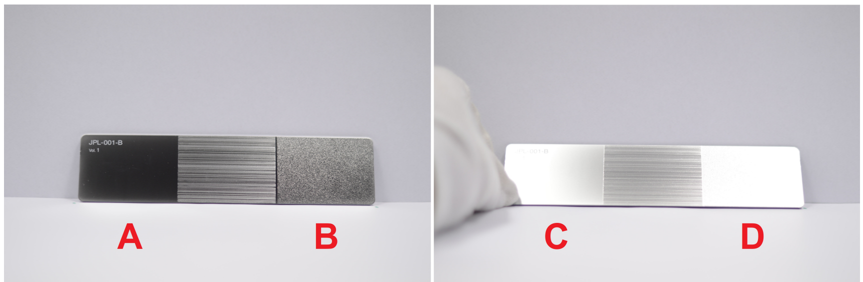

Fig. 12. The object shown in the left and right photos is the same. However, even a slight change in illumination angle leads to dramatic changes in its appearance. If the smoothest (region A) and the roughest (region B) surfaces are distinguishable in the left image, they (regions C and D) look nearly identical in the right one due to dynamic range limitations.

cific conditions by nature, and limits of particular statistics might not have vital importance, as long as a correlation between statistics and appearance is established for given (application-specific) conditions.

Finally, we should mention that above-discussed variation in luminance distributions was observed due to the curvature of the spherical objects. Our findings might not be applicable to other surfaces, especially to the planar ones. To demonstrate this, we tried photographing flat plastic and metallic samples from the JIDA Standard Sample dataset [1]. We have observed two interesting phenomena that made studying image statistics of these samples unreliable:

- Because the surface is flat, all points on small objects are under approximately the same illumination geometry that makes it impossible to see specular and diffuse areas separately, and the entire part of the patch looks rather homogeneous, essentially cutting down the luminance histogram to a single luminance value. This can be seen in the left image of Fig. 12, where the leftmost part of the patch (region A) is smooth and glossy, albeit homogeneous under given conditions.
- The samples, especially the metallic ones, are extremely prone to appearance changes even with a slight change in illumination geometry. This is demonstrated in Fig. 12.

Although haze and absence-of-textures on low gazing angles (further dimensions or types of gloss) could be observed on the flat patches, these phenomena are beyond the scope of this work and should be addressed in the future.

5 Conclusion and Future Work

We have taken photographs of real world objects and studied correlation between image statistics and actual physical surface properties. Although very clear positive skew of the luminance histogram is characteristic for smooth (and presumably glossy) surfaces, the robustness of the metric is challenged by complexity of the surrounding and semantic understanding of the scene and surface geometry. Furthermore, mean luminance can be correlated with contrast gloss, while change in variance across different conditions can be a predictor for translucency. It is worth mentioning that the dynamic range of our acquisition system was limited, and analysis of the high dynamic range data could reveal further interesting trends. Complex shapes and wider range of the materials should also be covered. While difference in perceptual gloss was assumed between smooth and rough surfaces, the statistics should be correlated with actual psychophysical measures in the future. Finally, more statistical measures, like entropy, and chromatic information should also be included in future studies and the performance of simple image statistics should be compared with that of the complex machine learning (e.g. *deep learning*) models. It has been demonstrated recently [26] that unsupervised learning techniques outperform image statistics and even supervised learning techniques in prediction of human perception. This is an interesting avenue that not only provides basis for reliable computer vision systems, but can also reveal curious mechanisms of the human vision.

References

1. JIDA Standard Samples. (Accessed on 21/11/2019 at), https://www.jida-online.shop/product/5
2. Anderson, B.L., Kim, J.: Image statistics do not explain the perception of gloss and lightness. Journal of Vision **9**(11:10), 1–17 (2009)
3. Arthur, D., Vassilvitskii, S.: k-means++: The advantages of careful seeding. In: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. pp. 1027–1035. Society for Industrial and Applied Mathematics (2007)
4. van Assen, J.J.R., Wijntjes, M.W., Pont, S.C.: Highlight shapes and perception of gloss for real and photographed objects. Journal of Vision **16**(6:6), 1–14 (2016)
5. Eugène, C.: Measurement of "total visual appearance": a CIE challenge of soft metrology. In: 12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement. pp. 61–65 (2008)
6. Fleming, R.W., Bülthoff, H.H.: Low-level image cues in the perception of translucent materials. ACM Transactions on Applied Perception (TAP) **2**(3), 346–382 (2005)
7. Gigilashvili, D., Pedersen, M., Hardeberg, J.Y.: Blurring impairs translucency perception. In: Color and Imaging Conference. vol. 2018, pp. 377–382. Society for Imaging Science and Technology (2018)
8. Gigilashvili, D., Thomas, J.B., Hardeberg, J.Y., Pedersen, M.: Behavioral investigation of visual appearance assessment. In: Color and Imaging Conference. vol. No. 1, pp. 294–299. Society for Imaging Science and Technology (2018)
9. Gigilashvili, D., Thomas, J.B., Pedersen, M., Hardeberg, J.Y.: Perceived glossiness: Beyond surface properties. In: Color and Imaging Conference. pp. 37–42. No. 1, Society for Imaging Science and Technology (2019)
10. Hajipour, A., Shams Nateri, A., Sadr Momtaz, A.: Estimation of fabric opacity by scanner. Sensor Review **34**(4), 404–409 (2014)

11. Hunter, R.S.: Methods of determining gloss. NBS Research paper RP **958** (1937)
12. Hunter, R.S., Harold, R.W.: The measurement of appearance. John Wiley & Sons (1987)
13. Kim, J., Anderson, B.L.: Image statistics and the perception of surface gloss and lightness. Journal of Vision **10**(9:3), 1–17 (2010)
14. Kim, J., Marlow, P., Anderson, B.L.: The perception of gloss depends on highlight congruence with surface shading. Journal of Vision **11**(9:4), 1–19 (2011)
15. Landy, M.S.: Visual perception: A gloss on surface properties. Nature **447**(7141), 158–159 (2007)
16. Marlow, P.J., Anderson, B.L.: Generative constraints on image cues for perceived gloss. Journal of vision **13**(14:2), 1–23 (2013)
17. Marlow, P.J., Kim, J., Anderson, B.L.: The perception and misperception of specular surface reflectance. Current Biology **22**(20), 1909–1913 (2012)
18. Marschner, S.R., Westin, S.H., Lafortune, E.P., Torrance, K.E., Greenberg, D.P.: Image-based brdf measurement including human skin. In: Rendering Techniques' 99, pp. 131–144. Springer (1999)
19. Molloy, M.: Confused by this shiny leg optical illusion? Here's how it works. The Telegraph (Accessed on 04/01/2019 at https://wwwtelegraphcouk/news/2016/10/31/confused-by-this-shiny-leg-optical-illusion-heres-how-it-works)
20. Motoyoshi, I.: Highlight–shading relationship as a cue for the perception of translucent and transparent materials. Journal of vision **10**(9:6), 1–11 (2010)
21. Motoyoshi, I., Nishida, S., Sharan, L., Adelson, E.H.: Image statistics and the perception of surface qualities. Nature **447**(7141), 206 (2007)
22. Pellacini, F., Ferwerda, J.A., Greenberg, D.P.: Toward a psychophysically-based light reflection model for image synthesis. In: Proceedings of the 27th annual conference on Computer graphics and interactive techniques. pp. 55–64. ACM Press/Addison-Wesley Publishing Co. (2000)
23. Pointer, M.: A framework for the measurement of visual appearance. CIE Publication pp. 175–2006 (2006)
24. Qi, L., Chantler, M.J., Siebert, J.P., Dong, J.: Why do rough surfaces appear glossy? JOSA A **31**(5), 935–943 (2014)
25. Sharan, L., Li, Y., Motoyoshi, I., Nishida, S., Adelson, E.H.: Image statistics for surface reflectance perception. JOSA A **25**(4), 846–865 (2008)
26. Storrs, K.R., Fleming, R.W.: Unsupervised learning predicts human perception and misperception of specular surface reflectance. bioRxiv p. 25 pages (2020)
27. Tanaka, M., Horiuchi, T.: Investigating perceptual qualities of static surface appearance using real materials and displayed images. Vision research **115**, 246–258 (2015)
28. Thomas, J.B., Deniel, A., Hardeberg, J.Y.: The plastique collection: A set of resin objects for material appearance research. XIV Conferenza del Colore, Florence, Italy p. 12 pages (2018)
29. Thomas, J.B., Hardeberg, J.Y., Simone, G.: Image contrast measure as a gloss material descriptor. In: International Workshop on Computational Color Imaging. pp. 233–245. Springer (2017)
30. Wetlaufer, L., Scott, W.: The measurement of gloss. Industrial & Engineering Chemistry Analytical Edition **12**(11), 647–652 (1940)
31. Wiebel, C.B., Toscani, M., Gegenfurtner, K.R.: Statistical correlates of perceived gloss in natural images. Vision Research **115**, 175–187 (2015)

Article I

Davit Gigilashvili, Jean Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (2020). “On the Nature of Perceptual Translucency.” In: *8th Annual Workshop on Material Appearance Modeling (MAM2020). Eurographics Digital Library*, pp. 17–20

Workshop on Material Appearance Modeling (2020)
H. Rushmeier and R. Klein (Editors)

On the Nature of Perceptual Translucency

Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg and Marius Pedersen

Deptartment of Computer Science, Norwegian University of Science and Technology, Gjøvik, Norway

Figure 1: *We encounter materials permitting some degree of subsurface light transport, described as transparent or translucent.*

Abstract

Translucency is an appearance attribute used to characterize materials with some degree of subsurface light transport. Although translucency as a radiative transfer inside the medium is relatively well understood, translucency as a perceptual attribute leaves much room for interpretation. Our understanding of the translucency perception mechanisms of the human visual system remains limited. No agreement exists on how to quantify perceived translucency, how to compare translucency of multiple objects and materials, how translucency relates to transparency and opacity, and what are the perceptual dimensions of it. We highlight the challenges in perception research arisen by these ambiguities and argue for the need for standardization.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [**Computer Graphics**]: Methodology and Techniques—
—Standards J.4 [**Social and Behavioral Sciences**]: Psychology—

1. Introduction

When we speak of translucency, we usually mean materials that at some degree permit subsurface light transport. Several optical material properties are used in the *Radiative Transfer Equation* [Cha60] to characterize the light propagation inside the medium, such as, absorption and scattering coefficients, scattering phase function, and index of refraction.

The human visual system is adept at detecting subsurface light transport, perceiving materials to be translucent. For instance, we do not need prior training to judge whether a material transmits light, or to tell the difference between real human skin and a plastic dummy, between translucent glass and opaque metal. Although perceptual aspects of translucency is a topic of interest in academia [FB05] and industry (e.g. in 3D printing [BATU18, UTB*19]) alike, our knowledge about the psychovisual mechanisms of translucency perception remains limited. Fleming and Bülthoff [FB05] proposed that the human visual system relies on low level image cues to judge translucency. Gkioulekas *et al.* [GXZ*13] studied the impact of the scattering phase function on translucency perception, while Xiao *et al.* [XWG*14] demonstrated that perceptual translucency is not a constant property and it depends on the illumination direction. Despite those attempts, a lot of uncertainties remain about the concept of perceptual translucency. Below, we will discuss multiple challenges we have faced due to this ambiguity throughout the process of psychophysical studies of translucency perception, making results inconsistent and difficult to interpret.

DOI: 10.2312/mam.20201141

2. Open questions about perceptual translucency

2.1. Definition and conceptual understanding

Translucency is considered a major appearance attribute by the CIE [Poi06, Eug08] alongside color, gloss and texture. No single standard definition of translucency exists. ASTM - Standard Terminology of Appearance [AST17] defines translucency as *"the property of a specimen by which it transmits light diffusely without permitting a clear view of objects beyond the specimen and not in contact with it."*. According to Gerbino [GSTdW90], *"transparent substances, unlike translucent ones, transmit light without diffusing it."* Eugène [Eug08] also highlights diffusing-blurring nature of translucency, arguing that *"if it is possible to see only a "blurred" image through the material (due to some diffusion effect), then it has a certain degree of transparency and we can speak about translucency"*. However, the author believes that *"a single and simple definition of translucency is unlikely to be achieved."* According to the CIE [Poi06], *"translucency is a subjective term that relates to a scale of values going from total opacity to total transparency."* In non-scientific contexts, *translucent* as an adjective can be used to describe the scattering, as well as clear transparent media [web]. While these definitions usually refer to the physical property of light scattering, the term is still vague in terms of perception, as it does not reflect in what way physical properties relate to appearance (except for "blurring"), making it subject to individual interpretation.

2.2. Perceptual dimensions of translucency

One of the major challenges regarding translucency is to identify its perceptual dimensions. For example, various perceptual dimensions exist to describe color - such as, hue, chromaticity or lightness. The same is true for gloss. Hunter [Hun37] proposed six dimensions of gloss (specular gloss, contrast gloss, distinctness-of-reflected-image, absence-of-bloom, absence-of-surface-texture, and sheen). Pellacini *et al.* [PFG00] identified two perceptual dimensions of gloss: contrast and distinctness. It is not clear yet what would be similar perceptual dimensions for translucency, although there is evidence that they might exist. The authors of this paper have conducted psychophysical experiments studying translucency perception [GTHP18, GUT*19, GDPH20]. We have observed that the subjects find it challenging to interpret the term and to identify the dimensions for quantifying it. They could not decide which cue to prioritize: complexity of light and matter interaction, i.e. preservation of structure of the light - clarity of the image seen-through the material, or preservation of the radiometric values (the amount of transmitted light). What if we compare very dark transparent-looking material with little scattering against the lighter one with less absorption but higher scattering? (refer to Fig. 2). These observations are consistent with Eugène's [Eug08] proposal that *"the concept of translucency can perhaps be regarded as a generic and subjective term, combining the concepts of* ***clarity*** *("ability to perceive the fine details of images through the material") and* ***haze*** *("property of the material whereby objects viewed through it appear to be reduced in contrast")* - also admitting that much work is still needed to clear up these uncertainties.

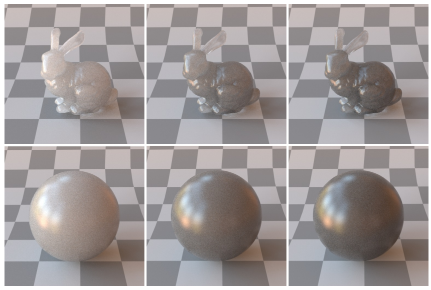

Figure 2: *Objects in the same column are made of the identical material. However, due to smaller scale and presence of thin parts, the Bunny has more cues evoking perception of translucency. Objects in the first column have high scattering and low absorption. In the second column - lower scattering and higher absorption. In the third column - same scattering as in the second column - but higher absorption. How can we compare their perceptual translucency?*

2.3. Relation with transparency and opacity

Another reason why the term leads to confusion is the lack of knowledge how it relates to transparency and opacity. Eugène [Eug08] proposes that translucency is related to transparency and opacity but does not discuss how. Gerardin *et al.* [GSF*19] propose that increasing subsurface scattering of the transparent material makes it translucent and eventually opaque, while adding absorption to a fully transparent material gradually makes it opaque, but never - translucent. This definition was not accepted by some of our subjects.

It is not clear whether transparency, translucency and opacity are on the same line of continuum, whether they are mutually exclusive or they can co-exist. Can a material possess some degree of transparency and translucency, or some degree of translucency and opacity at the same time? When do transparent materials start to be considered translucent, or when do translucent ones become opaque? Transparency and opacity seem to be ranges across the spectrum of light transmission properties rather than extreme discrete points. We have demonstrated that opacity is a subjective term and does not imply complete absence of transmission [GTHP18] (further supported by [GMH19]). It seems that the conceptual boundary between transparency, translucency and opacity is fuzzy - although the amount of translucency could be characterized with a bell-shaped curve that gradually increases, reaching a peak and then decreasing again while moving from transparency to opacity [Per] (refer to Fig. 3).

2.4. How to quantify perceptual translucency?

Limited knowledge on how to quantify translucency and how it relates to other perceptual properties of subsurface light transport (transparency and opacity) makes it challenging to apply magnitude estimation techniques [Tor58] to quantify translucency of a given material, or psychophysical scaling methods, such as pair comparison and rank order [Eng00], to compare the stimuli with one an-

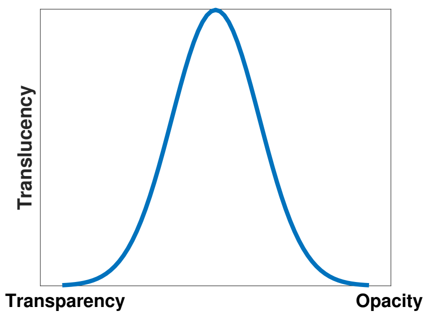

Figure 3: *Translucency might be gradually increasing, reaching its peak and decreasing between transparency and opacity. However, transparency and opacity are unlikely to be discrete points and translucency can co-exist with them.*

other. As there is an intuitive spectrum of glossiness properties from a Lambertian matte to a perfect mirror, it has been demonstrated to be feasible to estimate magnitude of glossiness [PFG00], or to identify glossier and less glossy objects when comparing multiple stimuli [THS17, GTPH19]. However, we faced a challenge with interpretation of the term when similar approach was applied to translucency. The subjects found it challenging to rank the stimuli by translucency, from the most translucent to the least translucent one [GTHP18]. What does *more translucent* mean? How would we tell which stimulus is more translucent? (e.g. in Fig. 1 and 2) Is it the one closer to transparency, opacity, or the center of the hypothetical transparency-opacity axis? Does higher scattering or absorption make materials more translucent? When does translucency peak, is correlation between scattering and translucency monotonous? These have been the questions we have not been able to answer.

The state-of-the-art works experimenting on translucency perception avoid quantifying translucency and abstain from comparing *more* and *less translucent* stimuli. They rather encapsulate this in matching and similarity detection tasks, asking observers to match the stimuli by appearance [XWG*14, XZG*19, FB05] and/or to select similar ones by translucency [GXZ*13, GUT*19]. While this task is less ambiguous and easier to interpret for the subjects, it has not been demonstrated up-to date that the human visual system can isolate translucency from total appearance. This creates the risk the observers making up their own rules matching the stimuli by total appearance, by lightness, or any property other than translucency. If the definition of translucency is not clear, how can they judge translucency similarity?

2.5. Translucency constancy of objects and materials

Similarly to our work [GTHP18], Nagai *et al.* [NOT*13] asked subjects to identify *more translucent* stimulus, interpreting it as having stronger subsurface scattering. However, definition of translucency as a material property does not adequately convey the complex nature of translucent appearance. We believe that in addition to physical material properties at least three other factors - illumination geometry, the size of the object and its shape should be considered. An object looks more translucent [XWG*14, FB05, GTHP18] and less opaque [GMH19] in back-lit conditions. It has been shown that scale and overall thickness of the object [FB05, UTB*19], as well as presence of thin regions [GTHP18, GUT*19] impact perceived translucency.

The majority of the observers in our studies [GTHP18, GUT*19] had difficulty comparing objects with different shapes due to the ambiguity between object-specific translucency and translucency as a shape-independent physical material property (e.g. what if a material is fully transparent, but complex shape, surface geometry or roughness do not permit to see-through the object - is it still transparent?) Moreover, it was problematic to come up with a single translucency measure for an object with a complex shape and varying thickness (refer to a female bust with thick torso and thin cloth areas in Fig.1). Hutchings [Hut94] proposes that heterogeneous material might have *"more than one colour, perhaps more than one translucency, gloss, or surface irregularity"* that no appearance profile system can deal with. Should translucency of an object be assessed as a whole, as a global attribute, or should it be taken as a local, region-specific one?

We believe perceptual translucency is a context-dependent attribute with limited constancy and mapping physical material properties with a visual attribute is a surjective but non-injective function - several different physical properties evoking identical perception of translucency. If we draw a parallel with color, material translucency could be analogous with spectral reflectance as an objective physical material property, and object translucency - with color, both being perceptual by definition. However, there exist physiological color matching functions with no interpretation, while no physiological functions have been found or described so far for perceived translucency. This could be explained with a fact that perception of translucency is a more complex phsychovisual phenomenon, involving spatial properties, contrast and various image cues [FB05]. As it is possible to fix physiological state, there exists a standard observer for color. However, physiology of translucency perception is not understood, leaving room for further research. While perceived translucency would more logically be compared with color appearance, no translucency counterpart is identified for colorimetry yet. It is likely that translucency measurement will be context-specific, customized to individual circumstances.

3. Conclusion

To summarize, our experience with psychophysical experiments has shown that there is an obvious need for translucency measurement, comparison and definition standards. The lack of an established procedure for perceptual translucency measurement makes tasks ambiguous and inconsistent. There is a clear disagreement among observers regarding its dimensionality. Although they always found a strategy to tackle a particular task, their "solutions" do not necessarily express what they perceive. A rigorous future work is needed to identify perceptual dimensions of translucency, if any. Revealing particular dimensions will make psychophysical measurements more consistent and easier to interpret. The definitions of translucency imply an absolute, objective attribute of a specimen. We believe the definition should reflect its situation-dependence and perceptual nature, proposing the following reformulation of [AST17]: *"translucency - the property of a specimen by which it evokes perception of subsurface light transport* ***under given conditions****."*

References

[AST17] ASTM E284-17 Standard Terminology of Appearance. ASTM International, West Conshohocken, PA. URL: https://doi.org/10.1520/E0284-17. 18, 20

[BATU18] BRUNTON A., ARIKAN C. A., TANKSALE T. M., URBAN P.: 3D printing spatially varying color and translucency. *ACM Transactions on Graphics (TOG) 37*, 4 (2018), 157:1–157:13. 17

[Cha60] CHANDRASEKHAR S.: Radiative transfer. (1960) Dover Publications Inc. New York, pp. 1–53. 17

[Eng00] ENGELDRUM P. G.: *Psychometric scaling: a toolkit for imaging systems development*. Imcotek, 2000. 18

[Eug08] EUGÈNE C.: Measurement of "total visual appearance": a CIE challenge of soft metrology. In *12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement* (2008), pp. 61–65. 18

[FB05] FLEMING R. W., BÜLTHOFF H. H.: Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception (TAP) 2*, 3 (2005), 346–382. 17, 19

[GDPH20] GIGILASHVILI D., DUBOUCHET L., PEDERSEN M., HARDEBERG J. Y.: Caustics and translucency perception. In *Material Appearance 2020, IS&T International Symposium on Electronic Imaging* (2020), Society for Imaging Science and Technology, pp. 033:1–033:6. 18

[GMH19] GIGILASHVILI D., MIRJALILI F., HARDEBERG J. Y.: Illuminance impacts opacity perception of textile materials. In *Color and Imaging Conference* (2019), Society for Imaging Science and Technology, pp. 126–131. 18, 19

[GSF*19] GERARDIN M., SIMONOT L., FARRUGIA J.-P., IEHL J.-C., FOURNEL T., HÉBERT M.: A translucency classification for computer graphics. In *Material Appearance 2019, Electronic Imaging* (2019), Society for Imaging Science and Technology, pp. 203:1–203:6. 18

[GSTdW90] GERBINO W., STULTIENS C. I., TROOST J. M., DE WEERT C. M.: Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance 16*, 1 (1990), 3. 18

[GTHP18] GIGILASHVILI D., THOMAS J.-B., HARDEBERG J. Y., PEDERSEN M.: Behavioral investigation of visual appearance assessment. In *Color and Imaging Conference* (2018), no. 1, Society for Imaging Science and Technology, pp. 294–299. 18, 19

[GTPH19] GIGILASHVILI D., THOMAS J.-B., PEDERSEN M., HARDEBERG J. Y.: Perceived glossiness: Beyond surface properties. In *Color and Imaging Conference* (2019), no. 1, Society for Imaging Science and Technology, pp. 37–42. 19

[GUT*19] GIGILASHVILI D., URBAN P., THOMAS J.-B., HARDEBERG J. Y., PEDERSEN M.: Impact of shape on apparent translucency differences. In *Color and Imaging Conference* (2019), Society for Imaging Science and Technology, pp. 132–137. 18, 19

[GXZ*13] GKIOULEKAS I., XIAO B., ZHAO S., ADELSON E. H., ZICKLER T., BALA K.: Understanding the role of phase function in translucent appearance. *ACM Transactions on graphics (TOG) 32*, 5 (2013), 1–19. 17, 19

[Hun37] HUNTER R. S.: Methods of determining gloss. *NBS Research paper RP 958* (1937), 19–39. 18

[Hut94] HUTCHINGS J. B.: Appearance profile analysis and sensory scales. In *Food Colour and Appearance*. Springer, 1994, pp. 142–198. 19

[NOT*13] NAGAI T., ONO Y., TANI Y., KOIDA K., KITAZAKI M., NAKAUCHI S.: Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study. *i-Perception 4*, 6 (2013), 407–428. 19

[Per] Unpublished personal communication with Prof. Holly Rushmeier of Yale University. 18

[PFG00] PELLACINI F., FERWERDA J. A., GREENBERG D. P.: Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (2000), ACM Press/Addison-Wesley Publishing Co., pp. 55–64. 18, 19

[Poi06] POINTER M.: A framework for the measurement of visual appearance. *CIE Publication. CIE 175:2006 ISBN: 978 3 901906 52 7* (2006). 18

[THS17] THOMAS J.-B., HARDEBERG J. Y., SIMONE G.: Image contrast measure as a gloss material descriptor. In *International Workshop on Computational Color Imaging* (2017), Springer, pp. 233–245. 19

[Tor58] TORGERSON W. S.: Theory and methods of scaling. 1958, Wiley: New York. 18

[UTB*19] URBAN P., TANKSALE T. M., BRUNTON A., VU B. M., NAKAUCHI S.: Redefining a in RGBA: Towards a standard for graphical 3D printing. *ACM Transactions on Graphics (TOG) 38*, 3 (2019), 1–14. 17, 19

[web] Merriam-Webster Dictionary. https://www.merriam-webster.com/dictionary/translucent. Accessed: 2020-11-06. 18

[XWG*14] XIAO B., WALTER B., GKIOULEKAS I., ZICKLER T., ADELSON E., BALA K.: Looking against the light: How perception of translucency depends on lighting direction. *Journal of Vision 14*, 3:17 (2014), 1–22. 17, 19

[XZG*19] XIAO B., ZHAO S., GKIOULEKAS I., BI W., BALA K.: Effect of geometric sharpness on translucent material perception. *bioRxiv* (2019), 795294. 19

Article J

Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen (n.d.). "Translucency perception: A review." In: *Accepted for publication in the Journal of Vision*, 45 pages

Translucency Perception: A Review

Davit Gigilashvili Norwegian University of Science and Technology
Department of Computer Science, Gjøvik, Norway

Jean-Baptiste Thomas Norwegian University of Science and Technology
Department of Computer Science, Gjøvik, Norway

Jon Yngve Hardeberg Norwegian University of Science and Technology
Department of Computer Science, Gjøvik, Norway

Marius Pedersen Norwegian University of Science and Technology
Department of Computer Science, Gjøvik, Norway

Translucency is an optical and a perceptual phenomenon which characterizes subsurface light transport through objects and materials. Translucency as an optical property of a material relates to the radiative transfer inside and through this medium, while translucency as a perceptual phenomenon describes the visual sensation experienced by humans when observing a given material under given conditions. The knowledge about the visual mechanisms of the translucency perception remains limited. Accurate prediction of appearance of the translucent objects can have a significant commercial impact in the fields such as 3D printing. However, little is known how the optical properties of a material relate to a perception evoked in humans. This article overviews the knowledge status about the visual perception of translucency and highlights the applications of the translucency perception research. Furthermore, this review summarizes current knowledge gaps, fundamental challenges and existing ambiguities with a goal to facilitate translucency perception research in the future.

Keywords: translucency perception, material appearance
The work has been funded by the Research Council of Norway.

Introduction

How different objects and materials appear to human observers is important not only in commerce, where customer choice and satisfaction are often influenced by the visual look of the product, but also in trivial daily tasks performed by humans. For instance, we use the visual appearance information to judge whether materials are fragile or elastic, whether food is spoiled or edible. By their appearance, we can effortlessly identify materials within seconds (Sharan et al., 2009; Wiebel et al., 2013). According to the International Commission on Illumination (the CIE - Commission Internationale de l'Eclairage) total appearance *"points out the visual aspects of objects and scenes."* (Pointer, 2006) Translucency is among the most essential visual attributes of appearance, along with color, gloss and texture (Pointer, 2006; Eugène, 2008), remaining the least studied one among those (Anderson, 2011). Although the color information incident on the human retina encodes important information about the objects and materials, overall sensation also depends *"on the appearance of that colour due to the relationship between the light transmitted, the light reflected, and the light scattered by the body of the object"*. (Pointer, 2003) Translucency is seen as a phenomenon *"between the extremes of complete transparency and complete opacity"* (*Eugène, 2008*).

doi: Received: May 26, 2021 ISSN 1534–7362 © 20?? ARVO

According to the "ASTM E284-17 Standard Terminology of Appearance" (2017) translucency is *"the property of a specimen by which it transmits light diffusely without permitting a clear view of objects beyond the specimen and not in contact with it."*.

The etymology of the term is related to the Latin words: "trans" (through) and "lux" (light) - implying light penetration inside the body of the material. (Kaltenbach, 2012) Translucent appearance is usually the result of a visual stimulus incident onto a retina from the objects permitting some degree of the *subsurface light transport*. Translucency is impacted by multiple intrinsic and extrinsic factors. The intrinsic factors are the physical parameters found in the *radiative transfer equation* (Chandrasekhar, 1960), such as the index of refraction, absorption and scattering coefficients, as well as the scattering phase function. They define how the light propagates through the media. A photon can get absorbed or scattered, i.e. redirected towards a different direction when there is a change in the index of refraction, either at the external surface of the object, or inside its volume (Tavel, 1999). How this passage of light through a material relates to a visual sensation of translucency remains unclear to date. The extrinsic factors include, but are not limited to, the illumination direction (Fleming & Bülthoff, 2005; Xiao et al., 2014), object shape (Fleming & Bülthoff, 2005; Gigilashvili, Thomas, et al., 2018) and the color of the surface a translucent object is placed on (Gigilashvili, Dubouchet, et al., 2020). The human visual system (HVS) is remarkably good at detecting subsurface light transport - we can easily tell the difference between a translucent glass and an opaque metal, translucent wax and opaque stone. We can distinguish translucent human skin from an opaque plastic dummy, translucent milk from opaque chalk. One of the fundamental problems is to understand how the HVS interprets the surface-reflected and subsurface-scattered light from the stimuli incident on the human retina. The exact visual and cognitive mechanisms of this ability are far from being fully understood. Since no model has yet been able to predict perceived translucency of a given material in an accurate and robust manner, translucency perception remains a topic of active research in academia and industry alike.

We would like to highlight that the primary focus of this article is *translucency*, not *transparency* - a better understood concept and visual attribute. While the two concepts are sometimes used interchangeably (e.g. *Merriam-Webster Dictionary* (n.d.)), it is usually accepted that *"transparent substances, unlike translucent ones, transmit light without diffusing it."* (Gerbino et al., 1990). According to the CIE, *"if it is possible to see an object through a material, then that material is said to be transparent. If it is possible to see only a "blurred" image through the material (due to some diffusion effect), then it has a certain degree of transparency and we can speak about translucency."* (Eugène, 2008) This implies that a given material might possess some degree of transparency and some degree of translucency at the same time.

The contribution of this article is threefold:

1. Summary of the state-of-the-art about the perception of translucency and the review of the recent developments in the field.

2. Discussion of the different applications that could benefit from the translucency perception research and overview of the importance of the topic in and across different disciplines.

3. Outline of the major knowledge gaps and research challenges in order to facilitate future work.

The manuscript is organized as follows. We briefly summarize the motivation for translucency perception research in the next section. In four subsequent sections we review the state-of-the-art and demonstrate the findings on the example of real and synthetic stimuli. Firstly, we provide a historical discourse on how the knowledge status has developed over time. Secondly, we overview the role of transparency in translucency perception. Thirdly, we discuss which factors impact perceived translucency. Fourthly, the potential cues for translucency perception are analyzed. Afterwards, we discuss the current challenges in the translucency perception research and outline the most important questions remaining open which is followed by a concluding section.

Background and Motivation

Translucency plays a significant role in a multitude of fields and applications. Thus, it is a research interest in different disciplines. In this section, we first overview the applications and the interdisciplinarity of the problem. Afterwards, we discuss the gap between the optical and the perceptual properties of a material - motivating the research from the human vision point of view.

Applications

In order to highlight the importance of understanding underlying visual mechanisms of translucency perception, below we will summarize the major applications where the translucency perception research can make impact.

A broad range of customer products look translucent, either customers expecting a translucent look from the products, or the degree of translucency itself can be an indicator of product's quality. This raises the need for studying translucency in the respective industries. For example, the foods, such as beer, meat and dairy products, are translucent. Therefore, translucent appearance plays an important role in the **food industry**, not only impacting customer satisfaction (Hutchings, 1977, 2011), but also contributing to the safety assurance (Chousalkar et al., 2010; Ray & Roberts, 2013). **Decorative paint manufacturing** is another example, as the hiding power of the colorants impacts the appearance and the overall quality of the paints (Krewinghaus, 1969; Midtfjord et al., 2018; Zhao & Berns, 2009).

Translucency has an implication for aesthetic purposes as well. Generation, reproduction and perception of translucent appearance has long been a topic of interest in **visual arts** and **cultural heritage**. Translucent building materials play an important role in the modern day **architecture** and are used to generate various visual effects of the exterior as well as interior **design** (Murray, 2013; Kaltenbach, 2012). The translucent look of a marble makes it an appealing material actively used both in architecture and **sculptures** (Barry, 2011), while translucency of glass is widely taken advantage of in the **glass art** (Kaltenbach, 2012). A special case is **painting** where the tradition of photorealistic depiction of the scenes exists from the medieval times and even though the scenes do not conform to the laws of physics, the artists still have been capable of generating vividly impressive and realistic depictions of the environment (Cavanagh, 2005), seemingly following the rule-of-thumb, heuristic "recipes" (Di Cicco, Wiersma, et al., 2020a). Recently, several studies have addressed perception of **painterly materials** (Zuijlen et al., 2020) with an emphasis on translucency in the **marine art** (Wijntjes et al., 2020) and **still life paintings** (Di Cicco, Wiersma, et al., 2020a; Di Cicco, Wijntjes, & Pont, 2020). Translucency is an important attribute for perception of visual realism and aesthetics of the artworks, especially those, depicting sea scenes, fruits and human skin. Understanding how painters generate the vivid sensation of translucency without conforming to the laws of physics can reveal interesting perceptual mechanisms of the HVS. This demonstrates that in addition to the physically-based simulations of the visual stimuli in computer graphics, translucency perception research can also greatly benefit from studying artworks, and vice-versa.

Translucent appearance is also actively studied in the **aesthetic medicine** and **cosmetology**. The interdisciplinary works in **material science** and **dentistry** emphasise the importance of proper translucent look of the dental implants and restorative materials (Liu et al., 2010; Wilson & Kent, 1971; Seghi et al., 1989; Lopes Filho et al., 2012; Anfe et al., 2008). On the other hand, face powders and moisturizers are used to enhance an appealing translucent look of the human skin (Giancola & Schlossman, 2015; Emmert, 1996), which can be studied by simulation of cosmetics and human skin rendering (Li et al., 2018) in **computer graphics**.

While **computer graphics** is often used as a tool for studying translucency perception (e.g. (Xiao et al., 2014; Urban et al., 2019)), perceiving translucency and accurate reproduction of translucent appearance is itself an important topic for the computer graphics community, especially when photorealism is at stake (Frisvad et al., 2020). One of the most significant, yet challenging, topics is accurate rendering of the human skin, which not only plays an essential role in the movies, video games and other segments of the entertainment industry, but also extends to the fields of computer vision (face detection and edge detection (Gkioulekas et al., 2015)), medicine and cosmetology (Igarashi et al., 2005; Li et al., 2018). While a considerable progress has been made in this direction, **skin rendering**, which inherently implies the accurate reproduction of translucent appearance, is a topic of active ongoing research (d'Eon & Irving, 2011) and remains especially challenging due to the multilayer nature of a human skin (Frisvad et al., 2020; Nunes et al., 2019).

One of the most novel fields which can benefit from translucency perception research is **3D printing**. 3D printing technologies have reached a level of development where translucency has become an important visual attribute, increasingly attracting an attention in the 3D printing community. The recent advances in multimaterial 3D printing enable generation and reproduction of material translucence by mixing transparent and colored opaque printing materials, which expands the appearance gamut of the 3D printing hardware (Brunton et al., 2018). However, object shape and scale dramatically impact perceived translucency, e.g. smaller objects transmit more light than the larger objects made of the identical material. In order to obtain a desired translucent look, mixing ratios of the printing materials should be adapted to these extrinsic factors, which itself needs a deeper understanding of the translucency perception process (Urban et al., 2019). A seminal contribution to this direction has been made by Urban et al. (2019) who proposed a hardware- and software-independent perceptual translucency metric for the 3D printing applications.

These fields might have established their own standards for measuring particular optical properties of the light permeable materials, such as scattering and extinction coefficients. However, the research on translucency perception is needed to understand how those objective measures can be used to predict what the customers will see. Moreover, the measurements are usually done for a small number of predefined shapes, conditions and geometries, which might not correspond to the real-life encounters and might generalize poorly. Therefore, it is important to know in what way customers' perception gets affected by the extrinsic factors, such as the shape of the object, illumination direction or motion. Understanding translucency perception and its contributing factors will make replication and matching of the total appearance easier. This will facilitate many appearance-related tasks, such as archiving and conservation in cultural heritage, as well as the development of the perception-aware rendering techniques in computer graphics.

Physics and Perception - the Gap

The primary reason why instrumental measurement of the perceptual translucency remains beyond reach is the fact that the definition of the perceptual attributes is vague (see sub-section *Inconsistent definition and conceptual ambiguity*) and their physical correlates are not identified. Even though the techniques of material property acquisition have advanced and the photorealism of the computer-generated imagery is impressive, the link between the measured physical properties of the materials and their visual appearance is far from being fully understood. Photosensitive measurement instruments might not be able to capture the appearance perceived by the HVS and cannot provide a quantitative correlate of visual sensation (SABIC Innovative colorXpress, n.d.). In other words, even if we achieve an accurate measurement, modelling and simulation of the optical properties of a given material, we might be able to create a "digital twin" of a real-world object, but we still will not be able to accurately predict how this material, either the real or the virtual, will look to the HVS - limiting our capability to generate desired visual effects from scratch and to replicate the appearance across different objects, scenes and conditions. This largely motivates the attempts of *soft metrology* and the rigorous research on visual appearance in different disciplines.

The knowledge gap is especially apparent when it comes to finding the correlation between the physical properties of subsurface light transport and the perception of material translucence. While there is a long tradition of research on colors - providing a reasonably deep understanding of color vision and color appearance, the perception of translucency has rarely been explored up until recently.

Indeed, translucency as an optical property of a material can be measured instrumentally (Pointer, 2003). The physical accuracy of rendering in computer graphics is constrained by the accuracy of the input physical material properties, dubbed as *"the input problem"* by Rushmeier (1995, 2008). This makes accurate measurement of the optical properties especially important. The most comprehensive and up-to-date survey regarding the acquisition of the optical properties of translucent materials has been done by Frisvad et al. (2020).

However, no technique has been proposed to date for an instrumental measurement of perceptual translucency. In other words, we have not been able *"to obtain numbers that are representative of the way objects and materials look"* (Hunter & Harold, 1987). Multiple application-specific instruments measure transmission-related visual attributes (BYK Gardner GmbH., n.d.), playing an important role in a broad range of industries, from solar cell manufacturing (Preston et al., 2013) to petroleum and edible product quality assurance (*Lovibond Tintometer*, n.d.). The two most common attributes studied in relation to translucency are **clarity** - *"defined in terms of the ability to perceive the fine detail of images through the material"*, and **haze** - *"defined as a property of the material whereby objects viewed through it appear to be reduced in contrast"* (Pointer, 2003). Haze is usually associated with a wide angle scattering (when the angle between the incident illumination and the transmitted light is more than 2.5 degrees, according to the *ASTM D 1003 - Standard Test Method for Haze and Luminous Transmittance of Transparent Plastics* (2003)) of light that causes blur and loss of contrast of the see-through image, while the clarity usually results from a narrow angle (less than 2.5 degrees) scattering. Analysis of the measurement procedures is beyond the scope of this paper, but it is important to highlight that no clear link between translucency as an appearance attribute, on the one hand, and clarity and haze, on the other hand, has been established. Pointer (2003) argues that *"the concept of translucency can perhaps be regarded as a descriptor of the combined effects defined above as clarity and haze. This implies that it is a more general term and, perhaps, should be limited to use as a subjective term, keeping clarity and haze as descriptors of objective, or measurable, correlates."* In the subsequent sections, we will analyze what we know and do not know about perceiving material translucence.

Historical discourse

Translucent appearance has long been encapsulated in a more general problem of visual appearance of objects and materials. The early theories of the visual appearance proposed that the HVS might invert optical processes in the scene to deduce the physical material properties and thus, the appearance (Pizlo, 2001; D'Zmura & Iverson, 1993; Poggio & Koch, 1985). Although this hypothesis is nowadays largely disputed (Fleming & Bülthoff, 2005; Chadwick et al., 2019), it remains debatable to what extent and complexity can we talk about "inverting" and estimating physical properties in the scene (Anderson, 2011). The later works proposed that the HVS might be using the heuristic low-level image cues and statistics (Fleming & Bülthoff, 2005; Motoyoshi et al., 2007; Motoyoshi, 2010; Chadwick & Kentridge, 2015) for assessing material properties, including translucency. According to the recent proposal by Fleming (2014), the HVS might be learning a generative model which predicts the variation of appearance across different natural illumination conditions. The recent developments in the material appearance research include unsupervised machine learning techniques to first predict human perception and then get deeper insight into it (Assen et al., 2020; Storrs & Fleming, 2020; Fleming & Storrs, 2019; Prokott & Fleming, 2019).

The fact that subsurface light transport plays an important role in visual appearance has been obvious from the very first attempts to measure appearance (Hunter & Harold, 1987). It has been important to understand how the light diminishes when passing through the thin layers of materials that either absorb or scatter light, for instance, when several layers of paint or coatings are applied on a given surface and how this affects the final color. Multiple models have been proposed in the first half of the twentieth century (using a term *turbid materials*). The Kubelka-Munk (KM) theory was one of the most widespread as well as simplest among those (Kubelka, 1931, 1948; Vargas & Niklasson, 1997). Kubelka-Munk coefficients K and S of a given paint film describe the portion of the light that gets absorbed and scattered, respectively, per unit thickness travelled through the paint material (Krewinghaus, 1969). Although it remains used for color matching calculations in the industries handling multilayered thin translucent materials, such as ink and dyed paper manufacturing (Yang & Kruse, 2004), its limitations are noteworthy - the model considers just two fluxes of light travelling upwards and downwards, and assumes that the light is not scattered laterally (Hašan et al., 2010) (although there have been attempts to extend it to the lateral light transport (Donner & Jensen, 2005)). Therefore, these kind of simplified models are not applicable to objects with complex geometry and subsurface light transport. Moreover, they might characterize material properties, but they are not suitable for characterization and prediction of translucency appearance.

Early attempts of studying visual perception of subsurface light transport were limited to perception of transparency - which, in some sense, was used as an umbrella term to describe light-transmissive materials. Proposed models consider a target transparent material as a thin filter which modulates the color of the background pattern seen through it and which can be described with a simple algebraic relationship (Metelli, 1970, 1974, 1985; Beck & Ivry, 1988; Gerbino et al., 1990; Gerbino, 1994). However, these models did not account for subsurface scattering. For details on perception, depiction and generation of transparency refer to the reviews in (Sayim & Cavanagh, 2011; Singh & Anderson, 2002a; Fleming & Bülthoff, 2005); regarding the perception of thick, complex-shaped transparent objects see the work by Fleming et al. (2011). Although relatively well-understood, transparency still remains a topic of active research (see (Faul & Ekroll, 2012; Falkenberg & Faul, 2019)). Object and background separation in transparent materials pose an important challenge in the ever emerging field of computer vision (Anderson, 2011).

While these works explain the perceptual mechanisms of see-through materials, the background is not always visible through the objects and the cues the HVS relies on for transparency perception are simply absent. This is especially true for the materials with high subsurface scattering, when none of the background can be detected through the object and the luminance gradient on its body is the only indicator that the light penetrates inside the volume. Many materials we interact with on a daily basis, such as wax, marble, textile, meat, cream or milk, are not see-through and cannot be approximated with the perceptual models of transparency. Therefore, the cues used by the HVS for perceiving translucency of the highly scattering media might be fundamentally different from those of transparency. This gave birth to the translucency perception research as a separate topic from transparency perception. The advances in the translucency perception research can be attributed to the rapid advance in computer graphics (Fleming & Bülthoff, 2005; Anderson, 2011). The difficulty to vary subsurface scattering properties systematically impeded generation of the proper visual stimulus datasets for conducting psychophysical experiments or analyzing image statistics. The progress in modelling subsurface scattering (such as Jensen et al. (2001)) made generation of translucent visual stimuli cheap, fast and fully controllable.

Koenderink & Doorn (2001) described in 2001 that the shading patterns differ dramatically between opaque and translucent media and that the "shape from shading" paradigm, which assumes Lambertian opaque surfaces, is not applicable to translucent objects. They raised an interesting question on how the HVS calculates the shape of the translucent objects and discussed an example of atmospheric objects, such as clouds, where shape judgement is entirely speculative. They used diluted and undiluted milk images to demonstrate how the radiance distribution over the material body depends on the mean free path of the photon (which is calculated as $\frac{1}{\alpha+\sigma}$, where α and σ are absorption and scattering coefficients, respectively). They also pointed out that the appearance of translucent objects varies with the point of observation as the number of photons emerging from an object body differs among different spacial positions. They also drew a parallel with the painters who are able to render a realistic appearance of translucent objects and argued that humans understand translucency in a qualitative way rather than by the means of calculating underlying physics.

This idea was later augmented by Fleming & Bülthoff (2005) in their seminal work, which paved the way for the last two decades' translucency perception research. They argued that instead of inverting optics, the HVS relies on the low-level image cues for calculating translucency. They examined and described different factors, such as object scale, color saturation, presence of specular reflections, potentially affecting perceived translucency. They identified that some regions, such as edges, contain richer information regarding material translucence. They demonstrated that translucency depends on the illumination geometry and back-lit objects look more translucent. Finally, they analyzed how the candidate image statistics, such as the moments of luminance histogram and intensities of the shadowed regions co-vary with the illumination geometry.

The intensities of the shadowed regions seem to be one of the most significant visual characteristics differentiating translucent and opaque materials. Motoyoshi (2010) proposed that the HVS might be calculating luminace statistics of the non-specular regions of the image to understand translucency. The author experimentally demonstrated that blurring and decreasing the contrast in the non-specular regions of the opaque material generates a translucent look.

Later works attempted to identify the impact of the various intrinsic and extrinsic factors on perceived translucency, such as the role of a scattering phase function (Gkioulekas et al., 2013; Xiao et al., 2014) and illumination direction (Xiao et al., 2014). Further works identified the spatial regions which are the most informative for understanding translucency (Nagai et al., 2013; Gkioulekas et al., 2015). Similarly to Fleming & Bülthoff (2005), Gkioulekas et al. (2015) also observed that edges contain a vital portion of the information about the subsurface light transport and discussed a potential use of the edge profiles as a physical correlate of translucency. Marlow et al. (2017) found that the lack of co-variance between the shape and shading information correlates well with the perceptual translucency. They demonstrated that illusory translucency can be evoked on optically opaque objects when the diffuse light field generates the shading which is not covariant with the surface geometry. The study has an interesting implication that translucency perception might be adjoined with the shape perception. The recent study by Chadwick et al. (2019) demonstrated that translucency perception is anatomically independent from color and texture perception.

The rapid development in the 3D printing technologies, which permit accurate generation of the physical objects with complex subsurface light transport properties (Brunton et al., 2018, 2020), on the one hand yielded an opportunity to use the physical objects instead of the computer-generated imagery in psychophysical experiments (Vu et al., 2016), and, on the other hand, increased an industrial demand on the translucency perception research (Urban et al., 2019; Gigilashvili, Urban, et al., 2019). Urban et al.

(2019) have recently proposed a perceptually uniform measure *Alpha* for 3D-printing applications, which can also account for an object scale. Gigilashvili, Thomas, et al. (2018, 2021) argued that when observing displayed images observers cannot enjoy the fully realistic experience they have on a daily basis when interacting with translucent materials. The authors believe that although having full control of the scene and the optical parameters, these kind of experiments might not reveal all behavioral patterns and thus, the visual mechanisms for translucency assessment. They used handcrafted physical objects (Thomas et al., 2018) for translucency assessment tasks and analyzed the behavioral patterns qualitatively. They observed that the dynamic cues, such as moving objects in relation with a textured background and head movements, as well as comparison of the given object's appearance between back-lit and front-lit illumination conditions, are used frequently by human observers while judging translucency. They also found that in addition to the appearance of a given object, the extrinsic cues elsewhere in the scene, such as caustics projected by an object onto a different surface, might also facilitate judgement of translucency (Gigilashvili, Dubouchet, et al., 2020). The advantages and disadvantages of using physical and digital stimuli are discussed later in the manuscript.

Translucency of see-through media

Transparency, translucency and opacity relate to the same phenomenon - the subsurface scattering of light (or the lack of thereof). The internal scattering gradually makes a perfectly transparent medium more translucent and eventually opaque (Gerardin et al., 2019; Gigilashvili, Thomas, et al., 2020). The boundary among them is fuzzy, implying that transparency and translucency are not mutually exclusive. Some degree of transparency and some degree of translucency can co-exist in the same stimulus. As noted above, translucent materials scatter light, while perfectly transparent ones do not (Gerbino et al., 1990). However, in some cases the light gets partly scattered and partly transmitted directly. If the amount of scattering is sufficiently low (as in the top row of Figure 8) or the object is sufficiently thin (as in the bottom row of Figure 10), the background is visible through a translucent object. In this case, the existing transparency models might, to some extent, contribute to the explanation of perceived translucency.

Internal scattering affects the clarity of the background image. Blur of the see-through image produces a translucent look (refer to Figure 1 and also Figure 19 (c) in (Singh & Anderson, 2002b)). It has been demonstrated that a change in the internal scattering produces a larger apparent translucency difference when the background is visible and blurred, than it does for highly scattering materials (Gigilashvili, Urban, et al., 2021). Singh & Anderson (2002a) extended transparency research to thin see-through filters that scatter light. Scattering blurs the image and usually decreases the contrast. In most cases the two parameters co-vary. The authors demonstrated that the blur alone decreases perceived transmittance when the Michelson contrast is fixed (Michelson contrast is defined as $(I_{max}-I_{min})/(I_{max}+I_{min})$, where I_{max} and I_{min} are maximum and minimum luminances, respectively (Legge et al., 1990)). Although they also found that the apparent contrast is smaller due to blur even if the Michelson contrast is kept constant, the decrease in perceived transmittance cannot be fully attributed to that. They propose that both blur and contrast of the transmitted image are the cues that increase the perception of opacity and decrease perceived translucency. A similar observation was made by Gigilashvili, Pedersen, & Hardeberg (2018) who studied blur from the image quality point of view and found that blurring removes the transmission cues and impairs translucency perception.

Visibility of the background through a medium is indicative of the subsurface light transport and can inform the HVS about translucency (e.g. see Figures 5, 8 and 10.). Seeing through a medium has been broadly studied in the context of transparency. The visual stimulus reaching a human retina through a transmissive material is a mixture of the contributions by the background and the transparent overlay. The HVS perceives the background as a single surface, even though the colors of the background in a plain view and those seen through a transparent medium might differ considerably. We somehow understand and estimate the properties of a transparent medium superimposed on a background. In order to infer transparency and distinguish transparent substances from opaque ones, the HVS relies on the regularities that exist between the colors of the background in a plain view and those seen through a transparent medium. Transparency is perceived when the lightness and chromatic compatibility exists between the overlay and the background. Modelling transparency perception has developed in two primary directions. Some works model transparency in a form of an *additive color mixture* (Metelli, 1970; Singh & Anderson, 2002b). Example of the additive model is the episcotister model by Metelli (1970, 1974, 1985). The idea of the episcotister

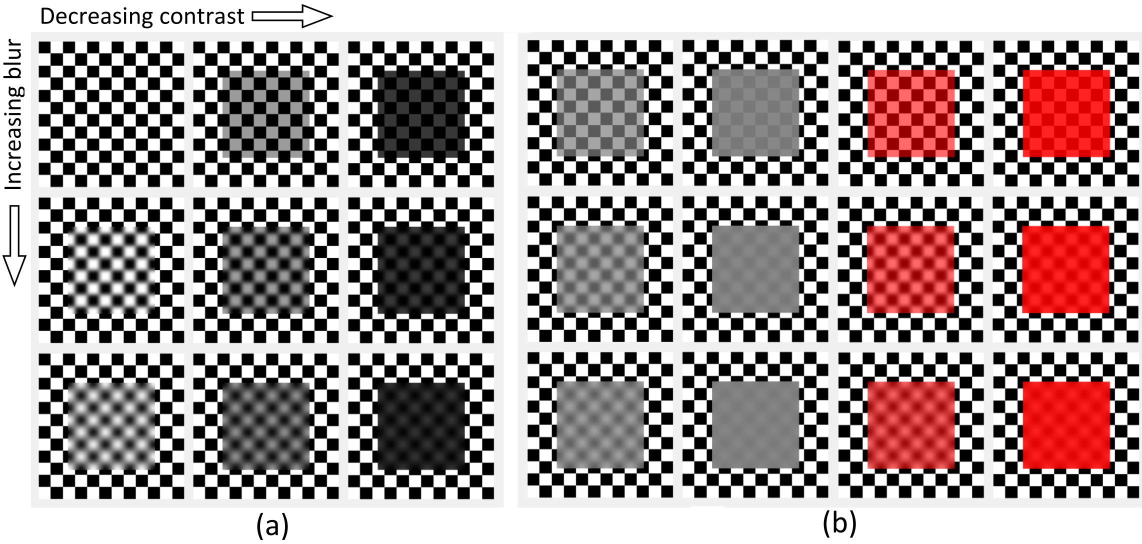

Figure 1: **(a)** A vivid impression of transparency and translucency has been produced by simple image manipulations. The contrast was reduced by decreasing lightness of the white patches of the checkerboard that yields an impression of an absorbing transparent filter with no reflection. Application of the Gaussian blur generated translucent look for all levels of contrast. Top row - no blur; Middle row - σ=12; Bottom row - σ=20. **(b)** In the left two columns, the contrast is decreased by decreasing the lightness of the white patches and increasing the lightness of the black patches, as if the filter had direct reflection. A convincing translucent appearance is generated, even when no blur is applied (top row). Translucency is stronger and more convincing with reflectance (additive) component than it was for the absorption-only scenario (in the top row, compare the rightmost image in **(a)**, and the second one from the left in **(b)**). The two rightmost columns demonstrate the chromatic case, where the hue shift also produces a hazy look and contributes to translucency appearance.

is the following: a disc with a sector cut out is rotating with high speed and is seen as a transparent overlay over an opaque background. The colors of the disc and the background simply add algebraically, and the proportions depend on the angle of the cut-out sector. While colors are mixed over time in Metelli's model, additions can happen spatially as well – for instance, an opaque mesh with small holes looks partly transmissive as a whole (Singh, 2020). The same principle has been later extended to the chromatic cases as well (D'Zmura et al., 1997; Hagedorn & D'Zmura, 2000). D'Zmura et al. (1997) studied the relation between colors at the background-overlay junctions and found that a shift in colors and change of the contrast are responsible for transparency perception. For instance, if the colors either converge to towards a point or are translated in the color space, they induce the percept of transparency, while rotations and shear do not lead to the same effect. Additive models approximate well the phenomena such as fog (Hagedorn & D'Zmura, 2000) or the media shown in the top row of Figure 8.

However, many transparent materials we encounter on a daily basis, such as glass, plastic or beverages, involve more complex optical phenomena. The transparency of the media similar to those shown in the bottom row of Figure 8 can be described with the *filter models*, which involve a *subtractive color mixing*. The filter models have been proposed both for the achromatic (Beck et al., 1984) as well as chromatic stimuli (Khang & Zaidi, 2002; Faul & Ekroll, 2002, 2011). This approach models the transparent overlay as an optical filter, which absorbs part of the light propagating through it, but also reflects some of the incident illumination at the vacuum-filter interface as per Fresnel equations. The color seen through the filter is a combination of the transmitted and reflected components.

There are two primary reasons why considering transparency perception models are also important for translucency:

First of all, it has been demonstrated that if particular regularities between background and transparent overlay colors are absent (D'Zmura et al., 1997; Faul & Ekroll, 2002), the filter is perceived opaque. Therefore, we believe that those kind of compatibility between the filter and background colors is also significant for translucency perception (it is worth noting that similar kind of chromatic compatibility is needed for gloss perception as well (Nishida et al., 2008)). The future work should reveal to what extent is the perception of translucency dependent on these regularities and whether translucency can be perceived in the cases when the filter and background colors are incompatible for inducing transparency perception (e.g. assuming fluorescence).

Secondly, a vivid perception of translucency can be evoked by transparent filters even in the absence of blur (i.e. if the contours in the background image remain undistorted). This means that when the background is visible, translucency can be observed even without any internal scattering. This can be ascribed to the decreased contrast and the color shift in the see-through image (Figure 1). If the transparent filter absorbs (subtractive color mixing) or reflects light (additive component), the contrast in the see-through image is reduced. Human observers are usually able to identify the additive component as a mirror reflection of the environment. Hence, the reflections from the surface usually evoke perception of gloss (as in the bottom row of Figure 8). However, Faul & Ekroll (2011) have demonstrated that specular reflections under uniform diffuse illumination evoke perception of translucency instead of gloss, proposedly because surface scattering is mistaken for volume scattering (see Figure 1 (b)). They also extended their prior work on filter models (Faul & Ekroll, 2002) and proposed an alternative parametrization of filter's physical properties – thickness, absorption and refractive index. They propose hue (H), saturation (S), transmittance (V) and clarity (C), to quantify the perceptual dimensions of transparency. The dimensions are related to the physical parameters; for instance, transmittance decreases exponentially with the filter thickness, and clarity is related to the index of refraction. Although the model does not account for subsurface scattering, V and C yield a broad range of appearances across the transparency-opacity continuum. The index of refraction determines the amount of the direct reflection from the surface. If it is equal to the refractive index of the immersing medium, no light is reflected at the interface, yielding the maximum clarity. However, a high reflection from the surface yields hazy translucent appearance (see Figure 1). A more perceptually uniform version of this space has been recently proposed by Faul (2017). The author made another interesting observation: the filter reflections and the resulting lack of clarity induce the perception of transparency and translucency when the luminance contrast in the background is large. However, the effect becomes weaker on low contrast backgrounds. For instance, if a homogeneous background was used instead of a checkerboard, the filters shown in Figure 1 would have appeared uniform opaque patches. Faul (2017) proposes motion as one of the factors for disambiguating this kind of stimuli. This and other factors contributing to apparent translucency or facilitating perception of translucency is discussed in the next section.

Factors impacting translucency

Translucency as a visual attribute is impacted by different intrinsic and extrinsic factors. Below we will overview the knowledge status on them.

Intrinsic parameters

Absorption and scattering coefficients

Wavelength-dependent absorption and scattering coefficients are fundamental parameters that describe the radiative transfer through a medium. Scattering (σ_s) and absorption (σ_a) coefficients signify the scattering and absorption events per unit distance traveled by a photon, respectively. The sum of the absorption and scattering coefficients is called extinction or attenuation coefficient (σ_t). The extinction coefficient σ_t is given as a sum of the scattering and absorption coefficients (σ_s+σ_a, respectively). σ_t for perfectly transparent material is equal to zero. High σ_a means that less photons escape the material and the object gets a darker shade; per contra, high σ_s is responsible for blurry and shiny appearance. It is worth mentioning that in addition to volume scattering (scattering inside the medium), a scattering event can also take place at the surface (will be discussed in the following sections). Xiao et al. (2014) demonstrated that the increase in the optical density of the materials affects translucent material matching in a monotonous and linear way under all illumination geometries. The effect of different absorption and scattering coefficients is shown in Figure 2.

Cunningham et al. (2007) studied aesthetic correlates of physical attributes and found that absorption and scattering are embedded onto a 1-dimensional manifold where they are significantly

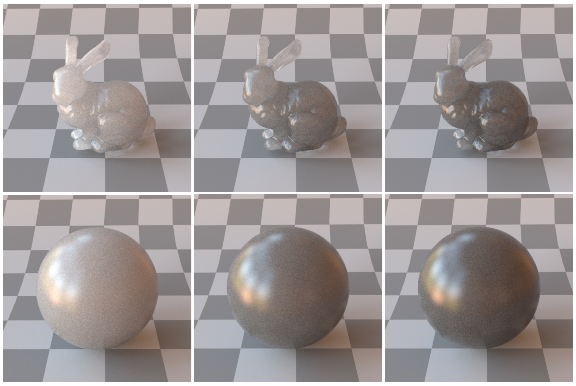

Figure 2: Objects in the same column are made of the identical material. However, due to smaller scale and presence of thin parts, the Bunny has more cues evoking perception of translucency. Objects in the first column have high scattering and low absorption. In the second column - lower scattering and higher absorption. In the third column - same scattering as in the second column - but higher absorption. How can we compare their perceptual translucency? Which of these six objects or materials are the most and the least translucent? [Reproduced from (Gigilashvili, Thomas, et al., 2020)]

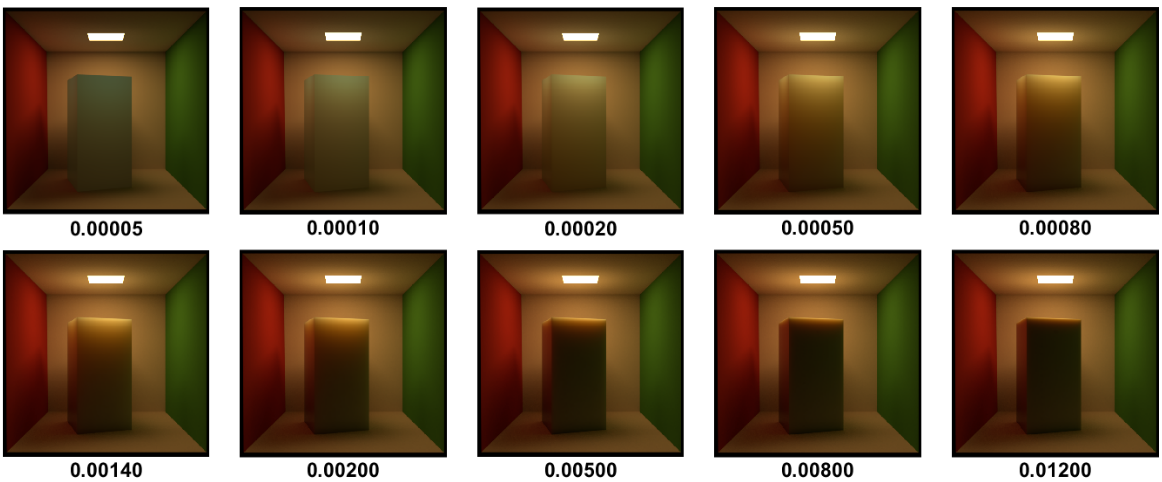

Figure 3: We rendered the box with skimmed milk optical properties as measured by Jensen et al. (2001) and implemented in Mitsuba (Jakob, 2010) in a Cornell Box (Niedenthal, 2002) (a broader variety of measured scattering properties can be found in the work by Narasimhan et al. (2006)). The optical density was varied with a scale parameter (shown below the image). It is apparent that the penetration depth decreases monotonically with the optical density. Therefore, only the edges are bright in the optically thick materials and the contrast with the rest of the object is large. On the other hand, photons spread easily through optically thin materials yielding relatively homogeneous luminance distribution.

correlated with the semantic labels of ”brightness” and ”blackness”. Koenderink & Doorn (2001) illustrated that materials with high mean free path look relatively uniformly shaded as the photons propagate through the material easily. On the other hand, if the mean free path is short, the penetration depth is shorter (Motoyoshi, 2010) and the radiant energy is visible near the edges on the side of the incident beam, while the rest remains relatively dark. This is illustrated in Figure 3. How intensity varies as a function of the distance from the surface, is illustrated in Figure 4.

Chadwick et al. (2018) demonstrate that although imperfectly, human observers are still able to unmix absorption and scattering in milky tea images. They tried to identify potential image cues used by observers and found that mean saturation explains well the variation in observer responses on the milkiness estimation task (which is accounted for scattering). On the other hand, value (V of the HSV) and the spatial saturation gradient were needed to explain the tea strength (absorption) responses. Interestingly, the cross-individual variation was large - different observers seemingly rely on different perceptual functions, or simply interpret the concepts differently. Urban et al. (2019) proposed a perceptually uniform translucency metric, which encapsulates the observation that the HVS is more sensitive to absorption-scattering differences in optically thin materials than in optically thick ones. The same was observed by Gigilashvili, Urban, et al. (2019, 2021). They found that if a material is nearly transparent, even a slight change in absorption and scattering coefficients is easily detected by humans, while larger steps are needed to notice the difference in more opaque materials.

Vu et al. (2016) observed that for textureless, flat thin 3D-printed shapes transmittance is more perceptually important than lateral light transport. They quantified the ratio of transparent and scattering white material in the mixture on a 255-level *gamma* scale, where low gamma corresponds to a higher portion of the scattering colorant and found that within the range 0-180, i.e. more than 70% of the physical parameter-space, transmittance was negligibly small (and perceptually opaque), while in the remaining range human observers were sensitive to colorant ratios, as the transmittance and the perceptual correlate were well explained with the Stevens' power law (Stevens, 1960).

Despite those attempts, the question on how exactly absorption and scattering coefficients contribute to perceptual correlate of translucency remains largely unresolved. One of the problems is that the perceptual dimensions of translucency are not known and the relation with transparency and opacity remains fuzzy. One of the recent attempts to structure translucency in a physical parameter space was made by Gerardin et al. (2019). They proposed a 3-dimensional translucency classification space for computer graphics - a cube where dimensions correspond to absorption, scattering and surface roughness. They claim that by increasing scattering, a transparent material gradually becomes translucent and then eventually opaque. However, by increasing absorption, a transparent material gradually becomes opaque, but never translucent.

Finally, the amount of the radiant energy that emerges from an object can be result of not only subsurface scattering (or surface reflection), but also emission as well (Tominaga et al., 2017). To the best of our knowledge, no study has investigated translucency perception on fluorescent materials. How well the HVS can separate the light emerging from a material into transmitted and emitted components, or whether we can tell the difference between translucent and fluorescent stimuli should be answered in the future.

Scattering Phase Function

Although the likelihood and the number of scattering events are essential, the direction a scattered photon is redirected to can also be important. If multiple scattering is assumed (Jensen et al., 2001), where diffuse approximation can be applicable, the impact might not be that strong. However, it can have a striking impact on the thin parts of the object, where only few scattering events take place (although in some cases, a phase function can impact thick parts too, refer to Figure 5).

Gkioulekas et al. (2013) have conducted a comprehensive study on the role of a phase function in translucent appearance. They argue that a similar translucent appearance can be yielded with the contrasting phase functions and conclude that a perceptual translucency space is composed of a lower number of dimensions than the physical parameter space. They generated a broad range of phase functions by linearly combining multiple Henyey-Greenstein and von Mises-Fisher lobes. Afterwards they conducted psychophysical experiments and came up with a 2-dimensional perceptual space of phase functions, where each dimension modulates diffusion - i.e. milky appearance and sharpness - i.e. glassy appearance, respectively. The contribution is significant for material design and has expanded the gamut of possible translucency, as many of the appearances would not have been re-

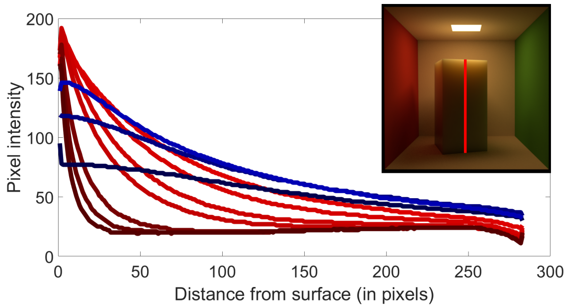

Figure 4: Image intensity as a function of the distance from the incidence surface. The cross-section where the intensities are measured is marked with a red strip in the top right corner. Optically thicker materials are shown in red (darker the shade, denser the material). The intensities are high at the boundary and they increase in the near vicinity, reaching local maxima - as proposed by Gkioulekas et al. (2015) (to be discussed later), then they monotonously decrease as the depth increases. Optically thin materials are shown in blue (lighter shade corresponds to thinner material), as they behave differently. They do not have a high intensity near the edge and the decrease slope is smaller. This supports the proposal by Koenderink & Doorn (2001).

Figure 5: The images vary in the phase function, while all other intrinsic properties are kept constant. Single lobe Henyey & Greenstein (1941) phase function is used with a varying value of *g*. The parameter *g*, is usually defined in the range of [-1 1], where negative values imply backwards scattering (back to the direction the light is incident from), positive values mean scattering forward, while 0 corresponds to the isotropic scattering. In the columns left to right *g* is equal to -0.9, -0.5, 0 (isotropic), 0.5 and 0.9, respectively. The top row is rendered in the front-lit illumination geometry, while the bottom row is back-lit. Because of the low optical density of the material, the direct transmission is high in the back-lit condition and the impact of the scattering directionality is negligible. The opposite is true for the front-lit condition. In case of back-scattering, more photons are redirected towards the camera, while the forward-scattering phase function redirects photons away from the camera. The appearance varies strikingly and ranges from almost Lambertian diffusive (due to high backwards scatter near the surface) to blurrier translucent looking (*g*=-0.5 and isotropic) and to darker, opaque-looking one. Please note that in case of forward scattering, thicker parts of the Bunny look more opaque, while thinner parts look translucent, as the forward scattering phase function facilitates transmission from the background towards the camera.

producible with a single lobe phase function. However, the robustness of the space is partially compromised in back-lit illumination geometry. Xiao et al. (2014) have extended the work and found that although the illumination direction usually affects the perceived magnitude of translucency, this impact is not significant for some phase functions. They found that phase function's location in the perceptual space (which was proposed by Gkioulekas et al. (2013)) defines whether an illumination direction impacts per-

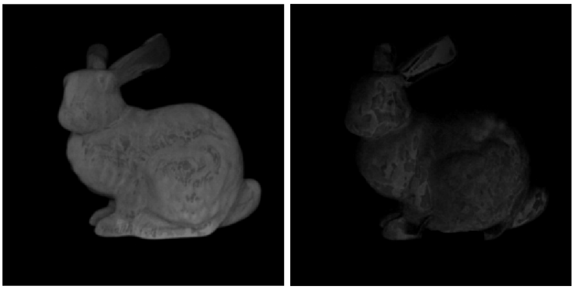

Figure 6: The intensity difference between the two extremes of the forward and backward scattering lobes in front-lit (left) and back-lit (right) illumination conditions. In the the front-lit condition, the difference is striking, while it is less apparent for the back-lit illumination condition. This can be accounted to the fact that due to low optical density of the material, direct transmission is high when illuminated from back and the scattered light accounts for a smaller portion of the resulting appearance.

ceived translucency. The similar correlation has been found between a phase function and translucency constancy (Xiao et al., 2014). The general trend is that the impact of lighting directionality is stronger for phase functions producing sharp glassy results than for more diffusing ones, which is intuitive - nearly isotropic phase functions which scatter light in all directions will be less affected than the ones that redirect photos strictly towards particular directions. Although Xiao et al. (2014) argue that the role of the phase function is also dependant on the object shape, the exact covariance between the shape and the impact of the phase function needs to be addressed in more detail.

Figure 5 illustrates a simple case of how the phase function alone can impact appearance, while all other parameters remain fixed. The images are rendered with a single lobe Henyey & Greenstein (1941) phase function, which takes a parameter g to define the directionality of the scattering. In the front-lit illumination geometry (top row), backward scattering resulted in brighter and more diffuse look, as the photons were scattered back towards the camera. On the contrary, forward scattering redirects photons away from the camera, resulting in dark opaque-looking appearance (although, note that thin parts look see-through, as the background reflections are forward-scattered *towards* the camera). On the other hand, the impact is negligible for the back-lit illumination condition (bottom row), because strong directional backlight results in high direct transmission and the magnitude of scattering change has weak impact on the resulting appearance. Figure 6 illustrates that the difference between the two extreme cases of the phase functions is striking for front-lit conditions (left image), while remains subtle for back-lit conditions (right image).

Index of refraction

The index of refraction is one of the most understudied intrinsic material properties in the context of translucency perception. At the boundary of the media, the difference between their refractive indices defines the angle the light ray is refracted with. Therefore, the refractive index has a strong impact on the background distortion in see-through images (proposedly also contributing to shape perception (Schlüter & Faul, 2019)). Fleming et al. (2011) have shown that humans are surprisingly good at estimating refractive indices of transparent materials - proposedly relying on a background distortion cue (although subject to biases due to object's thickness and distance to the background). Afterwards, Schlüter & Faul (2014) argued that instead of estimating an abstract refractive index, the HVS rather performs image-based matching where the both background distortion and the specular reflections are contributing. Regardless these attempts, the role of the refractive index in the appearance of non-see-through materials remains understudied. Additionally, difference in the refractive indices of the two bounding media modulates the magnitude of the Fresnel reflection and transmission, more refractive objects usually appearing glossier (Fleming et al., 2011; Schlüter & Faul, 2019) (also impacting caustics (Lynch et al., 2001; Kán & Kaufmann, 2012)). This is illustrated in Figure 7. While the subsurface scattering properties of a material remain constant, a high refractive index can render a mirror-like look and decrease perceived translucency (which is rooted in the decreased Fresnel transmission). If the difference

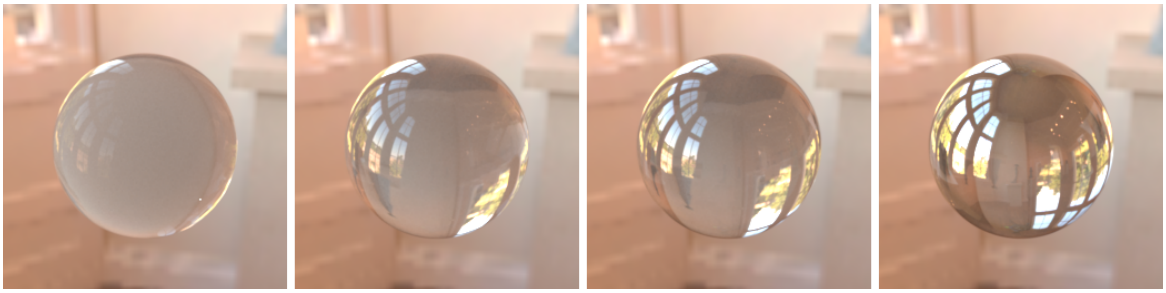

Figure 7: The only optical property that varies among the four images in the index of refraction (1.10, 1.33 (water), 1.50 (glass) and 2.41 (diamond), from left to right, respectively). A low refractive index ends in the lower Fresnel reflection and higher portion of the light penetrating the subsurface. Therefore, scattering in the subsurface is more apparent and the leftmost image looks more translucent. On the other hand, a high refractive index leads to higher reflection ratio and lower transmission, which yields glossy specular appearance rather than translucent one (refer to the rightmost image).

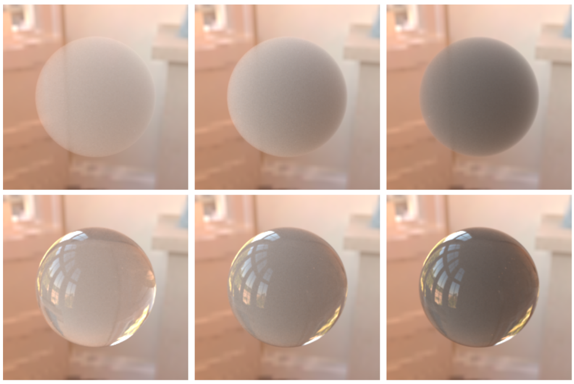

Figure 8: Glossiness is not essential for sensation of translucency. In the top row, the difference between the ambient vacuum and the object refractive indices is negligible, which results nearly no refraction and thus, no specular reflections. Despite the absence of the glossiness cues, the object still appears translucent, but the material looks more like a smoke or a sponge. In the bottom row, specular reflections are added, while the scattering properties inside the participating medium is identical to those of the top row. Material looks more glassy and more realistic, as the bottom row objects are more likely to be encountered in the real life than their top row counterparts. However, we cannot comment whether glossiness actually increases perceived magnitude of translucence.

Figure 9: The transmission image in the left photograph is upside-down which indicates that it is the result of the refraction through a convex lens. If we simply rotate the sphere upside-down, then the transmission image will look more like an opaque mirror-reflection. This was first demonstrated by Kim & Marlow (2016).

between the refractive indices of the bounding media is negligible, hardly any specular reflections are generated and a smokey-looking participating medium appears (see the top row in Figure 8 and compare with the bottom row in the same figure).

Observers' knowledge of the geometrical optics and the refraction phenomenon can facilitate distinction between the transparent media and mirror-like reflectors. While the convex lens refracts the light and transmission image is superimposed on the object upside-down, the convex mirror reflects the environment upright. Kim & Marlow (2016) have observed that rotating an image of a transparent sphere upside down creates an illusion of reflection, instead of transmission. This effect is illustrated in Figure 9.

The refractive index also determines the internal reflections (when the light is reflected backwards when it is trying to leave the translucent material), which impacts the amount of radiant energy emerging from the material - thus, also translucency cues. The extreme case is the total internal reflection – when the light traveling from a medium with a higher refractive index is fully reflected backwards – thus, no refraction happens and no light emerges from that medium to the medium with lower refractive index. The total internal reflection takes place when the angle of incidence is larger than the critical angle. Therefore, it is more likely to happen on complex surface geometries, rather than smoother ones. This could be one of the reasons for the appearance difference between the smooth and the complex Lucy shapes in Figure 13.

Marlow & Anderson (2021) have shown that if the illumination and observation angles nearly match, refraction can affect translucency, as the portion of the light exiting the material is reflected internally and is redirected towards the convex and away from the concave regions. The effect is relatively weaker when the difference between the indices of refraction of the bounding media is low and non-existent when the difference between the observation and illumination angles is large.

Finally, **polarization** of the incident light can also play a role in the Fresnel reflection and transmission. Gkioulekas et al. (2015) have used cross polarization photography to remove undesired specular reflections. They argue that specular reflections affect the location of the maxima and compromise the robustness of their radiance edge profiles for translucency prediction (to be discussed later). However, it might not be important for rough surfaces. Polarization is a broadly unexplored extrinsic property that deserves attention in translucency perception research.

Extrinsic factors

Object scale and structural thickness

If the object is enlarged, the distance a photon needs to travel increases. This means that for a given extinction coefficient, the number of absorption and scattering events goes up and less photons escape the material unscattered. The opposite is true, if the object is smaller. Therefore, object scale has an impact on the translucent appearance (Fleming & Bülthoff, 2005). This has serious consequences for 3D printing. Urban et al. (2019) have proposed *Alpha* - a psychophysics-based perceptually uniform translucency metric. However, the authors highlight that the metric should be scaled with the object size and provide a proper implementation of this. How object scale impacts appearance for a fixed optical material properties is illustrated in Figure 10 (also compare Bunny with a sphere in Figure 2). Photons need to travel a shorter distance at the edges - making them bright and thus, a characteristic cue for distinguishing translucent and opaque materials (Fleming & Bülthoff, 2005; Gkioulekas et al., 2015). Gkioulekas et al. (2015) have observed that the radiance profile at the edges are surprisingly robust and invariant towards illumination changes, making them a reliable "signature" for a material translucence.

Depending on the structural thickness, translucency appearance of a given object made of a homogeneous material can vary considerably. Refer to the Figure 11. While the torso of the bust usually looks darker and less see-through, the thin parts of the dress transmit more light in all illumination conditions and look especially shiny when back-lit. The same is true for the ears of the Bunny (Figure 5). It has been shown that presence of the thin parts can facilitate detection of translucency differences (Gigilashvili, Urban, et al., 2019, 2021), proposedly attributed to the fact that the HVS is more sensitive towards the changes in optically thin materials (Urban et al., 2019). This is further substantiated by Sawayama et al. (2019) who propose that a rugged surface of the object facilitates discrimination of translucency. Both findings indicate that the parts where a photon needs to travel the shortest distance contain the most information about material translucence. Also, materials with a heterogeneous structural thickness might overall look more translucent and less opaque when they have thin parts. This is true both for solid objects (Gigilashvili, Thomas, et al., 2018, 2021), as well as liquids (see the role of wavetips in sea paintings (Wijntjes et al., 2020)).

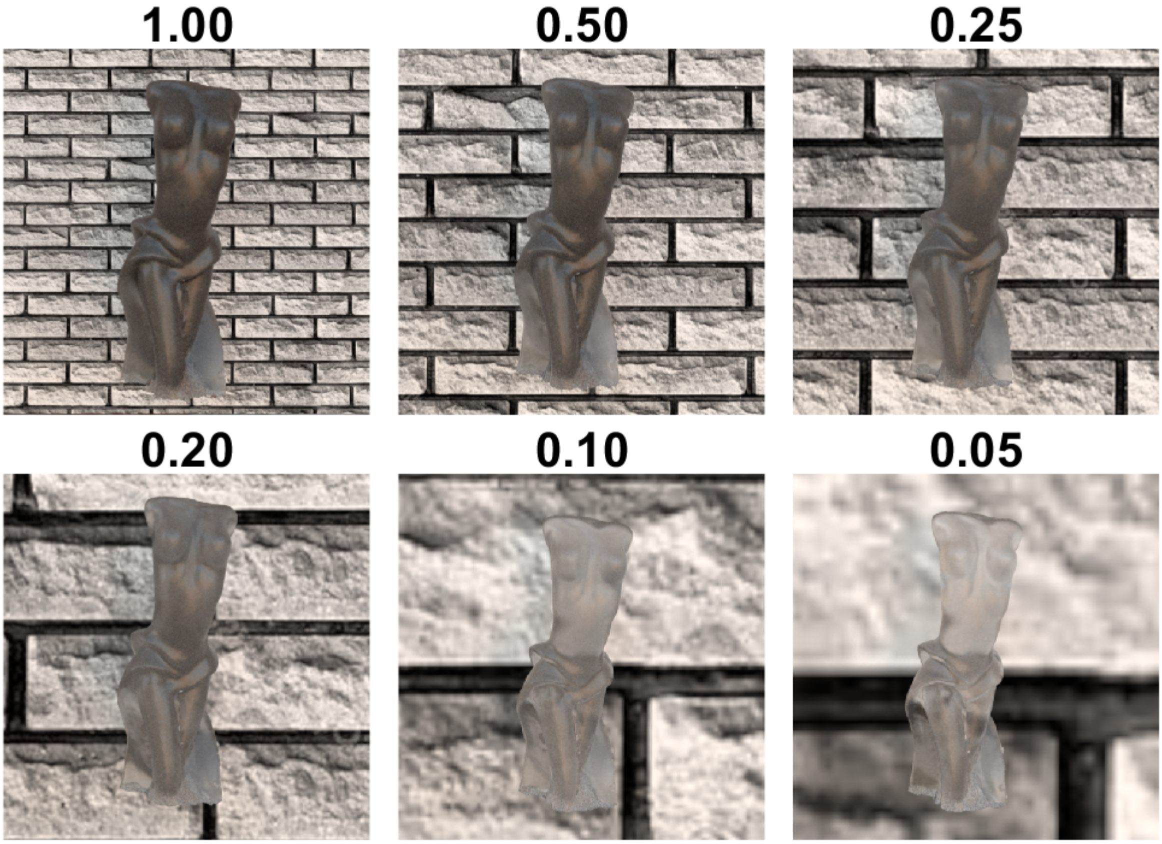

Figure 10: The illustration of how the object scale impacts perceived translucency of an object. While all six figures differ in scale, they have an identical shape and are made of an identical material. The smaller scale of an object means that a photon needs to travel a shorter distance to go through the material, i.e. for given scattering and absorption coefficients, the likelihood of scattering and absorption events decreases. This makes larger objects look more opaque and smaller ones look more light-transmissive. The numbers correspond to the scale relative to the top-left object. The background texture can also facilitate understanding the scale differences. We also illustrate that when the object scale varies, perceived translucency is also strongly impacted by the resolution of the image. If we put these six figures in a single scene, side-by-side (e.g. if we put the 0.05 version next to the original one in the 1.00 scene), smaller ones might look opaque, as the luminance variation will not be detected due to the contrast sensitivity limitations.

Surface roughness and geometry

Micro- and macro-scale surface geometry, although both scatter light, have qualitatively different effects on appearance. The microfacet-level surface roughness impacts refraction (Xiao et al., 2014), blurs the background image and evokes the perception of translucency even for the materials with zero subsurface absorption and scattering (Gigilashvili, Dubouchet, et al., 2020). It has been observed to be positively and monotonously correlated with translucency, when the transparency is seen as the other extreme (Gigilashvili, Dubouchet, et al., 2020). In the translucency classification system for computer graphics, proposed by Gerardin et al. (2019), surface roughness is one of the fundamental dimensions in the 3-dimensional parameter space. The authors argue that an increase in surface roughness makes transparent object translucent, but never opaque, because regardless the roughness of the surface, some photons still manage to go through (if the material has large mean free path). This phenomenon is shown in Figure 12.

According to the literature, translucency can impact perceived

Figure 11: The three frames are taken from a video. Refer to **Supplementary Material 1** for the video. The object is identical, however, the illumination geometry varies from back-lit (left) to side-lit (middle) and front-lit (right). The video provides a vivid illustration of how the perceived translucency changes with the change of the illumination directions. Moreover, it demonstrates that motion facilitates perceiving material translucence. Finally, the object shape enables us to observe how the presence of the thin parts provides additional cues about the light transmission properties of a material.

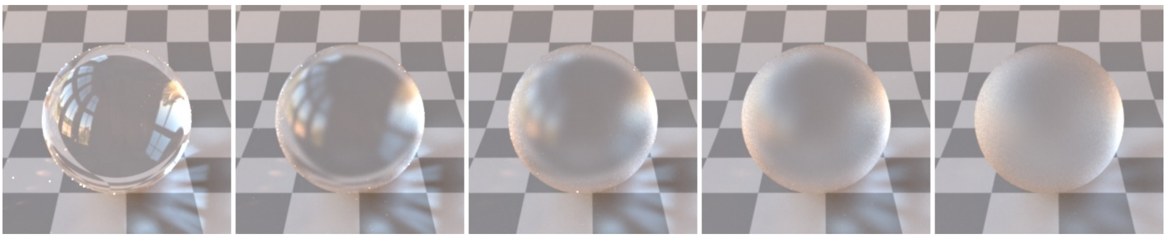

Figure 12: In addition to the subsurface scattering, surface scattering also blurs the background and generates translucent appearance. Sharpness of the specular highlights provide a strong cue for estimating surface scattering (Pellacini et al., 2000; Thomas et al., 2017). However, when the surface scattering is high, estimating subsurface scattering properties becomes increasingly difficult (e.g. see the right image: can we tell whether a subsurface is composed of a transparent or scattering material?). The root mean square (RMS) slope of microfacets equals to 0, 0.05, 0.10, 0.15 and 0.25, from left to right, respectively.

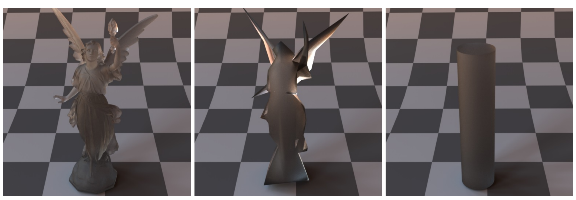

Figure 13: All three objects are made of the identical material. The Lucy (left) has the sharpest edges, while the sharpness and the surface curvature decreases gradually for low resolution Lucy (middle) and a cylinder (right). However, it is difficult to speculate which one yields the most vivid perception of translucency.

macro-scale surface geometry of the object - translucent objects appearing less sharp (Chowdhury et al., 2017). Interestingly, Xiao et al. (2020) have found the correlation the other way round too - experimenting with different levels of surface relief and claiming that presence of sharp edges make materials appear less translucent. They partially attribute this to the local contrast generated by the shadows due to high surface reliefs. However, the surface relief on a relatively flat surface is a tiny subset of the potential surface geometries which yield sharp edges. For instance, refer to Figure 13. The Lucy (on the left) has the sharpest edges and the most fine details; the low resolution Lucy (Gigilashvili, Shi, et al., 2021) (a smoother version of Lucy with a smaller number of vertices) has fewer and less sharp edges; while the cylinder is the least sharp among the three. All three objects are made of an identical material. If the proposal by Xiao et al. (2020) generalizes well to all geometries, then the ranking from the most translucent to the least translucent should be the following: a cylinder, low resolution Lucy, high resolution Lucy. It is difficult to claim the latter definitively. On the other hand, we can even speculate that the thin edges of Lucy make it appear more translucent (see the section above), its complex surface geometry causes more blur, while other shapes are structurally thicker, flatter, more specular and less blurry. In an earlier work, Xiao et al. (2014) also argue that complex shapes (e.g. presence of thin and thick parts) generate a greater range of translucency cues and lead to the faster failure of the translucency constancy.

Finally, a complex surface geometry might generate more specular highlights, caustics and interreflections - making more difficult to see-through and yielding illusion of subsurface scattering (Gkioulekas et al., 2015). Think of a transparent glass vase that is shiny, due to its complex shape, and looks as if it scattered light under the surface (see more on this in (Todd & Norman, 2019)). This phenomenon is illustrated in Figure 14. The sphere and the Lucy are made of an identical material. However, the low curvature and the simple shape of the sphere permits seeing-through it (it looks transparent), while the light scatters on the surface of Lucy and hence, it looks more translucent and less see-through.

Illumination direction

Illumination direction has one of the most striking effects on the magnitude of perceived translucency. If you have ever taken your food and looked through it towards the sunlight, you should have noticed that it starts glaring (see Figure 15). This effect can be taken advantage of in art and architecture. Also refer to Figure 11 which illustrates the frames from the video (refer to **Supplementary Material 1** for the video). Even though the material is identical, the difference in perceived translucency is apparent among the three conditions (compare left, middle and right images in Figure 11). Koenderink & Doorn (2001) have argued that translucency is viewpoint-dependent and *"transillumination" of the light* through the material is a strong cue for translucency. Most of the materials look more translucent when the light source and the observer are located in different hemispheres - i.e. when a sample is back-lit from the observer's viewpoint. This effect was first illustrated by Fleming and Bülthoff and has been further substantiated experimentally by Xiao et al. (2014), who observed that most materials look more translucent when back-lit and material matching is easier in back-lit conditions than in the front-lit one. Interestingly, Fleming & Bülthoff (2005) report that the information is not diagnostic enough for material discrimination when they are front-lit. This observation is however challenged by Xiao et al. (2014), who argue that this can be attributed to using a simplistic torus shape by the authors, while in Xiao et al. (2014) experiments with the complex shape of the Stanford Lucy enabled discriminating materials even in the front-lit conditions. Gigilashvili, Thomas, et al. (2018, 2021) have observed that humans prefer a back-lit condition for assessing material translucence. They argue that the magnitude of the differences between translucent and opaque materials is larger in back-lit condition - making it a desired geometry for comparing objects. Per contra, in the study of the dental porcelain translucencies (Liu et al., 2010), authors argue that sensitivity towards translucency differences does not differ significantly between front-lit and back-lit illumination conditions. However, the noticeability thresholds are lower for back-lit conditions (with *p-value* $\approx$ *0.06)* It has been also observed that textiles that normally look opaque might look translucent when back-lit (Gigilashvili, Mirjalili, & Hardeberg, 2019) - having implications for clothing and curtain manufacturing. Gkioulekas et al. (2013) noted that the illumination direction has the strongest effect on the appearance space where they embed different phase functions. As noted above (Figure 5), the parameter of the Henyey-Greenstein phase function has the weaker effect under the back-lit illumination condition (compare the top and bottom rows). On the other hand, Marlow & Anderson (2021) observed that the intensities produced by sub-

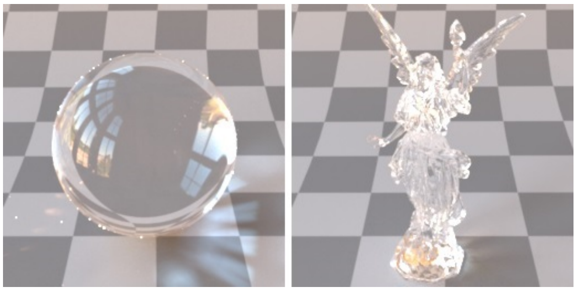

Figure 14: The sphere and the Lucy are made of the identical material. However, while the simple surface geometry and the low local curvature enable observing transmittance image through a sphere, the complex surface geometry and the high local curvature of the Lucy result in more specular reflections, interreflections and caustics. Eventually, although the extinction coefficient of Lucy is 0, its surface geometry makes it impossible to separate surface scattering from subsurface scattering highlights, evoking feel of translucency rather than transparency. This is especially apparent in tonemapped low dynamic range images, such as those.

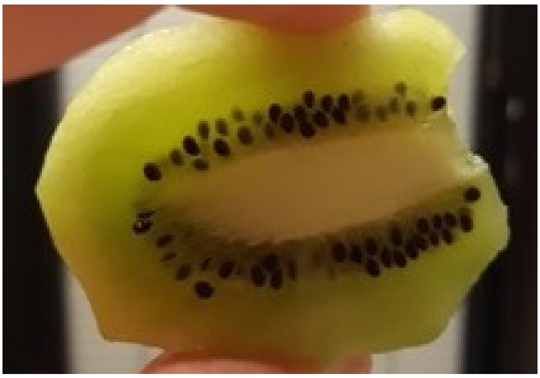

Figure 15: Most fruits look translucent when seen in back-lit illumination geometry. Bright edges and the luminance gradient indicate that the flesh is translucent, while the seeds look solid opaque black.

surface scattering remain relatively stable when the observer and the light source remain in the same hemisphere and the illumination angle changes from orthogonal to low grazing angles.

Illumination structure

The impact of illumination structure on the perception of translucency is not well explored. Although Xiao et al. (2014) argue that it is important to study translucency in the natural complex illumination and not under simplistic point light sources, as in (Fleming & Bülthoff, 2005; Motoyoshi, 2010; Nagai et al., 2013). Intuitively, a collimated beam should penetrate deeper than the diffuse ambient light inside the material and thus, is expected to generate higher magnitude of translucency. This was illustrated by Gigilashvili, Mirjalili, & Hardeberg (2019). They observed that textile samples were considered opaque under diffuse illumination, while some of them were re-classified as translucent when a high-luminance directional lamp was introduced in the scene. Presence of the shadows, which are thought to be one of the most important cues for assessing translucency (discussed later), also depend on the illumination structure. For instance, in case of a directional light, the only way shadowed and concave regions can get light is via subsurface scattering, while in case of diffuse and more natural illumination, shadowed regions can receive light also from the ambience, which can impact how translucency or opacity of the material is interpreted (Fleming & Bülthoff, 2005; Xiao et al., 2014; Nagai et al., 2013) (also see white diffuse front-lit and translucent

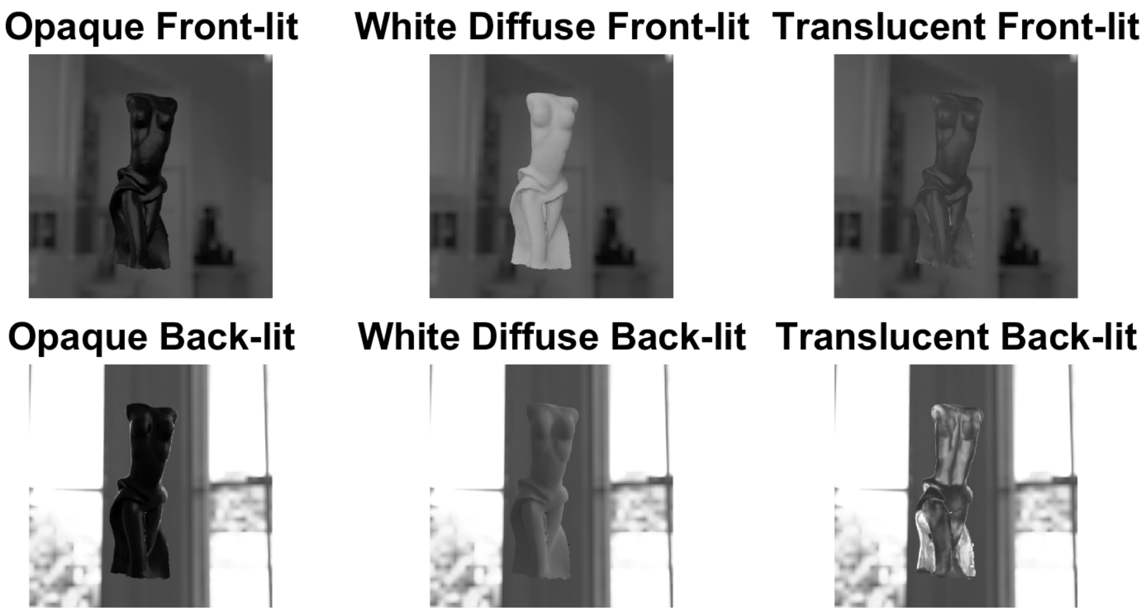

Figure 16: The way the illumination geometry modulates object appearance differs strikingly between translucent, black somewhat specular opaque and white Lambertian objects.

front-lit in Figure 16). Motoyoshi (2010) argues that sometimes it can be difficult to understand whether the blurry appearance is a result of subsurface scattering or diffuse illumination. Marlow et al. (2017) argue that for distinguishing translucency and opacity the HVS uses the covariance between surface and shading. If the surface and shading do not co-vary and the regions which were expected to be shadowed look lighter, a sensation of translucency is generated. They illustrated that if embedded in a proper light field which generates or eliminates this covariance, it is possible to render an illusory translucency on optically opaque media and the other way round. However, it is important to explore how often this can be encountered in the natural conditions. Fleming et al. (2003) observed that matching accuracy of the surface reflectance properties decreases under non-realistic illumination, and the random patterns of illumination might not generate glossy appearance at all. Similar phenomena could potentially be true for translucency. However, gloss has been shown to be less dependent on illumination than observed by Fleming et al. (2003), when complex shapes are used and the Fresnel effects are accounted for (Faul, 2019).

Caustics

While all previous research (e.g. (Fleming & Bülthoff, 2005; Motoyoshi, 2010; Marlow et al., 2017)) attempted to identify translucency cues on the object body proper, Gigilashvili, Thomas, et al. (2018, 2021) noticed that for assessing translucency, human observers put an importance on the cues elsewhere in the scene - primarily, the caustic patterns that are cast by an object onto another surface. The shadows cast by translucent and opaque objects differ (compare the top and the bottom rows in Figure 17). In some particular scenarios, caustics might be the only indicator of translucency, while from the object body alone, it might be impossible to infer that (compare the middle object in the bottom row between the left and right images of the Figure 17). Gigilashvili, Dubouchet, et al. (2020) have shown experimentally that placing an object on a black surface and eliminating the caustic pattern cast onto that decreases perceived magnitude of translucency. Whether this is solely attributed to the absence of caustics, or the impact of the ambient surface on the overall luminance of the object also contributes to that effect should be the topic of the future research.

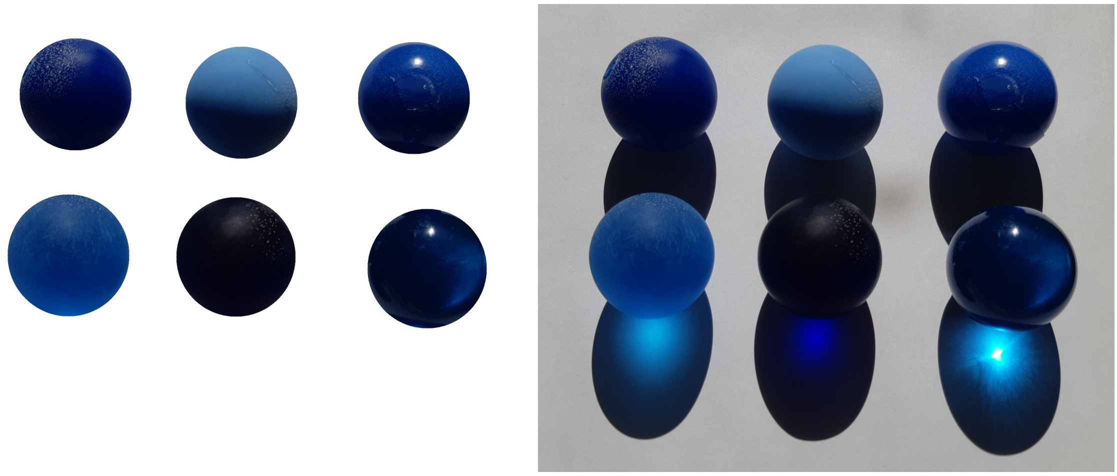

Figure 17: While the object body might look fully opaque (e.g. middle object in the bottom row), the external caustics provide rich information about the subsurface light transport properties of a material. The figure is reproduced from Gigilashvili, Mirjalili, & Hardeberg (2019). The animated version can be seen in **Supplementary Material 2.**

High-level cognitive understanding

It has been shown that appearance perception is not a one-way pipeline but rather a loop where the low-level vision is not simply an input for mid- and high-level vision, but also gets impacted by them (Anderson, 2011; Bartleson, 1960). Chadwick et al. (2018, 2019) made two interesting observations: although translucency perception is anatomically independent from color perception, observers with normal vision are still better at judging translucency in color images than in their grayscale counterparts - which could potentially be attributed to the easier identification of the familiar materials; secondly, people estimate absorption and scattering properties better in the stimuli existing in reality than in synthetic, virtual materials - proposedly attributed to a better training and experience with interacting with the real materials. Prior experience might be a considerable factor when assessing translucency. For instance, Liu et al. (2010) studied translucency perception of dental porcelains and found that the experts with *"more than 10 years of shade-matching experience"* discriminate levels of translucency better than novice students. On the other hand, Motoyoshi (2010) reported that there has been no difference between the observers who had seen and who had not seen the experimental stimuli before the experiment. Nagai et al. (2013) observed cross-individual differences in translucency cues. The authors used psychophysical reverse-correlation methods and found that different people looked at different regions of the objects to assess translucency, however, the exact reason remains unknown. The vast majority of the observers looked at the face of the Stanford Buddha shape used in the experiment, even though it might not have been the most informative region in terms of image statistics. According to the authors, this could be attributed to the fact that the human face catches attention easily (Hershler & Hochstein, 2005).

The high-level cognitive information seemingly plays a role in the perception of painterly translucency. It has been shown that depiction and perception of translucency is related with the perceived realism and "convincingness" of the artworks (Di Cicco, Wiersma, et al., 2020b; Wijntjes et al., 2020). Wijntjes et al. (2020) hypothesize that high-level cognitive factors might be contributing to perception of translucency in the sea paintings, such as a priori expectation that the water in the Caribbean scenes should be more transparent and translucent, than in the depictions of the non-tropical regions. Gigilashvili, Thomas, et al. (2018, 2021) noticed that observers try to identify materials when assessing their translucency and glossiness. Convincingness of translucency is enhanced with glossiness, proposedly due to the memory of familiar objects (Fleming & Bülthoff, 2005). Material perception has been shown to be a multimodal process relying on multisensory information (Spence, 2020). If material identification contributes to translucency perception, this opens up a new question, whether the senses other than vision play a role in the perception of translucency, either directly or indirectly. Marlow et al. (2017) have demonstrated that translucency perception to some extent implies understanding and estimating surface geometry. Additionally, when observing an object with varying thickness, we are able to perceive the object as made of a single, homogeneous material and not a composite of different materials, even though the luminance statistics and other translucency cues might differ considerably among these regions. All these observations indicate that translucency might not be dependent solely on the low-level vision cues but high level cognitive factors might be contributing to that as well. The fact that people, for instance, understand and use caustics (Figure 17) for inferring translucency, already involves a high level cognition of the scene. How much perceived magnitude of translucency is impacted by the high-level vision should be answered by future research.

Motion and scene dynamics

A fundamental problem in translucency perception is separating reflected and transmitted energy in the proximal stimulus on the retina. In this process, the HVS can obviously benefit from understanding the distal stimulus - the scene and the ambience. Motion has been demonstrated to be important for gloss perception and gloss constancy (Doerschner et al., 2011; Wendt et al., 2010; Hartung & Kersten, 2002), especially for separating specular highlights and surface texture. Additionally, motion can help understanding the object shape and geometry. On the other hand, the energy emerging from an object after subsurface scattering depends on the spatial location - making translucency viewpoint-dependent, as noted by Koenderink & Doorn (2001). Therefore, observing a translucent object from different viewing geometries should intuitively provide additional information about the bidirectional surface scattering reflectance distribution function (BSSRDF). Van Assen et al. (2018) have demonstrated that motion is important for perceiving viscosity and elasticity of translucent liquids and spreadable materials. Fleming (2014) hypothesizes that the HVS learns and predicts how appearance of a given material varies across different conditions, inherently implying motion in the learning process. Gigilashvili, Thomas, et al. (2018, 2021) analyzed human behavior when they were asked to assess translucency. They observed that humans frequently use motion-related cues - they move the fingers behind the object, move the object over a textured surface, move it relative to the light source and compare the object's appearance between front-lit and back-lit conditions - in short, it is natural for humans to change the background, observe how much it has impacted the appearance of an object and infer light transmission properties from it. This is qualitatively related to the phenomenon of change blindness in image quality - the change being more apparent when the subsequent frames are toggled back and forth, rather than being observed on independent occasions (Le Moan & Pedersen, 2017). Fleming (2014) hypothesizes that the brain might be building a statistical generative model of appearance that first learns and then predicts how appearance of a given material varies across different natural illumination conditions. If this hypothesis is true, interaction and dynamics would be an inherent part of the learning process from the infancy age. However, we do not know what part of it is learned and what is inherited.

The impact of motion is demonstrated in the video available in **Supplementary Material 1**. The video shows that motion relative to the illumination has a considerable impact on the luminance distribution on the object body and makes perception of translucency more convincing. Xiao et al. (2014) argue that motion might enhance material and translucency constancy. Intriguingly, although translucency constancy fails due to the illumination direction change, the continuous motion in the video (**Supplementary Material 1**) enables material constancy - we understand that it is the same material and its appearance changes due to the illumination, not due to the change in the optical properties of the material.

To the best of our knowledge, Tamura et al. (2018) have been the only ones to empirically study the role of scene dynamics on transmission perception. They found that the relative motion of the image superimposed on the object is significantly important for distinguishing reflective opaque mirrors from translucent glass materials. This once again highlights that still images might not be able to reveal the full range of the cues used by the HVS.

The role of other appearance attributes

Color

When talking about color, it is crucial not to mix up the chromatic and achromatic components. Lightness or brightness are directly correlated with the absorption and scattering, which make materials look darker or brighter, respectively (Cunningham et al., 2007; Koenderink & Doorn, 2001; Urban et al., 2019; Chadwick et al., 2018) (see Figure 2). As many translucent materials we interact with on a daily basis, such as milk, cream, cheese, and snow, have a whitish bright diffuse-looking appearance, Gigilashvili, Thomas, et al. (2021) have observed that many observers associate lightness with milkiness and translucency. However, lightness information is certainly subject to spatial and geometric constraints (Marlow et al., 2017). For instance, brighter edges (Fleming & Bülthoff, 2005; Gkioulekas et al., 2015; Di Cicco, Wiersma, et al., 2020b; Wijntjes et al., 2020) and shadowed areas (Motoyoshi, 2010; Fleming & Bülthoff, 2005; Marlow et al., 2017) are direct indicators of translucency. For completeness' sake, we should mention that Sawayama et al. (2019) found that a mean color difference between the images is, indeed, not informative enough to discriminate translucency.

On the other hand, little is known how chromaticity contributes to translucency. Chadwick et al. (2019) have worked with an observer who has a color-deficiency of a cortical origin. They demonstrated that color and translucency processing happens in the different parts of the brain and thus, are anatomically independent. However, they also observed (Chadwick et al., 2018, 2019) that the color normal observers perform better on color images rather than on grayscale ones - potentially explained by higher-level cognitive processing related to the material identification and realism. Di Cicco, Wijntjes, & Pont (2020) have recently shown that perceived translucency of painted citrus fruits is significantly correlated with their color saturation. Fleming & Bülthoff (2005) have illustrated that saturation might enhance the effect of translucency. Namely, if the saturation and lightness intensity are correlated positively, translucency looks like a warm glow, while it looks icy translucent in case of the negative correlation. This phenomenon is illustrated in Figure 18. Moreover, perception of wetness, which is optically related to translucency, has been also shown to be related with saturation (Sawayama et al., 2017). However, it is noteworthy that saturation alone cannot evoke perception of translucency. Besides, absorption and scattering coefficients of the most natural materials are wavelength-dependent, a phenomenon used extensively in 3D printing (Brunton et al., 2018, 2020) and art (Thomas et al., 2018; Gigilashvili, Thomas, et al., 2021). Therefore, the amount of the light emerging after the subsurface light transport will be dependent on the spectral power distribution of the illuminant. For instance, if the material which fully absorbs red wavelengths is illuminated with a red light, it might look opaque, not translucent. Although the effect might be negligible and rare under natural illumination, potential aesthetic effects generated with spectral translucence deserve future search and exploration.

Gloss

It has been shown that translucency impacts apparent gloss (Gigilashvili, Thomas, et al., 2019; Gigilashvili, Shi, et al., 2021). However, the correlation the other way round is not clear and straightforward. Moreover, Schmid et al. (2020) argue that the neural aspects of gloss perception should be addressed in the context of material identification, highlighting resemblance of the visual features between material recognition and glossiness perception. Schlüter & Faul (2019) argue that specular reflections have an important implication for perception of transparency. There are several indications in the literature that glossiness might be increasing perceived magnitude of translucency. This phenomenon has been observed by Motoyoshi (2010) (although no effect was observed by Nagai et al. (2013)). Furthermore, Yu et al. (2019) have proposed a highlight-generation method for rendering translucent appearance. While the primary intention was to enhance the perception of the fine details, interestingly, the perceived magnitude of translucency was also enhanced. Translucency and glossiness have been observed to be positively correlated in paintings (Di Cicco, Wiersma, et al., 2020b; Wijntjes et al., 2020; Di Cicco, Wijntjes, & Pont, 2020). Fleming & Bülthoff (2005) have observed that glossiness enhances the realism of translucent appearance, potentially attributing to the fact that many translucent materials are also glossy (e.g glass, marble, liquids), and we "expect" translucent objects to be glossy. However, translucency and gloss cannot alone explain each other, as many glossy materials, such as metals, are not translucent (Fleming et al., 2013; Tanaka & Horiuchi, 2015) and many translucent materials, such as smoke, cotton and textiles, are not glossy (Koenderink & Doorn, 2001). For instance, in Figure 8 the top row of the objects, which lack specular reflections to due a negligible change in the refractive index look smokey, or spongy, but still vividly translucent. The bottom row possesses the identical subsurface scattering properties, but adds specular reflections due to a large change in refractive index. We can argue that the bottom row looks more realistic and more likely

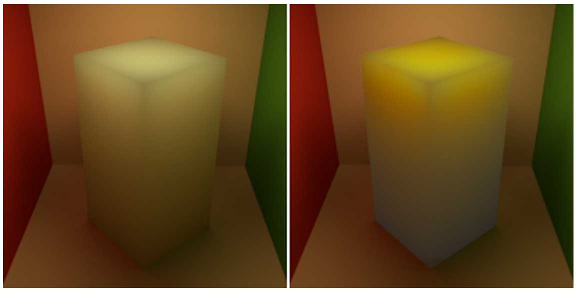

Figure 18: The mean saturation is equal in both images. However, the figure in the left image has a negative correlation between saturation and value (of HSV), while the right one has a positive correlation. This, as observed by Fleming & Bülthoff (2005), makes their translucence glow icier and warmer, respectively.

to be encountered in real life, but any estimation of the perceived magnitude of translucency, unless the difference between the refractive indices is large (see Figure 7), would be purely speculative. In some cases, the correlation between gloss and translucency can be straightforwardly negative, as the surface roughness, which decreases the magnitude of glossiness (Pellacini et al., 2000; Thomas et al., 2017) itself evokes the perception of translucency (refer to Figure 12). Moreover, the increase in the refractive index generates a stronger Fresnel reflection, i.e. stronger glossiness and less transmittance (Koenderink & Doorn, 2001), as illustrated in Figure 7. Finally, glossiness and specular highlights can facilitate understanding the shape (Todd & Norman, 2003; Norman et al., 2004; Fleming et al., 2004; Fleming & Bülthoff, 2005; Xiao et al., 2014; Marlow & Anderson, 2021). As the shape comprehension is proposedly related to translucency perception (Marlow et al., 2017; Marlow & Anderson, 2021), gloss can play a supplementary role in this manner too. Marlow & Anderson (2021) have recently identified co-variance between the intensity gradient produced by the sub-surface scattering and the shape of the specular reflections, both helping the recovery of the 3D shape and material properties. Moreover, they have experimentally shown that a light permeable surface covered with convex and concave regions is perceived more translucent when physically accurate specular reflections are superimposed. However, the effect is weakened or lost if the reflections are rotated and thus, incongruent with the subsurface scattering gradient.

Cues for translucency perception

Above-discussed intrinsic and extrinsic factors are impacting the proximal stimulus in a way that the HVS can deduce subsurface scattering and light transmission in the images. Whereas the scene dynamics and the temporal aspects enhance translucency detection, it is possible to perceive translucency from still images, which makes the researchers conclude that there should be some diagnostic features and statistics in the 2D images, which separate translucent media from the opaque ones. For example, it has been proposed that the skewness of the luminance histogram might be correlated with perceived gloss (Motoyoshi et al., 2007) (but see (Anderson & Kim, 2009; Kim et al., 2011)). There have been attempts to identify similar measures diagnostic for translucency and to propose at least partial models of translucency perception. Singh & Anderson (2002a) argued that in see-through scattering media, both apparent contrast and apparent blur of the background contribute to the perception of translucency. However, the cues on the objects which did not permit seeing a background through them, remained largely unexplored. Although no full model of translucency perception exists, and none is close being as complete as the Metelli-type models of transparency, several interesting observations have been made in the past 15 years which reveals some interesting characteristics of the translucency perception mechanisms. We overview these partial models and also provide some illustrations based on the *bust* renderings from the *Plastique Artwork Collection* (Thomas et al., 2018), which is rendered in the Mitsuba-embedded natural illumination (Jakob, 2010). Using this shape for the demonstrations has two practical implications: firstly, it has a varying degree of structural thickness, sharp edges and fine details - providing a broad range of translucency cues; secondly, a behavioral study has been conducted on the physical replica of this shape (Gigilashvili, Thomas, et al., 2018, 2021), which permits comparison of the real and synthetic stimuli in the future. We believe that this shape could become a standard for translucency perception research in parallel with Stanford Lucy (Stanford University Computer Graphics Laboratory, 1994). In the demonstrations below, we will mostly rely on a comparative analysis of six intensity images of the *Plastique bust* shape: a highly translucent material (referred to as *"translucent"*), highly absorbing somewhat specular black opaque material (*"black opaque"*) and a Lambertian-looking white diffuse opaque material (*"white diffuse"*) in back-lit and front-lit illumination conditions. These images are shown in Figure 16.

Fleming and Bülthoff

Fleming & Bülthoff (2005) were the first ones who tried to model the perception on the non-see-through scattering media. They have noticed that the intensity gradients differ between opaque and translucent objects, where the largest difference is noticeable near the edges. Bright and blurry edges are usually characteristic to translucent objects. Simplistic image manipulations by adding those features to a Lambertian surface using a high-pass filter enabled the authors to generate some degree of translucency, although not very realistic looking (see Figure 14 in (Fleming & Bülthoff, 2005)).

They also observed that the contrast between the specular and non-specular regions is smaller for translucent objects and on the example of a simple torus image, they demonstrated that the histogram of an opaque object is more skewed (for example, compare the first two columns in Figure 27). They observed that pixelwise-correlation between translucent and opaque images is far from linear and it alone cannot be a predictor for translucency. We have plotted how the intensity values change for each pixel of an identical material across two different conditions (see Figure 19) and between different materials under the same illumination (Figure 20). Similarly to Fleming and Bülthoff, we also observe that the correlation is not random, but highly non-linear (see Figures 19 and 20). For instance, when the illumination direction changes, the slope is steeper for a translucent object, while an opaque object intensities remain less impacted. The effect illumination geometry has on a pixel's intensity of a given material strongly depends on the spatial location of this pixel. We also noticed that in a back-lit illumination condition, the correlation between translucent object intensities and the opaque ones is mostly random, because of the high magnitude transmission component. On the other hand, in the front-lit condition, the non-specular spatial locations of a translucent material are lighter than their black opaque material counterparts, but darker than white diffuse ones (cf. captions of Figures 19 and 20).

Nevertheless, Fleming & Bülthoff (2005) were able to enhance translucency by applying a carefully selected "N-shaped" filter and to enhance opacity by applying a sigmoid filter to the intensity values. However, they note that this approach can only work when lighting is fixed and spatial correspondence between the pixels is unchanged. The authors illustrated isophotes - the contours of equal lightness and concluded that neither luminance distribution histogram, nor the spatial isophotes, can predict translucency alone, but it is rather more likely that the HVS relies on a combination of the luminance and spatial information. We came up with the qualitatively similar non-linear filters (shown in Figure 21) and tried to use them for making opaque objects translucent and translucent objects opaque (refer to Figures 22 and 23). We noticed that while the approach might work to some extent (especially, in the front-lit condition), it fails in the thin parts, especially in the back-lit condition. As Fleming and Bülthoff used a simple torus shape in their study, we tried the approach on a simple shape as well, such as a parallelepiped cube, which also produced considerable artifacts near the edges (refer to Figure 23). In the back-lit condition, the thin parts are bright, which is even further enhanced with a sigmoid filter. However, interestingly, unlike the front-lit condition, we do not see highlights in the thin areas as specular reflections. We somehow understand that they are a result of the subsurface light transport, which makes us believe that in addition to the low-level image cues, the higher order cognitive processes of the scene and geometry understanding also play a role in the translucency perception pipeline (refer to the captions of Figures 22 and 23).

Fleming and Bülthoff also demonstrated experimentally that back-lit objects look more translucent than the front-lit ones and tried to identify which image cues explain this psychophysical variation best. They found that neither pointwise-correlation, nor the first four moments of a luminace histogram (mean, variance, skew-

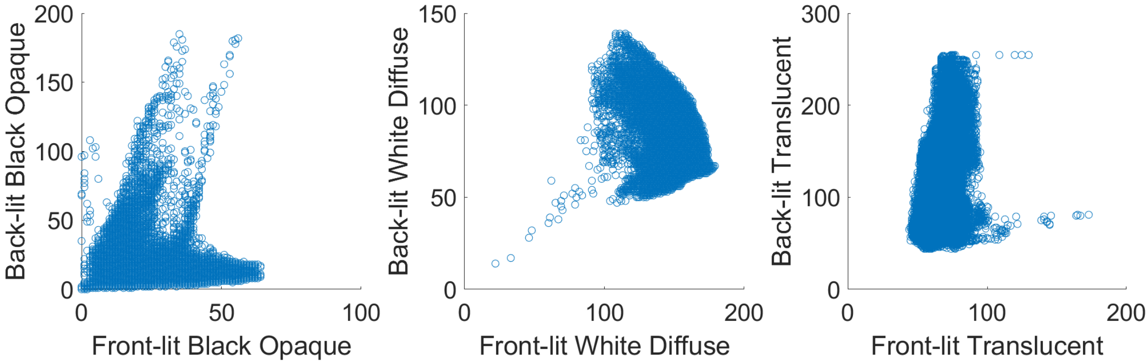

Figure 19: The correlation of the intensities between the identical pixels of the same object under two different lighting conditions. The plots show that although far from being linear, the dependence is not random. Most pixels of the black opaque object are simply darkened as they fall in the shadow when the light is moved on the back side. Some intensities, mostly on the edges, go up, because the backlight is not incident fully perpendicularly, and some of the light comes from the side angles as well. For the diffuse white object, the relationship is usually negative, except for some pixels on the edge, that brighten under back-light and are thus, positively correlated. For the translucent object, the slope is steep and the values simply go up when the object is placed under the back-light. We identified that the behavior of the pixel intensity is strongly dependant on its spatial location. However, the overall trend differs between the three objects and the change of pixel-wise intensities between the illumination geometries might to some extent indicate to the subsurface light transport.

ness and kurtosis) are predictors of translucency. In order to test this observation, we have rotated a bust figure with 180 degrees from back- to front-lit condition and visualized the summary statistics of a luminance histogram as a function of the rotation angle (refer to Figure 24). Similarly to Fleming and Bülthoff, we also noticed that they are non-monotonous, and while some trends can be identified, they are prone to bias due to the object shape and the distal stimuli in the scene composition, which makes them unlikely and unrobust cues for translucency perception (refer to the caption of Figure 24).

As neither histogram nor spatial information alone are enough for predicting translucency, we tried whether simple histogram matching between front-lit and back-lit conditions of the same translucent material could affect their appearance. Histogram matching affects the magnitude of intensities, but is also to some extent "spatially aware". The resulting images are shown in Figure 25. Although some artifacts were produced (e.g. near the edges), matching the front-lit object with its back-lit counterpart enhanced its opacity, as the high transmission pixels in a back-lit object, can be interpreted as specular reflections in the front-lit scenario. When the back-lit image was matched with the front-lit histogram, it started looking less transmissive, but still highly translucent.

Furthermore, Fleming & Bülthoff (2005) argue that the shadowed areas manifest the largest difference between the front-lit and back-lit scenarios and if subsurface scattering is the only way a photon could get to a bright region in the image (otherwise, it would have been in a shadow), that can be used as a cue to translucency. While this might be commonplace for directional lighting conditions, which renders sharp shadows on opaque objects, Xiao et al. (2014) argue that in diffuse and more natural light fields, the shadowed areas and surface concavities also receive light from the ways other than subsurface scattering. For instance, refer to Figure 26, which highlights the regions where a translucent object has a higher intensity than its opaque counterparts under the same illumination. A front-lit translucent object has larger intensities than its black opaque counterpart in nearly all regions apart from the specular reflections. However, under the same condition, that is not true for a white Lambertian-looking opaque object, which owes its larger intensities in shadowed areas to its highly scattering surface, direct front-side illumination and interreflections. Therefore, the lightness of the shadowed areas alone can not be an indicator of translucency either. On the other hand, it is worth noting that the intensity difference between back-lit and front-lit versions

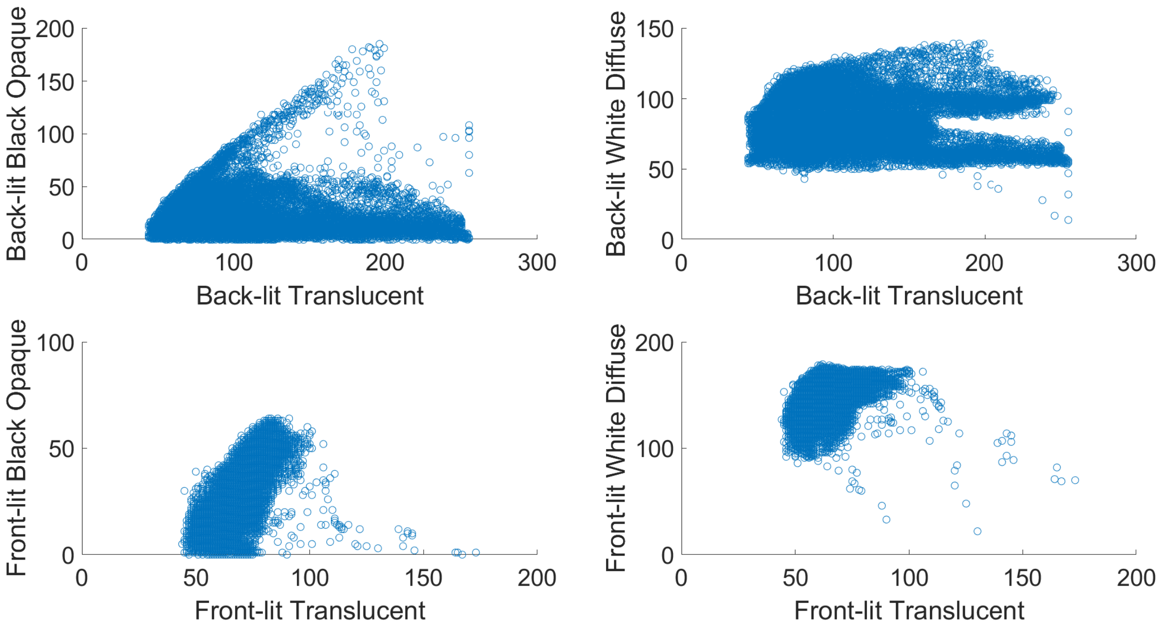

Figure 20: The figure illustrates pixel-wise correlation of the intensities under a given illumination geometry between a translucent object and other two non-transmissive materials. While the correlation looks mostly random under back-lit, it becomes more visible when objects are front-lit. The pixels become dimmer on the black opaque object, because its non-specular areas simply absorb light, while non-specular areas of a translucent object either back-scatter some of it, or transmit from a background towards the camera. The opposite is true for the white diffuse material, because more light gets scattered towards the camera by a white opaque object and no energy is lost due to the subsurface scattering away from the camera (the similar phenomenon was observed by Nagai et al. (2013).)

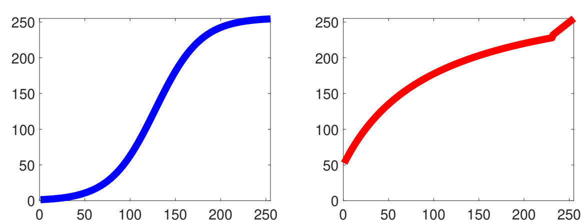

Figure 21: The sigmoid function (left) stretches low and high intensity values towards the extremes, which increases the overall luminance contrast. The "N-shaped" curve (right) scales up lower intensities, while keeps the highlights intact - decreasing the contrast between specular and shadowed areas. Fleming & Bülthoff (2005) observed that under fixed illumination conditions, similar functions can be used to enhance opacity and translucency, respectively.

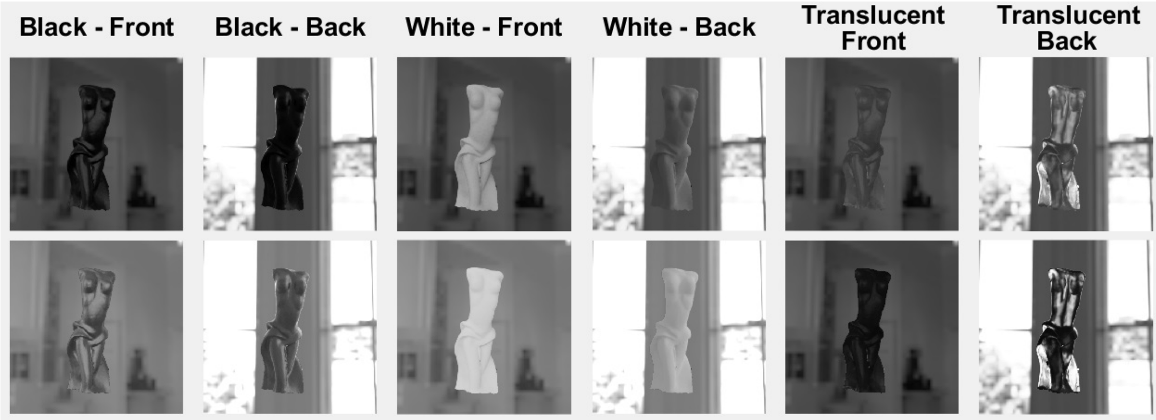

Figure 22: The non-linear functions (shown in Figure 21) applied to the intensities. We tried to make opaque objects more translucent with an "N-shaped" function, and a translucent object more opaque, with a sigmoid function. The top row illustrates the original image intensities and the bottom row shows the results after application of the non-linearity. The front-lit black opaque object has become slightly more translucent-looking as the contrast between the specular and non-specular regions has decreased. However, it can also be interpreted as an opaque object of a simply lighter shade. The back-lit black object does not look transmissive, but rather white diffuse material (compare with *White - Back* in the top row), because scaling up darker shades makes it more reflective but is unable to generate the transmission gradient similar to that of a translucent object (compare with rightmost images in both rows). The translucent look of a front lit white diffuse material has been considerably enhanced, because making shadowed areas lighter creates the feel that *"photons could not get there without subsurface scattering"*. Interestingly, under backlight, although looks less opaque, it does not process gradient characteristic for transmission or subsurface scattering either, making it look somewhat unrealistic. On the other hand, as the immediate background of the object is a wall, not the light source proper, its dim color can also be interpreted as a thin transparent filter. Finally, the front-lit translucent object became more opaque by eliminating the lighter shades in the non-specular areas and no cue has been left that could hint the HVS to the subsurface light transport. However, the approach failed in back-lit illumination geometry. Although the increased contrast between lights and darks make it look more solid, the highlights that are usually thought to be specular reflections, are transmission components in this case and scaling them up strengthens the perception of transmission. This was observed by Motoyoshi (2010), who noticed that in transparent materials and thin parts, the contrast is reversed or random. This illustrates that simple context-blind non-linear scaling does not control translucency-opacity appearance. In the rightmost image in the bottom row, we do not perceive highlights as specular reflections. We somehow understand that this is the result of light transmission. Therefore, the higher-level cognitive mechanisms of the scene and shape understanding seem to be involved in the translucency perception process.

of the same translucent material (also shown in Figure 26) could be one of the reasons why back-lit objects look more translucent than their front-lit versions. The authors conclude that the HVS relies on these kind of image cues rather than inverse optics. Indeed, *"there is simply not enough information available to invert the actual physics of image formation"*, as well-noted by Anderson (2011), but whether the HVS is completely unaware of the laws of physics, remains yet to be explored.

Motoyoshi

Motoyoshi (2010) has observed that specular regions remain relatively intact by the subsurface scattering and what varies across different levels of translucency is the appearance of the non-specular regions. Similarly to the earlier work (Fleming & Bülthoff, 2005), Motoyoshi noted that the non-specular regions usually get blurrier and lighter when subsurface scatter increases. Let's refer to Figure 27, which illustrates the absolute difference between translucent and opaque objects with different levels of specularity. This demonstration supports Motoyoshi's observation that the difference between highly translucent and highly opaque objects is minimal in the areas of specular reflections, which makes us conclude that the spatial regions diagnostic for translucency can be dependent on the surface roughness and the extent of specular coverage.

Furthermore, Motoyoshi also noted similarly to Fleming & Bülthoff (2005) that the pixel intensity correlation between translucent and opaque materials is far from linearity, but still not random. They separated the image into different spatial-frequency

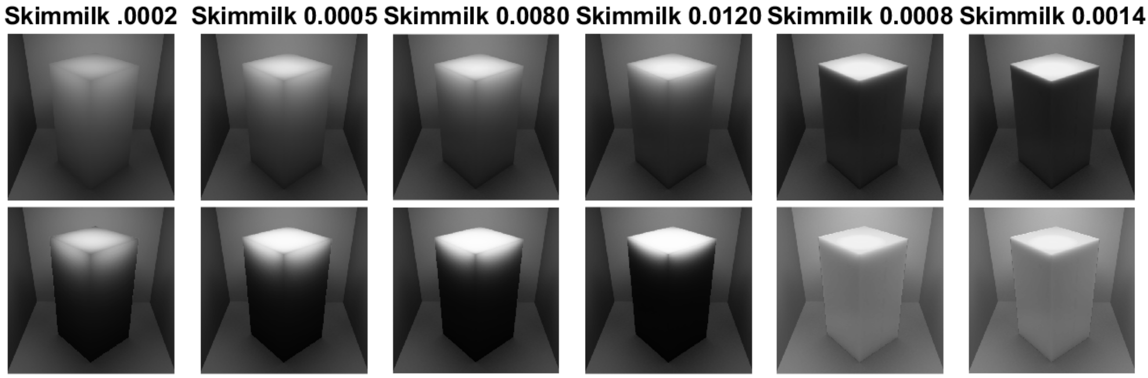

Figure 23: The nonlinearities identical to that in Figure 22 applied to simpler shaped objects. The top row illustrates the original intensities, while the bottom row is the non-linear filtered result. The contrast enhancement darkened the shades in the bottom of the box, where light penetration is little. However, it made the near-edge areas brighter, which in the first four columns, still produces somewhat unrealistic feel of translucency. The translucent look feels more and more unrealistic with the increase in the optical density (e.g. compare the first and the fourth images in the bottom row). On the other hand, an interesting result was produced by scaling up the lower intensities in nearly opaque objects (two columns on the right). As the side of the box that faces away from the illumination looks lighter, it overall evokes a feel of highly scattering bright material. This effect is stronger in column 5. Interestingly, keeping the original highlights on the side which directly faces the incident illumination made them look more like specular reflections.

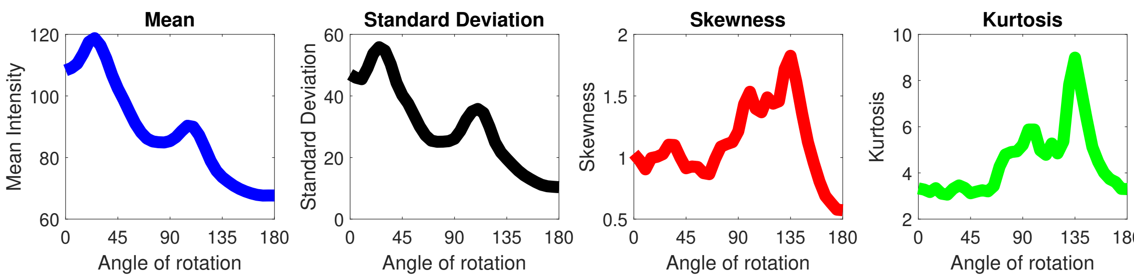

Figure 24: The back-lit bust figure is rotated with 180-degrees all the way to the front-lit condition. The plots show how the first four moments of the intensity histogram change as a function of the angle of rotation. In the original frame, the wall is the immediate background of the object. Once it is "flies" over the window, its mean intensity and standard deviation go up. They generally go down with the angle of rotation, but one local maxima is noticeable around 120 degrees, because there is another window in the scene, which once again "lights up" the object. The skewness and kurtosis have an apparent peak around 135 degrees. This is the result of a highlight produced by internal caustics. When the object is lit from the left, most of its body looks relatively darker, but a bright strip of the caustic pattern is created on its right side, as a result of photon accumulation. This is visible in the middle frame of Figure 11. This highlight generates the unexpected skew in the histogram. For the front-lit condition, the skewness and kurtosis drop dramatically, as the overall object looks blurry and more homogeneous. It is worth noting that these statistics are non-monotonous, and too dependent on the object shape and scene composition that makes their robustness as translucency cues questionable.

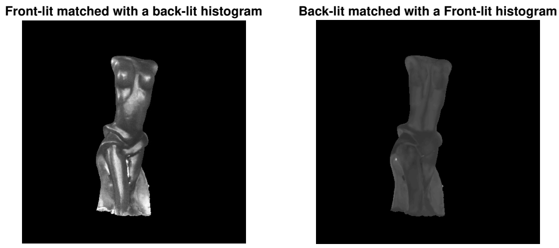

Figure 25: If we match the histogram of a front-lit translucent object with its back-lit counterpart, it starts looking opaque. This can be accounted to the highlights that were the result of trasnmission under the backlight but look more like specular reflections under the front-side illumination. However, some artifacts are still visible near the edges, which look unnaturally specular. On the other hand, due to the strong transmission gradient, we have not been able to produce an opaque look with a back-lit object, but blur in the highlight areas produces less transmissive but highly scattering look.

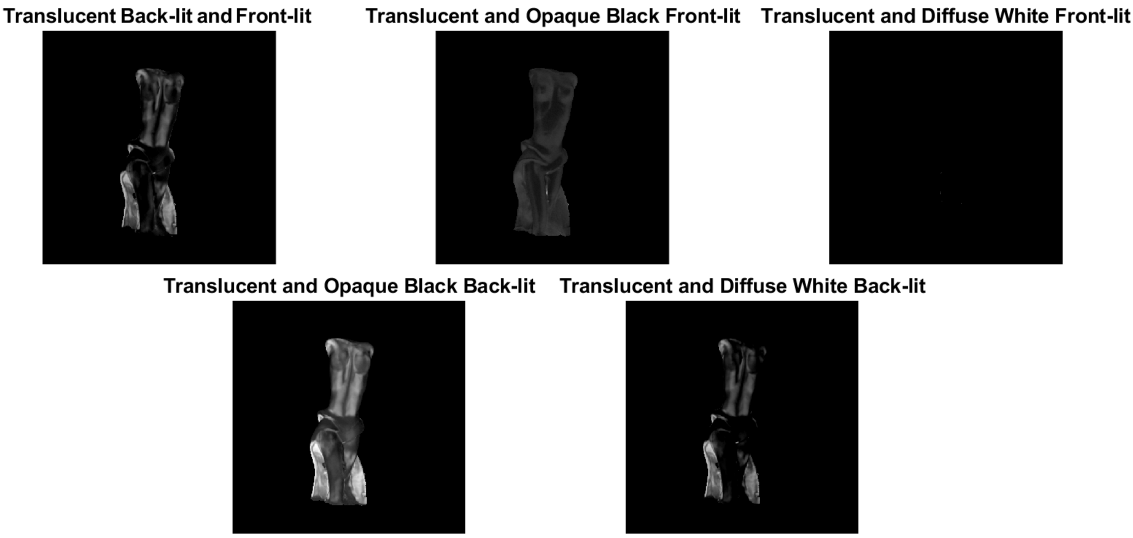

Figure 26: The intensity difference highlights the spatial regions where more energy emerges from a translucent object than from its opaque counterparts (note that this is not an absolute difference between the two images). The first image additionally shows the difference between back-lit and front-lit conditions. Under back-lit condition the intensity is higher in virtually every region when compared with a black opaque object, in thin and geometrically flatter areas when compared with diffuse white and itself under front-light. The front lit translucent object has higher intensity in non-specular regions only when compared with a black absorbing material, while none of its regions have higher intensity than a white front-lit Lambertian object.

sub-bands using a Gaussian band-pass filter. Afterwards, they manipulated and measured root mean square (RMS) contrast in each of the frequency domains. The relationship between the contrast in the non-specular regions and translucency is non-monotonous. At first, the RMS contrast decreases as we move from opacity towards translucency. However, for transparent and highly transmissive media, the contrast is either reversed or totally random. This can be attributed to the fact that the contribution of the background increases. This phenomenon can be observed in the thin parts of the dress shown in Figure 11. This is what histogram alone cannot capture without being aware of the spatial information. They have further shown that although the contrast in both low spatial and high spatial domains contribute to translucent appearance, the latter is more important and is able to yield translucent appearance

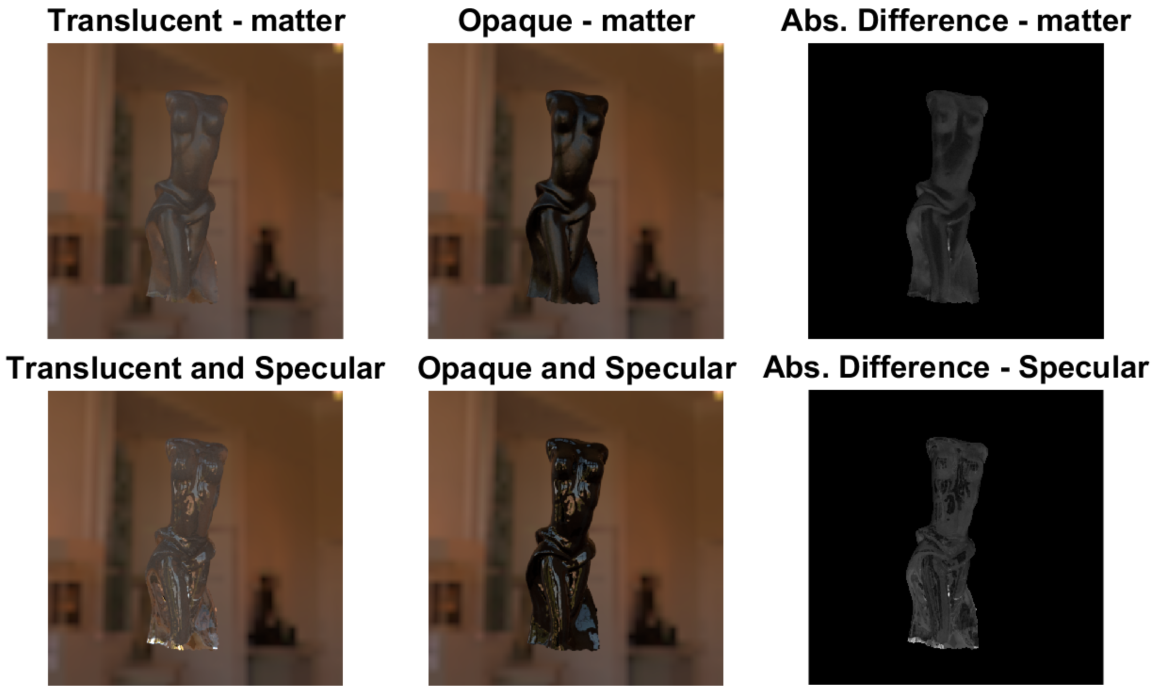

Figure 27: The absolute difference between the two images shows that the translucent and opaque objects have identical intensities in the specular regions. Besides, the contrast between the specular and non-specular regions as well as the spatial coverage of the regions where a translucent object and opaque object differ in intensities is modulated by surface roughness and thus, glossiness of the object.

even if the contrast in the low spatial frequency is held constant (refer to Figures 5-6 in (Motoyoshi, 2010)). This observation has implications in the image-based material editing and it has been demonstrated to be important for the image-based translucency transfer (Todo et al., 2019). The observation that blurring non-specular regions is associated with translucency, while specular highlights remain intact, also explains why N-shaped non-linearity, which generates larger changes for lower intensity inputs, has been able to enhance the perceived degree of translucency.

Xiao *et al.*

Xiao et al. (2020) have shown that a sharp surface relief enhances perceived opacity and argue that this can be attributed to sharper and darker shadows generated by these areas, which on the one hand agrees with the previous findings (Fleming & Bülthoff, 2005; Motoyoshi, 2010) that blurriness and brightness (mean luminance) of the shadowed regions can play a role in translucency perception, but on the other hand, should be taken with care, as the sharp and fine details of the surface can also lead to interreflections and bright appearance, as this is the case for the Lucy in Figure 14. In another work (Xiao et al., 2014), they observed that the thin parts and fine details of the Lucy contribute most to translucent material discrimination, supporting similar observations by other researchers (Fleming & Bülthoff, 2005; Nagai et al., 2013; Gkioulekas et al., 2015). One objective measure for these kind of sharp details could be surface curvature, which to some extent captures both sharp-fine details and outer edges of the object (refer to Figure 28). However, such metric does not capture flat thin areas (see the blue region in the bottom right corner of the figure), which although have low curvature, still appear different, which makes them a diagnostic cue for distinguishing between opaque and transmissive materials.

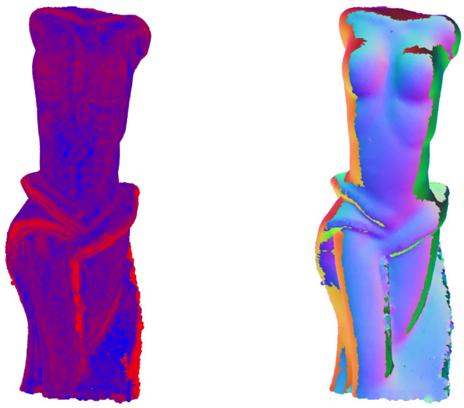

Figure 28: The left figure illustrates the surface curvature of the object (red areas - high curvature; blue areas - low curvature). The curvature can be interesting in two ways: it highlights sharp details and surface concavities which are expected to be in a shadow in case of opaque objects; on the other hand, the sharp edges near the thin areas, which transmit light easily, are also characterized with high curvature. However, note that non-edge parts of the flat thin regions have low curvature, but still transmit large amount of light (see the bottom right corner of the figure - the edge is red, but most of its thin dress part is in blue). The right image is a pseudocolor map of the surface normals - the points where the normals are facing the same direction are colored with similar colors. On opaque surfaces, the normals and thus, the shape can be estimated from the shading information.

Gkioulekas *et al.*

One particular instance of thin regions, the edges, generally have been observed to be informative about material translucence (Fleming & Bülthoff, 2005; Xiao et al., 2014). Therefore, Gkioulekas et al. (2015) tried to take advantage of this and utilized the radiance information near the edges to deduce the subsurface scattering properties of a material. They limited the study to the edges which are the result of surface discontinuity - such as those at the boundary of the two facets of a cube. They split the edges into four qualitative regions on the two facets and simulated a broad range of materials to observe how the radiance information in those regions varies. They found that each material has its surprising "signature" radiance profile at the edges (e.g. refer to Figure 29). Each radiance profile encapsulates information about reflection, refraction and scattering properties of a material. They analyzed from an optical point of view, how single scattering (single bounce of a photon), mid-order scattering and high-order scattering contribute to the energy incident on the camera sensor. A typical radiance profile is illustrated in Figure 3 of (Gkioulekas et al., 2015). For instance, a relatively high extinction coefficient puts intensity maxima closer to the boundary on the side facing away from the illumination direction, because the penetration depth decreases (this phenomenon is illustrated in Figure 29 and can be also observed in Figure 3). It is although noteworthy that the high extinction coefficient eliminates the maxima completely due to opacity (Figure 29). High albedo, i.e. a higher portion of scattering and a lower portion of absorption in a given extinction coefficient, impacts intensity of these extrema, not their locations. When the albedo is high, high order scattering contributes more than single scattering and the intensity decreases. The angular variance of the phase function also impacts the location of the local maxima. These local maxima can be noticed as small peaks in Figure 4. Afterwards they demonstrated that different scattering effects can generate matching radiant profiles - i.e. edge profile "metamers". On the other hand, when scattering properties are fixed, the profile is unique to a given refraction and illumination direction. We can, indeed, match refraction and illumination effects on the one facet - e.g. if the refraction index changes, we can change the illumination angle accordingly to get the "original" reflectance angle; however, this change will impact the other facet, generating a different radiance profile. The idea of using radiance profiles for translucency discrimination was novel. The authors demonstrated that their findings generalize well across many illumination geometries and broad range of translucent materials that makes this work one of the most signifi-

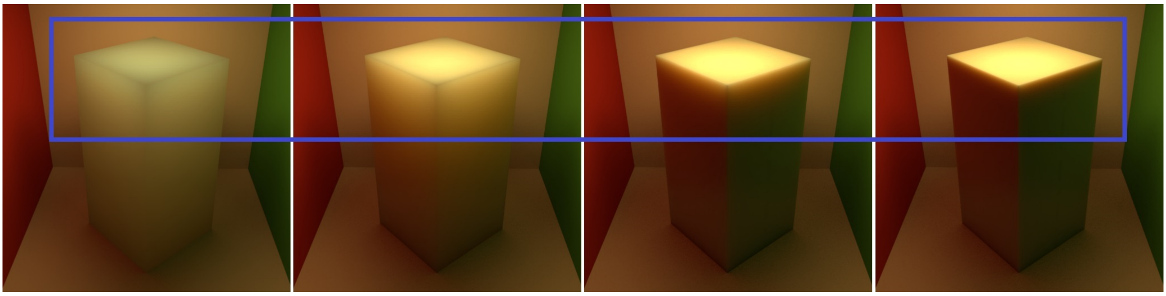

Figure 29: The objects are rendered with a skimmed milk material (Jensen et al., 2001) and an extinction coefficient is scaled to different levels. The intensity distribution at the edges varies considerably between different optical densities. As noted by Gkioulekas et al. (2015) the local maxima are moved closer to the edge, when optical density increases (observe darker strips across the edges in the left two images). However, when optical density is too high (the rightmost image), the material becomes opaque and the facet which is not directly illuminated looks dark and homogeneous. The blue frame highlights the areas where the edges are most informative.

cant contributions to the topic. On the other hand, it is also worth noting that the study was limited to the convex edges and might not generalize to concavities with strong interreflections (such as Lucy in Figure 14).

The authors argued that the edge radiance profiles are robust to the real world artifacts and can be reliable indicators for edge detection and material identification algorithms in computer vision systems. While they seem a robust indicator for machines, it remains unknown whether the HVS relies on similar edge profiles for translucency perception. Psychophysical experiments need to be conducted in the future to explore this question. It is interesting to observe whether image manipulations and mapping textures of different radiance profiles near the edges of different surfaces affect observers' estimations of the subsurface scattering properties. Additionally, proper eye-tracking measurements could also reveal the saliency of the edge profile components when subjects are performing translucency-related visual tasks.

Marlow *et al.*

Marlow et al. (2017) have shown that the co-variance between surface orientation and shading is related to opacity, while the lack of it produces translucent appearance. They mapped an identical texture of the luminance gradient onto the surfaces with different apparent 3D shapes. They observed that the interpretation of the material properties from a given luminance gradient is impacted by the perceived 3D shape. Particularly, if the image intensities co-vary with the perceived surface orientation, the material appears opaque; otherwise, it appears translucent (refer to the Movies S1 and S2 in (Marlow et al., 2017)). Additionally, they illustrated that when the light field the material is embedded in "accidentally" eliminates this co-variance between surface and shading of an opaque object, a vivid and convincing illusion of translucency is observed. The fact that perceived 3D shape impacts the apparent translucency implies that the luminance contrast, mean luminance or similar statistics per se are not enough to explain the perception of translucency, and the HVS is likely to be exploiting surface geometry and 3D shape information as well. However, it is not clear how the HVS calculates the geometry from the retinal images. In our opinion, one way the HVS might be quantifying this is the relation between the surface curvature and the surface normals (Figure 28), on the one hand, and the magnitude and direction of the shading gradient, on the other hand (Figure 30). While the gradient orientation largely depends on the surface 3D geometry in the diffuse opaque objects, it is more random in objects with a high degree of subsurface light transport. Moreover, the gradient magnitude is largest in highly curved areas in the opaque object, while that is not necessarily true for the translucent ones (see the caption in Figure 30 and compare with Figure 28).

Marlow and Anderson

Marlow & Anderson (2021) have recently shown that translucent materials are also subject to photogeometric constraints. The authors argue that there is a co-variance among the luminance gradient produced by the sub-surface scattering of light, the shape of the specular reflections and the shape of the self-occluding contours, and this co-variance provides information about material properties and the 3D shape of the object. The co-variance can be rooted in the fact that all three components – sub-surface scattering, specular reflections and self-occluding contours are affected by the same objective geometric priori – the 3D surface curvature. For instance, both the luminance gradient produced by the sub-surface scattering and the shape of the specular reflections are usually aligned with the direction of the lowest surface curvature – making them aligned with one another as well.

First of all, the authors demonstrate that the intensity gradient produced by the sub-surface scattering is affected by the 3D shape. In opaque materials the location of the luminance extrema depends on the surface orientation in space, as the luminance extrema are located on the sides of the convex and concave regions – whichever side faces the illumination is brightest and whichever faces away from it is in the shadow. Contrastingly, in translucent materials, the intensity gradient is related with the local surface curvature; the decrease in the extinction coefficient usually smoothens the gradient, decreases the luminance contrast (which is consistent with other works (Motoyoshi, 2010)) and moves locally brightest and darkest intensities closer to the peaks of the convexities and concavities, respectively. The authors also provide an optical explanation for this: the light attempting to exit the material is redirected towards convexities and away from concavities due to the internal reflections. These observations on translucent materials generalize well to a broad range of frontal and side illumination angles.

Information about 3D shape can facilitate the estimation of the material properties and vice versa. However, in real-life scenario, neither is hardly ever known to the observer. The HVS somehow manages to recover both 3D shape and properties of a material that according to Marlow & Anderson (2021) can be rooted in the above-mentioned co-variation among subsurface scattering, specular reflections and self-occluding contours. The authors conducted psychophysical experiments, which supported those hypotheses. They observed that presence of self-occluding contours and specular reflections increased the vividness of the perceived 3D shape of a bumpy translucent surface. On the other hand, the magnitude of perceived translucency was significantly increased by the specular reflections but was barely affected by self-occluding contours.

This work opens a new avenue for translucency perception research. Although previous work concluded that information on 3D shape is important for translucency (Marlow et al., 2017), this work is the first one to propose that the HVS might be recovering shape and material properties simultaneously, from the same photogeometric constraints. The major limitation of the work is that it does not cover back-lit objects. Identification of the similar photogeometric constraints for back-lit objects is considered very difficult or even impossible by the authors, leaving the question open.

Di Cicco *et al.*

Di Cicco, Wijntjes, & Pont (2020) have recently conducted a study on citrus fruit images. They used multidimensional scaling (MDS) and constructed a 2-dimensional perceptual space explaining the qualities related to translucency. This is an elegant example how translucency perception research can benefit from relying on artworks. They observed that color saturation, intensity gradient and highlights were visible features for translucent materials, while being also related to ”juiciness”. They argue that the intensity magnitude and sharpness in their case supports earlier findings that blur and contrast are important cues for translucency perception. The authors also identified *translucency-related regions*, similarly to Nagai et al. (2013), which in their case is the ”peeled side” of the fruit. This can be accounted to the fact that the pulp permits light penetration and bleeding around the edges. Although overall trends and cues are consistent with the previous findings, the peculiarity of the stimuli makes generalization to other translucent materials and photorealistic stimuli debatable.

Summary

To summarize, a full perceptual model of translucency which could simply take scene and material properties as an input and provide an estimation of a perceptual correlate remains beyond reach nowadays. The possibility of coming up with this kind of model anytime soon ranges from unlikely to impossible. Nevertheless, some partial models have collected interesting observations about translucency perception cues. These works complement each other and can be summarized as follows:

1. It seems that neither luminance nor spatial information alone is enough for estimating perceived translucency. The HVS seemingly uses some sophisticated combination of the both.
2. The spatial regions where a photon can go through easily look brighter and contain rich information about material

translucence. Examples of this kind of regions are edges, thin parts and sharp fine details of a surface geometry.

3. The regions which are usually shadowed in opaque objects are also informative about translucency, as they look brighter in translucent materials.

4. Points 2 and 3 can be generalized as follows: if in absence of subsurface light transport considerably smaller amount of light could have reached a particular region, this region can be diagnostic for material translucence.

5. Understanding how much light could or could not have reached a particular region inherently involves understanding the surface geometry and global correlation among different spatial regions.

6. It is not known how the HVS segments an image, how it identifies the informative regions and how it calculates the surface geometry. These calculations are not unique and vary across individuals. There can be multiple translucency cues in a proximal stimulus and different people can rely on different ones for yet unknown reasons.

We believe that in addition to standard psychophysics, where experimenters attempt to find a correlation between the varying physical parameters and the observer responses, it is also important to study translucency perception process from a behavioral perspective. The first step towards this has been done by Gigilashvili, Thomas, et al. (2018, 2021). In the subsequent section, we will analyze what visual mechanisms remain to be uncovered and what factors complicate the translucency perception research.

Challenges and knowledge gaps

Inconsistent definition and conceptual ambiguity

The exact meaning of translucency is not universally accepted and remains subjective (Pointer, 2003). This basic definition problem might make the scientific communication difficult and hinder the advance in the translucency perception research. We have particularized these problems in the recent position paper (Gigilashvili, Thomas, et al., 2020). Care is needed to avoid miscommunication of the empirical results and to ensure the reproducibility of the psychophysical experiments. Experimenters should make sure that the instructions are correctly understood and interpreted by their observers when the task concerns translucency perception - especially, when the experiments are conducted in languages other than English, as the translation of the term *translucency* might or might not differ from that of *transparency*. For example, Motoyoshi (2010) reports that there is no distinction between *transparent* and *translucent* in the Japanese language, which might have impacted his experimental results. However, he reports that observers assess translucent and transparent stimuli differently from each other, seemingly understanding the semantic difference between the two visual phenomena. This makes the author propose that the two concepts might be orthogonal. Scaling translucency remains a challenging and confusing task. To the best of our knowledge, Hutchings & Scott (1977) and Hutchings & Gordon (1981) (cited in (Hutchings, 2011)) have been the first ones to observe the confusion among the experiment participants while scaling translucency. The authors argue that *"care should be taken when using the term Translucency for scaling. An increase in translucency may mean an increase in transparency to some panelists while meaning the opposite to others"* (Hutchings, 2011). We have also observed a similar kind of problem in our experiments (Gigilashvili, Thomas, et al., 2018; Gigilashvili, Dubouchet, et al., 2020; Gigilashvili, Thomas, et al., 2021). The lack of knowledge on how to quantify translucency makes it challenging to measure it by magnitude estimation techniques (Torgerson, 1958) and psychophysical scaling methods, such as the pair comparison and rank order (Engeldrum, 2000). For example, it has been possible to quantify the magnitude of glossiness (Pellacini et al., 2000) or to differentiate more glossy and less glossy stimuli (Thomas et al., 2017; Gigilashvili, Thomas, et al., 2019). However, there is no universal agreement what "more translucent" means, neither can we tell "how much" translucency is in a given stimulus. When comparing multiple stimuli, which one is the most translucent (e.g. in Figure 2) - the one closest to transparency, closest to opacity or closest to a hypothetical peak between the two? Di Cicco, Wiersma, et al. (2020b) have observed that translucency was judged least consistently among all assessed parameters in the still life paintings of grapes, which might be attributed to the variation in the conceptual understanding, rather than the anatomical differences among observers. Nagai et al. (2013) defined *more translucent* in their experiments as having stronger subsurface scattering. Wijntjes et al. (2020) have defined translucency as *"the*

Figure 30: The top row illustrates the original images, the middle row shows the magnitude of the luminance gradient (the background is ignored), while the shades in the bottom row correspond to the gradient orientation. The back-lit translucent object has a high magnitude gradient, while it looks relatively homogeneous when front-lit. Under front light, a diffuse object produces more visible gradient (due to shading in the surface convexities) than a front-lit translucent one. In the front-lit condition, the gradient orientation closely follows the surface 3D geometry. When the objects are back-lit, the gradient orientation of an opaque object is strongly impacted by the partial side-illumination, while it looks more random for the back-lit translucent object, which also makes it difficult to recover its shape.

opposite of opaqueness, but... not limiting to pure transparency. For example, tea with milk is more translucent than a cup of white paint". Di Cicco, Wiersma, et al. (2020b) asked observers to quantify the magnitude of translucency of the painted grapes and defined the term in a similar manner: *"Translucency: how translucent do the grapes appear to you? Low values indicate that no light passes through the grapes and the appearance is opaque; high values indicate that some light passes through the grapes."* However, care should be taken in these cases as well, because we do not know whether the relation between scattering and translucency is monotonous. Materials with high and low scattering might be considered opaque and transparent, respectively - both having zero translucency. Many works avoid direct quantification of translucency in the psychophysical experiments and encapsulate it in the matching tasks asking observers to match the stimuli by appearance (Xiao et al., 2014; Fleming & Bülthoff, 2005) and/or by translucency (Gkioulekas et al., 2013; Xiao et al., 2020; Gigilashvili, Urban, et al., 2019). This, at first glance, simplifies the task. However, there is little empirical evidence that the HVS can fully isolate translucency from other attributes of total appearance. If the definition of translucency is ambiguous to the observers, how can they match materials by translucency and how can we guarantee that they are not making up their own rules for matching the stimuli, e.g. by lightness, or any property other than translucency? In order to identify what observers are basing their decisions on, the experimenters can calculate particular image statistics and check how well these statistics explain the variation in the observer responses (as done by Chadwick et al. (2019)). However, there is no guarantee that the actual statistics or cues used by observers will be correctly identified by the experimenters. Another workaround found in the literature is using the terms more familiar and less abstract than translucency. For instance, Chadwick et al. (2019) asked observers to assess *strength* and *milkiness* of the tea images. However, the association between the strength, milkiness and translucency is not clear either. Hutchings (2011) proposes using *extent of visibility* scale of Galvez & Resurreccion (1990) instead of referring to *"more translucent"* and *"less translucent"*. However, the scale is intended for assessing the appearance of the mungbean noodles in a plastic cup and for quantifying the visibility of the objects behind the noodle strands - thus, it is not readily applicable to the solid non-see-through materials. Furthermore, the inconstancy of translucency across different shapes makes it challenging to clearly separate translucency as a property of a given object and as a property of a material the object is made of. We observed (Gigilashvili, Thomas, et al., 2018; Gigilashvili, Urban, et al., 2019; Gigilashvili, Thomas, et al., 2021) that human observers find it challenging to compare or match translucency across different shapes, for two reasons: first of all, it is difficult to estimate optical properties of a material and to decouple its visual appearance from the shape-related effects (speaks of the limited ability to "invert optics" as it has been noted previously (Fleming & Bülthoff, 2005; Anderson, 2011; Chadwick et al., 2019)); secondly, the task is inherently ambiguous - translucency cues vary not only between the thick and thin objects, but also between the thick and thin regions of a particular object - making observers uncertain which region to assess and how to come up with a single translucency measure. According to Hutchings (1994), a heterogeneous material might have *"more than one colour, perhaps more than one translucency, gloss, or surface irregularity"* that no appearance profile system can deal with. The observers in the experiments by Nagai et al. (2013) pointed out that heterogeneous translucency which resulted from a varying shape, complicated the task, but it remained still viable according to the authors. This raises a question: should translucency of a complex-shaped homogeneous material be judged globally for a given object or material, or locally for each specific region of an object?

Challenges in experimental methods

One of the pivotal limitations of the experimental methods are the constraints related to the visual stimuli selection. Real objects, photographs or computer-generated imagery can be used to study translucency perception psychophysically. All of these methods come with their advantages and drawbacks, which are summarized in **Appendix 1 of (Gigilashvili, Thomas, et al., 2021)**. We advocate for using physical objects which make the experiments closer to the real-life scenarios, permitting binocular vision, interaction, motion cues, higher dynamic range and multisensory information (tactile, auditory, olfactory). We hypothesize that the behavioral patterns applied by observers on physical objects are close to their natural way of making judgments. On the other hand, we are aware of the trade-offs. Physical objects are difficult and expensive to model, measure and replicate. The experiments usually take longer (Maloney & Knoblauch, 2020) and the risk of damaging, the unpredictable effects of aging and the limited access across the scientific community hinder the reproducibility of the experiments. A descent alternative which permits interactivity, motion and binocular cues can be the immersive reality technologies.

A further aspect which is problematic from the experimental point of view is the lack of standardization. Normal conditions for observing translucency and a standard observer are not defined. For instance, the contrast sensitivity and the visual acuity might have a significant impact on the experimental results. However, the viewing conditions, such as the distance and the size of the visual field varies across different experiments which complicates the comparative analyses of their findings.

And last but not least, unlike color vision (Lafer-Sousa et al., 2015; Emery & Webster, 2019), the knowledge about cross-individual differences in translucency perception is virtually non-existent, as pooled experimental results are usually reported. Chadwick et al. (2018, 2019) have observed that the models explaining the variation in the psychophysical data differ among individuals. Similar cross-individual differences were also observed by Nagai et al. (2013) and Gigilashvili, Thomas, et al. (2018). Whether this could be attributed to the interpretation of the task, prior experience or anatomical differences, need to be answered in the future. Should we expect the translucency counterpart of *#TheDress* anytime soon which could expose these individual differences?

Visual mechanisms of translucency perception

The exact mechanisms of translucency perception remain largely unidentified. After compilation of the state-of-the-art works, we came up with the several important questions which we believe should be addressed in future works. Namely:

- Which image cues and regions does the HVS rely on and how does it identify, calculate and weight them?
- What is the role of shape and geometry perception and how does the HVS calculate them?
- To what extent is perceived translucency impacted by other appearance attributes, such as color, gloss, texture and fluorescence?
- What role do the identification of the familiar materials and other psychological priors play in translucency perception?
- How does the HVS use motion and scene dynamics to assess translucency?
- What is the physiology of translucency perception from the retinal to the cortical level and how much does it vary across individuals?

We envision that the future work can develop in three directions: the eye-tracking experiments can facilitate identification of the respective cues and key image regions; behavioral analysis (similar to (Gigilashvili, Thomas, et al., 2021)) might reveal how the judgments on translucency are made and which factors guide observers' actions; while neuroscience can shed light to the physiological and cognitive aspects in the perplexing process of translucency perception.

Eye tracking can potentially reveal the most salient cues to translucency and whether different observers rely on different cues, as noted by Nagai et al. (2013). Eye tracking is a more straightforward way than reverse correlation techniques used by Nagai et al. (2013) to learn where observers look in the process of translucency assessment. Additionally, the saccade paths measured with the eye tracking could reveal the sequence and the frequency of inspecting particular local regions. This could potentially reveal how different local regions relate to one another. On the other hand, we understand that eye tracking comes with the considerable limitations: the reason of the fixations might be unrelated to translucency – some regions might be salient for other reasons, e.g. human face can attract extra attention regardless the task; we will not capture the influence of the parafoveal vision and the cues which are not locally defined; the temporal resolution of the eye tracking equipment might be lower than the speed of the visual processing; the presence of eye tracking equipment might affect the naturalness of the interaction.

Summary and Conclusions

We have discussed translucency as one of the pivotal appearance attributes, which is increasingly important in a broad range industries and disciplines, including 3D printing, cosmetics, food industry and arts among many. Translucency results from the subsurface transport of light. While the techniques for measuring and modeling the optical properties of a material are relatively well-established, our understanding how they link to their perceptual correlates remains limited. The advance in translucency perception research is attributed to the development of computer graphics techniques which permit easier generation of the translucent visual stimuli. While the initial studies were limited to transparency perception, transparency models could not explain the perception of highly scattering media. The visual cues and perceptual mechanisms seem to be fundamentally different between the transparency and translucency of the see-through filters and translucency of highly scattering, non-see-through media. This resulted in emergence of a separate research topic - the perception of translucency in highly scattering media. In the past 20 years, multiple factors have been identified to be contributing to perceived translucency, such as the illumination direction, structural thickness of the object, as well as subsurface scattering properties. It is believed that the luminance distribution around the edges and in the shadowed regions, and its covariance with the surface geometry, might be used by the human visual system to infer translucency in highly scattering, non-transparent materials, while the HVS relies on apparent contrast and blur when the background is visible. Nevertheless, overall translucency perception research is still in its infancy.

We argue that the problems with the conceptual understanding and comprehension of the term impede the advance of the research and complicate the reproducibility of the tasks. We argue for the better standardization in this domain. Finally, we believe that eye tracking experiments could reveal which image regions and cues are significant, and advance in neuroscience could provide a deeper insight in the corresponding anatomical mechanisms for translucency perception.

Acknowledgments

The work has been funded by the MUVApp project (#250293) of the Research Council of Norway.

References

Anderson, B. L. (2011). Visual perception of materials and surfaces. *Current biology*, *21*(24), R978–R983.

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, *9*(11:10), 1–17.

Anfe, T. E. d. A., Caneppele, T. M. F., Agra, C. M., & Vieira, G. F. (2008). Microhardness assessment of different commercial brands of resin composites with different degrees of translucence. *Brazilian Oral Research*, *22*(4), 358–363.

Assen, J. J. R. van, Nishida, S., & Fleming, R. W. (2020). Visual perception of liquids: Insights from deep neural networks. *PLoS computational biology*, *16*(8), 29 pages.

ASTM D 1003 - Standard Test Method for Haze and Luminous Transmittance of Transparent Plastics. (2003).

ASTM E284-17 Standard Terminology of Appearance. (2017). Available from https://doi.org/10.1520/E0284-17

Barry, F. (2011). *Painting in stone: The symbolism of colored marbles in the visual arts and literature from antiquity until the enlightenment*. Unpublished doctoral dissertation, Columbia University.

Bartleson, C. J. (1960). Memory colors of familiar objects. *JOSA*, *50*(1), 73–77.

Beck, J., & Ivry, R. (1988). On the role of figural organization perceptual transparency. *Perception & psychophysics*, *44*(6), 585–594.

Beck, J., Prazdny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & psychophysics*, *35*(5), 407–422.

Brunton, A., Arikan, C. A., Tanksale, T. M., & Urban, P. (2018). 3D printing spatially varying color and translucency. *ACM Transactions on Graphics (TOG)*, *37*(4), 157:1–157:13.

Brunton, A., Arikan, C. A., Urban, P., & Tanksale, T. M. (2020, July 21). *Method for joint color and translucency 3D printing and a joint color and translucency 3D printing device.* Google Patents. (US Patent 10,717,235)

BYK Gardner GmbH. (n.d.). *Haze-gard transparency transmission haze meter.* (Accessed on 16/10/20 at: https://www.byk-instruments.com/us/en/Appearance/haze-gard-Transparency-Transmission-Haze-Meter/c/2345)

Cavanagh, P. (2005). The artist as neuroscientist. *Nature*, *434*(7031), 301–307.

Chadwick, A. C., Cox, G., Smithson, H. E., & Kentridge, R. W. (2018). Beyond scattering and absorption: Perceptual unmixing of translucent liquids. *Journal of Vision*, *18*(11:18), 1–15.

Chadwick, A. C., Heywood, C., Smithson, H. E., & Kentridge, R. W. (2019). Translucence perception is not dependent on cortical areas critical for processing colour or texture. *Neuropsychologia*, *128*, 209–214.

Chadwick, A. C., & Kentridge, R. W. (2015). The perception of gloss: A review. *Vision research*, *109*, 221–235.

Chandrasekhar, S. (1960). Radiative transfer. In (pp. 1–53). (1960) Dover Publications Inc. New York.

Chousalkar, K., Flynn, P., Sutherland, M., Roberts, J. R., & Cheetham, B. F. (2010). Recovery of salmonella and escherichia coli from commercial egg shells and effect of translucency on bacterial penetration in eggs. *International journal of food microbiology*, *142*(1-2), 207–213.

Chowdhury, N. S., Marlow, P. J., & Kim, J. (2017). Translucency and the perception of shape. *Journal of Vision*, *17*(3:17), 1–14.

Cunningham, D. W., Wallraven, C., Fleming, R. W., & Straßer, W. (2007). Perceptual reparameterization of material properties. In *Computational aesthetics* (pp. 89–96).

d'Eon, E., & Irving, G. (2011). A quantized-diffusion model for rendering translucent materials. *ACM Transactions on Graphics (TOG)*, *30*(4), 1–14.

Di Cicco, F., Wiersma, L., Wijntjes, M., & Pont, S. (2020a). Material properties and image cues for convincing grapes: The know-how of the 17th-century pictorial recipe by willem beurs. *Art & Perception*, *8*(3-4), 337–362.

Di Cicco, F., Wiersma, L., Wijntjes, M., & Pont, S. (2020b). Material properties and image cues for convincing grapes: The know-how of the 17th-century pictorial recipe by willem beurs. *Art & Perception*, *1*(aop), 1–26.

Di Cicco, F., Wijntjes, M. W., & Pont, S. C. (2020). If painters give you lemons, squeeze the knowledge out of them. a study on the visual perception of the translucent and juicy appearance of citrus fruits in paintings. *Journal of Vision*, *20*(13:12), 1–15.

Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R., Hartung, B., & Kersten, D. (2011). Visual motion and the perception of surface material. *Current Biology*, *21*(23), 2010–2016.

Donner, C., & Jensen, H. W. (2005). Light diffusion in multi-layered translucent materials. *ACM Transactions on Graphics (ToG)*, *24*(3), 1032–1039.

D'Zmura, M., Colantoni, P., Knoblauch, K., & Laget, B. (1997). Color transparency. *Perception*, *26*(4), 471–492.

D'Zmura, M., & Iverson, G. (1993). Color constancy. i. basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *JOSA A*, *10*(10), 2148–2165.

Emery, K. J., & Webster, M. A. (2019). Individual differences and their implications for color perception. *Current opinion in behavioral sciences*, *30*, 28–33.

Emmert, R. (1996). Quantification of the soft-focus effect: Measuring light-diffusing characteristics of cosmetic pigments and powders. *Cosmetics and toiletries*, *111*(7), 57–61.

Engeldrum, P. G. (2000). *Psychometric scaling: a toolkit for imaging systems development*. Imcotek.

Eugène, C. (2008). Measurement of "Total Visual Appearance": a CIE challenge of soft metrology. In *12th imeko tc1 & tc7 joint symposium on man, science & measurement* (pp. 61–65).

Falkenberg, C., & Faul, F. (2019). Transparent layer constancy is improved by motion, stereo disparity, highly regular background pattern, and successive presentation. *Journal of Vision*, *19*(12:16), 1–33.

Faul, F. (2017). Toward a perceptually uniform parameter space for filter transparency. *ACM Transactions on Applied Perception (TAP)*, *14*(2), 1–21.

Faul, F. (2019). The influence of fresnel effects on gloss perception. *Journal of Vision*, *19*(13:1), 1–39.

Faul, F., & Ekroll, V. (2002). Psychophysical model of chromatic perceptual transparency based on substractive color mixture. *JOSA A*, *19*(6), 1084–1095.

Faul, F., & Ekroll, V. (2011). On the filter approach to perceptual transparency. *Journal of Vision*, *11*(7:7), 1–33.

Faul, F., & Ekroll, V. (2012). Transparent layer constancy. *Journal of Vision*, *12*(12:7), 1–26.

Fleming, R. W. (2014). Visual perception of materials and their properties. *Vision research*, *94*, 62–75.

Fleming, R. W., & Bülthoff, H. H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception (TAP)*, *2*(3), 346–382.

Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, *3*, 347–368.

Fleming, R. W., Jäkel, F., & Maloney, L. T. (2011). Visual perception of thick transparent materials. *Psychological science*, *22*(6), 812–820.

Fleming, R. W., & Storrs, K. R. (2019). Learning to see stuff. *Current Opinion in Behavioral Sciences*, *30*, 100–108.

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, *4*(9:10), 798–820.

Fleming, R. W., Wiebel, C., & Gegenfurtner, K. (2013). Perceptual qualities and material classes. *Journal of Vision*, *13*(8:9), 1–20.

Frisvad, J. R., Jensen, S. A., Madsen, J. S., Correia, A., Yang, L., Gregersen, S. K., et al. (2020). Survey of models for acquiring the optical properties of translucent materials. *STAR*, *39*(2), 729–755.

Galvez, F. C. F., & Resurreccion, A. V. (1990). Comparison of three descriptive analysis scaling methods for the sensory evaluation of noodles 1. *Journal of Sensory Studies*, *5*(4), 251–263.

Gerardin, M., Simonot, L., Farrugia, J.-P., Iehl, J.-C., Fournel, T., & Hébert, M. (2019). A translucency classification for computer graphics. In *Material Appearance 2019, Electronic Imaging* (pp. 203:1–203:6). Society for Imaging Science and Technology.

Gerbino, W. (1994). Achromatic transparency. *Lightness, brightness, and transparency*, 215–255.

Gerbino, W., Stultiens, C. I., Troost, J. M., & Weert, C. M. de. (1990). Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 3–20.

Giancola, G., & Schlossman, M. L. (2015). Decorative cosmetics. *Cosmeceuticals and Active Cosmetics*, 191–219.

Gigilashvili, D., Dubouchet, L., Pedersen, M., & Hardeberg, J. Y. (2020). Caustics and translucency perception. In *Material Appearance 2020, IS&T International Symposium on Electronic Imaging* (pp. 033:1–033:6).

Gigilashvili, D., Mirjalili, F., & Hardeberg, J. Y. (2019). Illuminance impacts opacity perception of textile materials. In *Color and imaging conference* (pp. 126–131).

Gigilashvili, D., Pedersen, M., & Hardeberg, J. Y. (2018). Blurring impairs translucency perception. In *Color and imaging conference* (pp. 377–382).

Gigilashvili, D., Shi, W., Wang, Z., Pedersen, M., Hardeberg, J. Y., & Rushmeier, H. (2021). The Role of Subsurface Scattering in Glossiness Perception. *ACM Transaction on Applied Perception*, *18*(3), 10:1-10:26.

Gigilashvili, D., Thomas, J.-B., Hardeberg, J. Y., & Pedersen, M. (2018). Behavioral investigation of visual appearance assessment. In *Color and imaging conference* (pp. 294–299).

Gigilashvili, D., Thomas, J. B., Hardeberg, J. Y., & Pedersen, M. (2020). On the nature of perceptual translucency. *8th Annual Workshop on Material Appearance Modeling (MAM2020). Eurographics Digital Library*, 17–20.

Gigilashvili, D., Thomas, J.-B., Pedersen, M., & Hardeberg, J. Y. (2019). Perceived glossiness: Beyond surface properties. In *Color and imaging conference* (pp. 37–42).

Gigilashvili, D., Thomas, J.-B., Pedersen, M., & Hardeberg, J. Y. (2021). On the appearance of objects and materials: Qualitative analysis of experimental observations. *To appear in the Journal of the International Colour Association (JAIC)*, 33 pages.

Gigilashvili, D., Urban, P., Thomas, J.-B., Hardeberg, J. Y., & Pedersen, M. (2019). Impact of shape on apparent translucency differences. In *Color and imaging conference* (pp. 132–137).

Gigilashvili, D., Urban, P., Thomas, J.-B., Hardeberg, J. Y., & Pedersen, M. (2021). The impact of optical and geometric thickness on perceived translucency differences. *Under submission*, 17 pages.

Gkioulekas, I., Walter, B., Adelson, E. H., Bala, K., & Zickler, T. (2015). On the appearance of translucent edges. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 5528–5536).

Gkioulekas, I., Xiao, B., Zhao, S., Adelson, E. H., Zickler, T., & Bala, K. (2013). Understanding the role of phase function in translucent appearance. *ACM Transactions on Graphics (TOG)*, *32*(5), 1–19.

Hagedorn, J., & D'Zmura, M. (2000). Color appearance of surfaces viewed through fog. *Perception*, *29*(10), 1169–1184.

Hartung, B., & Kersten, D. (2002). Distinguishing shiny from matte [Abstract]. *Journal of Vision*, *2*(7), 551.

Hašan, M., Fuchs, M., Matusik, W., Pfister, H., & Rusinkiewicz, S. (2010). Physical reproduction of materials with specified subsurface scattering. In *Acm siggraph 2010 papers* (pp. 1–10).

Henyey, L. G., & Greenstein, J. L. (1941). Diffuse radiation in the galaxy. *The Astrophysical Journal*, *93*, 70–83.

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision research*, *45*(13), 1707–1724.

Hunter, R. S., & Harold, R. W. (1987). *The measurement of appearance*. John Wiley & Sons.

Hutchings, J. B. (1977). The importance of visual appearance of foods to the food processor and the consumer 1. *Journal of Food Quality*, *1*(3), 267–278.

Hutchings, J. B. (1994). Appearance profile analysis and sensory scales. In *Food colour and appearance* (pp. 142–198). Springer.

Hutchings, J. B. (2011). *Food colour and appearance*. Springer Science & Business Media.

Hutchings, J. B., & Gordon, C. (1981). Translucency specification and its application to a model food system. In *Proceedings of the fourth congress of the international colour association, west berlin.*

Hutchings, J. B., & Scott, J. (1977). Colour and translucency as food attributes. In *Color 77, proceedings of the 3rd congress of the international colour association, troy, new york* (pp. 10–15).

Igarashi, T., Nishino, K., & Nayar, S. K. (2005). The appearance of human skin. Technical Report: CUCS-024-05. , 85 pages.

Jakob, W. (2010). *Mitsuba renderer.* (http://www.mitsuba-renderer.org)

Jensen, H. W., Marschner, S. R., Levoy, M., & Hanrahan, P. (2001). A practical model for subsurface light transport. In *Proceedings of the 28th annual conference on computer graphics and interactive techniques* (pp. 511–518).

Kaltenbach, F. (2012). *Translucent materials: glass, plastics, metals*. Walter de Gruyter.

Kán, P., & Kaufmann, H. (2012). High-quality reflections, refractions, and caustics in augmented reality and their contribution to visual coherence. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 99–108).

Khang, B.-G., & Zaidi, Q. (2002). Cues and strategies for color constancy: Perceptual scission, image junctions and transformational color matching. *Vision Research*, *42*(2), 211–226.

Kim, J., Marlow, P., & Anderson, B. L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision*, *11(9)*(4), 1–19.

Kim, J., & Marlow, P. J. (2016). Turning the world upside down to understand perceived transparency. *i-Perception*, *7*(5), 1–5.

Koenderink, J. J., & Doorn, A. J. van. (2001). Shading in the case of translucent objects. In *Human vision and electronic imaging vi* (Vol. 4299, pp. 312–320).

Krewinghaus, A. B. (1969). Infrared reflectance of paints. *Applied optics*, *8*(4), 807–812.

Kubelka, P. (1931). Ein beitrag zur optik der farbanstriche (contribution to the optic of paint). *Zeitschrift fur technische Physik*, *12*, 593–601.

Kubelka, P. (1948). New contributions to the optics of intensely light-scattering materials. Part I. *JOSA*, *38*(5), 448–457.

Lafer-Sousa, R., Hermann, K. L., & Conway, B. R. (2015). Striking individual differences in color perception uncovered by 'the dress' photograph. *Current Biology*, *25*(13), R545–R546.

Legge, G. E., Parish, D. H., Luebker, A., & Wurm, L. H. (1990). Psychophysics of reading. xi. comparing color contrast and luminance contrast. *JOSA A*, *7*(10), 2002–2010.

Le Moan, S., & Pedersen, M. (2017). Evidence of change blindness in subjective image fidelity assessment. In *2017 IEEE International Conference on Image Processing (ICIP)* (pp. 3155–3159).

Li, C., Zhou, K., Wu, H.-T., & Lin, S. (2018). Physically-based simulation of cosmetics via intrinsic image decomposition with facial priors. *IEEE transactions on pattern analysis and machine intelligence*, *41*(6), 1455–1469.

Liu, M.-C., Aquilino, S. A., Lund, P. S., Vargas, M. A., Diaz-Arnold, A. M., Gratton, D. G., et al. (2010). Human perception of dental porcelain translucency correlated to spectrophotometric measurements. *Journal of Prosthodontics: Implant, Esthetic and Reconstructive Dentistry*, *19*(3), 187–193.

Lopes Filho, H., Maia, L. E., Araújo, M. V. A., & Ruellas, A. C. O. (2012). Influence of optical properties of esthetic brackets (color, translucence, and fluorescence) on visual perception. *American journal of orthodontics and dentofacial orthopedics*, *141*(4), 460–467.

Lovibond Tintometer. (n.d.). (Accessed on 14/12/20 at: https://www.lovibond.com/en/PC/Colour-Measurement/Applications)

Lynch, D. K., Livingston, W. C., & Livingston, W. (2001). *Color and light in nature*. Cambridge University Press.

Maloney, L. T., & Knoblauch, K. (2020). Measuring and modeling visual appearance. *Annual Review of Vision Science*, *6*, 519–537.

Marlow, P. J., & Anderson, B. L. (2021). The cospecification of the shape and material properties of light permeable materials. *Proceedings of the National Academy of Sciences*, *118*(14), 1–10.

Marlow, P. J., Kim, J., & Anderson, B. L. (2017). Perception and misperception of surface opacity. *Proceedings of the National Academy of Sciences*, *114*(52), 13840–13845.

Merriam-Webster Dictionary. (n.d.). https://www.merriam-webster.com/dictionary/translucent. (Accessed: 2020-11-06)

Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics*, *13*(1), 59–66.

Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*(4), 90–99.

Metelli, F. (1985). Stimulation and perception of transparency. *Psychological Research*, *47*(4), 185–202.

Midtfjord, H., Green, P., & Nussbaum, P. (2018). A model of visual opacity for translucent colorants. *Electronic Imaging*, *2018*(8), 210:1–210:6.

Motoyoshi, I. (2010). Highlight–shading relationship as a cue for the perception of translucent and transparent materials. *Journal of Vision*, *10*(9:6), 1–11.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, *447*(7141), 206–209.

Murray, S. (2013). *Translucent building skins: material innovations in modern and contemporary architecture*. Routledge.

Nagai, T., Ono, Y., Tani, Y., Koida, K., Kitazaki, M., & Nakauchi, S. (2013). Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study. *i-Perception*, *4*(6), 407–428.

Narasimhan, S. G., Gupta, M., Donner, C., Ramamoorthi, R., Nayar, S. K., & Jensen, H. W. (2006). Acquiring scattering properties of participating media by dilution. In *Acm siggraph 2006 papers* (pp. 1003–1012).

Niedenthal, S. (2002). Learning from the cornell box. *Leonardo*, *35*(3), 249–254.

Nishida, S., Motoyoshi, I., Nakano, L., Li, Y., Sharan, L., & Adelson, E. (2008). Do colored highlights look like highlights? *Journal of Vision*, *8*(6), 339.

Norman, J. F., Todd, J. T., & Orban, G. A. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science*, *15*(8), 565–570.

Nunes, A. L., Maciel, A., Meyer, G. W., John, N. W., Baranoski, G. V., & Walter, M. (2019). Appearance modelling of living human tissues. In *Computer graphics forum* (Vol. 38, pp. 43–65).

Pellacini, F., Ferwerda, J. A., & Greenberg, D. P. (2000). Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of the 27th annual conference on computer graphics and interactive techniques* (pp. 55–64).

Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision research*, *41*(24), 3145–3161.

Poggio, T., & Koch, C. (1985). Iii-posed problems early vision: From computational theory to analogue networks. *Proceedings of the Royal society of London. Series B. Biological sciences*, *226*(1244), 303–323.

Pointer, M. (2003). Measuring Visual Appearance- A Framework of the Future. Project 2.3 Measurement of Appearance. *NPL Report: COAM 19*.

Pointer, M. (2006). A framework for the measurement of visual appearance. *CIE Publication. CIE 175:2006 ISBN: 978 3 901906 52 7*.

Preston, C., Xu, Y., Han, X., Munday, J. N., & Hu, L. (2013). Optical haze of transparent and conductive silver nanowire films. *Nano Research*, *6*(7), 461–468.

Prokott, K. E., & Fleming, R. W. (2019). Predicting human perception of glossy highlights using neural networks [abstract]. *Journal of Vision*, *19*(10), 297b.

Ray, A., & Roberts, J. R. (2013). The structural basis of egg shell translucency and its role in food safety of table eggs. In *Proceedings of the 24th annual australian poultry science symposium* (p. 162).

Rushmeier, H. (1995). Input for participating media. In *In realistic input for realistic images (1995), acm press, acm siggraph '95 course notes. also appeared in the acm siggraph '98 course notes - a basic guide to global illumination.*

Rushmeier, H. (2008). Input for participating media. In *Acm siggraph 2008 classes* (pp. 1–24).

SABIC Innovative colorXpress. (n.d.). *Transparency and translucency.* (Accessed on 16/10/20 at: https://www.sabic-ip.com/staticcxp/user/en/LearnAboutColor/ColorBasicsDetail/scatter_opacity_transparency_translucency.html)

Sawayama, M., Adelson, E. H., & Nishida, S. (2017). Visual wetness perception based on image color statistics. *Journal of Vision*, *17*(5:7), 1–24.

Sawayama, M., Dobashi, Y., Okabe, M., Hosokawa, K., Koumura, T., Saarela, T., et al. (2019). Visual discrimination of optical material properties: a large-scale study. *BioRxiv*, 35 pages.

Sayim, B., & Cavanagh, P. (2011). The art of transparency. *i-Perception*, *2*(7), 679–696.

Schlüter, N., & Faul, F. (2014). Are optical distortions used as a cue for material properties of thick transparent objects? *Journal of Vision*, *14*(14:2), 1–14.

Schlüter, N., & Faul, F. (2019). Visual shape perception in the case of transparent objects. *Journal of Vision*, *19*(4:24), 1–36.

Schmid, A. C., Barla, P., & Doerschner, K. (2020). Material category determined by specular reflection structure mediates the processing of image features for perceived gloss. *bioRxiv*, 45 pages. (bioRxiv preprint doi: https://doi.org/10.1101/2019.12.31.892083, 2020)

Seghi, R. R., Hewlett, E., & Kim, J. (1989). Visual and instrumental colorimetric assessments of small color differences on translucent dental porcelain. *Journal of Dental Research*, *68*(12), 1760–1764.

Sharan, L., Rosenholtz, R., & Adelson, E. (2009). Material perception: What can you see in a brief glance? [Abstract]. *Journal of Vision*, *9*(8), 784.

Singh, M. (2020). Transparency and translucency. *Computer Vision: A Reference Guide*, 1–5.

Singh, M., & Anderson, B. L. (2002a). Perceptual assignment of opacity to translucent surfaces: The role of image blur. *Perception*, *31*(5), 531–552.

Singh, M., & Anderson, B. L. (2002b). Toward a perceptual theory of transparency. *Psychological review*, *109*(3), 492–519.

Spence, C. (2020). Shitsukan—the multisensory perception of quality. *Multisensory Research*, *1*(aop), 1–39.

Stanford University Computer Graphics Laboratory. (1994). *The Stanford 3D Scanning Repository.* (http://graphics.stanford.edu/data/3Dscanrep/)

Stevens, S. S. (1960). The psychophysics of sensory function. *American Scientist*, *48*(2), 226–253.

Storrs, K. R., & Fleming, R. W. (2020). Unsupervised learning predicts human perception and misperception of specular surface reflectance. *bioRxiv*, 44 pages.

Tamura, H., Higashi, H., & Nakauchi, S. (2018). Dynamic visual cues for differentiating mirror and glass. *Scientific reports*, *8*(1), 1–12.

Tanaka, M., & Horiuchi, T. (2015). Investigating perceptual qualities of static surface appearance using real materials and displayed images. *Vision research*, *115*, 246–258.

Tavel, M. (1999). What determines whether a substance is transparent? for instance, why is silicon transparent when it is glass but not when it is sand or a computer chip? *Scientific American*. (Accessed on 11/12/2020 at https://www.scientificamerican.com/article/what-determines-whether-a/)

Thomas, J.-B., Deniel, A., & Hardeberg, J. Y. (2018). The plastique collection: A set of resin objects for material appearance research. *XIV Conferenza del Colore, Florence, Italy*, 12 pages.

Thomas, J.-B., Hardeberg, J. Y., & Simone, G. (2017). Image contrast measure as a gloss material descriptor. In *International workshop on computational color imaging* (pp. 233–245).

Todd, J. T., & Norman, J. F. (2003). The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure? *Perception & Psychophysics*, *65*(1), 31–47.

Todd, J. T., & Norman, J. F. (2019). Reflections on glass. *Journal of Vision*, *19*(4:26), 1–21.

Todo, H., Yatagawa, T., Sawayama, M., Dobashi, Y., & Kakimoto, M. (2019). Image-based translucency transfer through correlation analysis over multi-scale spatial color distribution. *The Visual Computer*, *35*(6-8), 811–822.

Tominaga, S., Kato, K., Hirai, K., & Horiuchi, T. (2017). Appearance reconstruction of fluorescent objects for different materials and light source. In *Color and imaging conference* (pp. 34–39).

Torgerson, W. S. (1958). Theory and methods of scaling. 1958, Wiley: New York.

Urban, P., Tanksale, T. M., Brunton, A., Vu, B. M., & Nakauchi, S. (2019). Redefining a in RGBA: Towards a standard for graphical 3D printing. *ACM Transactions on Graphics (TOG)*, *38*(3), 1–14.

Van Assen, J. J. R., Barla, P., & Fleming, R. W. (2018). Visual features in the perception of liquids. *Current Biology*, *28*(3), 452–458.

Vargas, W. E., & Niklasson, G. A. (1997). Applicability conditions of the kubelka–munk theory. *Applied optics*, *36*(22), 5580–5586.

Vu, B. M., Urban, P., Tanksale, T. M., & Nakauchi, S. (2016). Visual perception of 3D printed translucent objects. In *Color and imaging conference* (pp. 94–99).

Wendt, G., Faul, F., Ekroll, V., & Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *Journal of Vision*, *10*(9:7), 1–17.

Wiebel, C. B., Valsecchi, M., & Gegenfurtner, K. R. (2013). The speed and accuracy of material recognition in natural images. *Attention, Perception, & Psychophysics*, *75*(5), 954–966.

Wijntjes, M., Spoiala, C., & De Ridder, H. (2020). Thurstonian scaling and the perception of painterly translucency. *Art & Perception*, *1*(aop), 1–24.

Wilson, A. D., & Kent, B. (1971). The glass-ionomer cement, a new translucent dental filling material. *Journal of Applied Chemistry and Biotechnology*, *21*(11), 313.

Xiao, B., Walter, B., Gkioulekas, I., Zickler, T., Adelson, E., & Bala, K. (2014). Looking against the light: How perception of translucency depends on lighting direction. *Journal of Vision*, *14*(3:17), 1–22.

Xiao, B., Zhao, S., Gkioulekas, I., Bi, W., & Bala, K. (2020). Effect of geometric sharpness on translucent material perception. *Journal of Vision*, *20*(7:10), 1–17.

Yang, L., & Kruse, B. (2004). Revised kubelka–munk theory. i. theory and application. *JOSA A*, *21*(10), 1933–1941.

Yu, H., Liu, P. X., & Hu, L. (2019). A highlight-generation method for rendering translucent objects. *Sensors*, *19*(4), 860:1–860:15.

Zhao, Y., & Berns, R. S. (2009). Predicting the spectral reflectance factor of translucent paints using kubelka-munk turbid media theory: Review and evaluation. *Color Research & Application*, *34*(6), 417–431.

Zuijlen, M. J. van, Pont, S. C., & Wijntjes, M. W. (2020). Painterly depiction of material properties. *Journal of Vision*, *20*(7:7), 1–17.

ISBN 978-82-326-5547-2 (printed ver.)
ISBN 978-82-326-5981-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology