Doctoral theses at NTNU, 2023:16

Altynay Kadyrova

Quality Assessment of 2.5D Prints

NTNU
Norwegian University of Science and Technology

Doctoral thesis

Altynay Kadyrova

Doctoral theses at NTNU, 2023:16

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 978-82-326-6503-7 (printed ver.)
ISBN 978-82-326-5726-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)

NTNU
Norwegian University of Science and Technology

NTNU

Altynay Kadyrova

Quality Assessment of 2.5D Prints

Thesis for the degree of Philosophiae Doctor

Gjøvik, January 2023

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 978-82-326-6503-7 (printed ver.)
ISBN 978-82-326-5726-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)

Doctoral theses at NTNU, 2023:16

PRINTED IN NORWAY
NO - 1598

NORDIC SWAN ECOLABEL
Printed matter
2041 0731

Printed by Skipnes Kommunikasjon AS

Abstract

Technology is developing rapidly around the globe and advancing its capabilities in every industry. It can provide users with high-tech solutions to complex tasks. With the help of technology, industry, presently, can acquire natural-looking reproductions of a variety of materials and objects. Thus, technology allows to create a replica that is almost identical to the original version. The printing industry uses its technologies to provide customers with visually-appealing products.

Elevated printing technology allows to add an elevation (i.e., height) to a 2D print. Such 2D print with surface elevation is called a 2.5D print. 2.5D printing (also called elevated printing) can fabricate natural-looking reproductions. Examples of 2.5D printing applications are packaging, signage, interior design, to name a few. It also has a high-end application area such as reproducing famous paintings for a variety of museums. Thus, it is essential to perform quality assessment to obtain high-quality 2.5D prints (reproductions). In other words, the quality assessment of 2.5D prints holds an important value for industry.

Overall, quality assessment plays a vital role in most application areas. More specifically, both subjective and objective quality assessment are crucial in quality inspection of products.

In this thesis, quality assessment of 2.5D prints both subjectively and objectively was investigated. For subjective quality assessment, several visual experiments were conducted where observers were given physical 2.5D prints for assessment. The research started with investigation of what quality attributes are relevant for quality assessment of 2.5D prints. The most used distinct quality attributes were found and they were proposed to be considered as the relevant ones for quality assessment of 2.5D prints.

In order to be of high quality, prints need to have a natural look of the content they describe. Therefore, naturalness perception of 2.5D prints by focusing on effects of several relevant parameters was investigated. More specifically, the effect of elevation and surface roughness, the effect of elevation, and the effect of several ink types on naturalness perception of 2.5D prints were considered. There was impact by elevation on naturalness perception of 2.5D prints. It was found that the effect of elevation

was content specific. Further, it was found that the optimal elevation that makes perception of naturalness of 2.5D prints was also content specific. In addition, 2.5D prints with matt inks were found to be perceived as more natural than with glossy inks by observers for specific content of images.

For objective quality assessment, capability of existing metrics to assess quality of 2.5D prints was investigated. It was found that the tested metrics were capable to work with 2.5D prints under certain conditions.

The research carried out in this thesis can benefit industry in designing both subjective and objective guideline(s) for quality assessment of 2.5D prints. This in turn can facilitate production of the high-quality and natural-looking 2.5D prints.

Abbreviations

ASTM	American Society for Testing and Materials
CIEDE2000	Commission Internationale de l'Éclairage Delta Empfindung 2000
FR	Full Reference
GDPR	General Data Protection Regulation
HVS	Human Visual System
iCID	improved Color Image Difference
IJC	Ink Jet Consumable
IQMs	Image Quality Metrics
ISO	International Organization for Standardization
MS-SSIM	Multi-Scale Structural Similarity Index
NR	No Reference
NSD	Norwegian Center for Research Data
QA	Quality Assessment
RR	Reduced Reference
RTAAV	Royal Talens Amsterdam Acrylic Varnish
RTI	Reflectance Transformation Imaging
S-CIELAB	Spatial-CIELAB
SSIM	Structural Similarity Index

Contents

Figures

Tables

Part I: Introduction to the research in the thesis

Chapter 1

Introduction

This chapter introduces the thesis's motivation, research goal, research questions, list of publications, ethical compliance, and thesis structure.

1.1 Motivation

Material reproduction has a considerable demand in various application areas such as medicine, education, and cultural heritage, to name a few. As a result, materials need to be reproduced realistically (i.e., have a natural look) with high quality. In other words, material reproduction and its Quality Assessment (QA) hold a significant value and importance. Various techniques such as 2.5D or 3D printing can be used to reproduce materials.

2.5D printing uses successive layers of ink to create a surface relief (i.e., elevation) [1]. In contrast to 2D printing, 2.5D printing can provide better reproduction of reflectance and surface coarseness [2]. Moreover, the quality of fine details can be better reproduced via 2.5D printing compared to 3D printing [1]. Therefore, there is a practical value of 2.5D printing in various application areas. A few examples of application areas where 2.5D printing has a common use are signage, interior design, cultural heritage, gadgets, packaging, and maps for visually impaired people.

In this thesis, 2.5D prints (Figure 1.1) were used and they were fabricated with Canon Arizona series flatbed printers to reproduce images of various materials. More specifically, QA of 2.5D prints was considered. To acquire high-quality and visually-appealing 2.5D prints, QA becomes essential. Hence, QA of 2.5D prints holds important value, especially for industry. Another important reason for working with QA of 2.5D prints is, to the best of our knowledge, a lack of standard protocols or guidelines for both subjective and objective QA of 2.5D prints.

Subjective QA can be done through visual or tactile experiments, or both where observers are involved. During the experiment, either digital

Figure 1.1: Example of 2.5D print. It shows the main feature of 2.5D print which is an elevated surface. Image: Canon Production Printing Netherlands.

images or physical prints, or both can be used and its outcome is the collection of subjective data. Objective QA can be performed in various ways: through measurements with instruments or devices, or through metrics. Metrics can be described as algorithms that take a digital image as input and output, for example, a numerical value regarding the quality of the image. If one works with physical prints, then prints need to be captured (i.e., digitized) prior to inputting to metrics. One of the important criteria for metrics is to be able to mimic the Human Visual System (HVS) so that they could represent what people perceive. While subjective QA can be somewhat time consuming and might require some resources, objective QA can be less resourceful. Nevertheless, subjective QA is still needed as it provides subjective data which represent the HVS responses, and subjective QA is still considered as the "gold standard". Thus, both subjective and objective QA are necessary in most application areas.

In the case of 2.5D prints, guidelines for both subjective and objective QA of 2.5D prints are demanded in order to effectively perform QA of 2.5D prints. To make such guidelines, investigation of approaches for QA of 2.5D prints is a reasonable step as a start. It is challenging to benchmark a metric without reliable subjective data, and it is difficult to collect reliable subjective data without a clear guideline(s) on QA of 2.5D prints. As a result, there is a demand for both subjective and objective guidelines for QA of 2.5D prints. The acquired findings from the research presented in the thesis can lay a foundation for developing both subjective and objective guidelines for QA of 2.5D prints. By elevation, a height is meant that is a distance from substrate to the raised point in a print for the rest of the thesis unless specified otherwise.

Presently, customers might not be fully satisfied with the quality of 2.5D prints, and therefore, there is motivation to contribute into QA of 2.5D prints to improve quality and repeatability.

1.2 Goal and research questions

In this thesis, various research questions were addressed by mainly focusing on subjective QA and naturalness perception of 2.5D prints with observers. Objective QA of 2.5D prints was also considered.

1.2.1 Main research goal

The main research goal of the thesis was to investigate approaches for QA of 2.5D prints with a longer term goal of this being useful for producing high-quality prints. Approaches were investigated because, to our knowledge, there is no guideline for QA of 2.5D prints presently. Both subjective and objective QA of 2.5D prints were covered but with more focus on the subjective side. The reason for this was that it is valuable first to focus on subjective QA of 2.5D prints as it provides subjective data which can be used for objective QA. The acquired findings can be considered in developing guidelines for QA of 2.5D prints by industry. Several objectives were defined to reach the main goal as described below:

- Investigation of relevant quality attributes for QA of 2.5D prints. The first stage of QA is to know what quality attributes one needs to focus on. Therefore, it is important to know what quality attributes are the most relevant ones for a specific application. To our knowledge, relevant quality attributes for QA of 2.5D prints have not been studied.
- Investigation of naturalness perception of 2.5D prints. Naturalness perception can be complex and, to the best of our knowledge, naturalness perception of 2.5D prints has not been investigated.
- Investigation of applicability of the existing metrics for QA 2.5D prints. There is a limited number of studies that have focused on objective QA of 2.5D prints.

1.2.2 Research questions

The thesis aimed to answer the following research questions with regard to the research objectives described above:

1. What are the relevant quality attributes for QA of 2.5D prints? (addressed in **Article A**)
2. What is the effect of various parameters on naturalness perception of 2.5D prints? (addressed in **Articles B, C,** and **D**)
    - What is the effect of elevation and surface roughness on naturalness perception of 2.5D prints? (addressed in **Article B**)
    - What is the effect of elevation on naturalness perception of 2.5D prints? (addressed in **Article C**)
    - What is the effect of various ink types on naturalness perception of 2.5D prints? (addressed in **Article D**)
3. Can existing metrics assess quality of 2.5D prints? (addressed in **Article E**)

The overview of the research questions with regard to the articles is illustrated in Figure 1.2. It consists of subjective QA and objective QA parts. The former part focused on quality and naturalness aspects of 2.5D prints while the latter part focused on metrics. There was an industry-based workflow in the research conducted in the thesis.

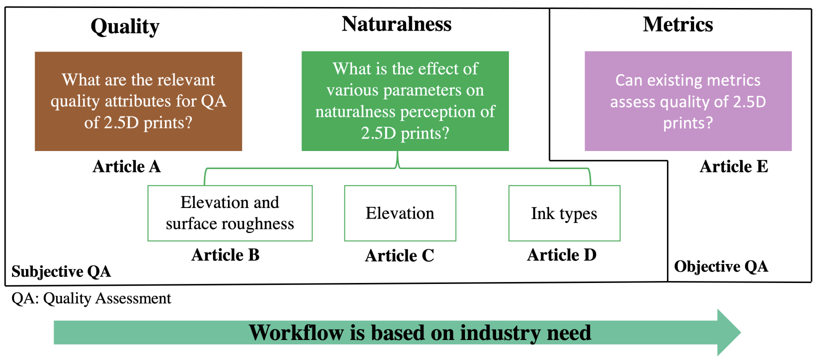

Figure 1.2: Overview of the research questions with regard to the articles.

1.3 List of publications

The thesis is based on five publications. They are enumerated alphabetically. There are four journal articles and one conference article.

Article A Kadyrova, A., Kitanovski, V., Pedersen, M. (2020, November). A study on attributes for 2.5D print quality assessment. In Color and Imaging Conference, Society for Imaging Science and Technology, vol. 2020, no. 28, pp. 19-24.

Article B Kadyrova, A., Pedersen, M., Westland, S. (2022, May). Effect of elevation and surface roughness on naturalness perception of 2.5D decor prints. Materials, 15(9), 3372.

Article C Kadyrova, A., Pedersen, M., Westland, S. (2022, May). What elevation makes 2.5D prints perceptually natural? Materials, 15(10), 3573.

Article D Kadyrova, A., Pedersen, M., Westland, S., Weijkamp, C. Effect of various ink types on naturalness perception of 2.5D prints. Under review in a journal, 18 pages.

Article E Kadyrova, A., Kitanovski, V., Pedersen, M. (2021, August). Quality assessment of 2.5D prints using 2D image quality metrics. Applied Sciences, 11(16), 7470.

In addition, one more article has been published but is not included as a part of the thesis.

Supporting Publication

Article F Kadyrova, A., Pedersen, M., Ahmad, B., Mandal, D. J., Nguyen, M., Zimmermann, P. H. (2022, January). Image enhancement dataset for evaluation of image quality metrics. Electronic Imaging, 34, 1-6.

1.4 Ethical compliance

In the research conducted in this thesis, ethical considerations were taken into account during the visual experiments involving observers regarding their personal data collection. In particular, approval from Norwegian Center for Research Data (NSD) for personal data collection and processing was acquired. All observers gave their consent for the participation in the experiments and their data collection. Observers' video data (in **Article A**) as well as audio data (in **Articles A, B, C,** and **D**) were collected. Their data were treated anonymously and were deleted at the end of the research.

1.5 Thesis organization

The thesis contains of two parts: a part consisting of introduction to the research in the thesis (Part I) and a part consisting of included articles (Part II). Part I has five chapters. Background of the research is given in Chapter 2 while summaries of each included articles with their interconnections are given in Chapter 3. Chapter 4 is dedicated for the discussion of the research. Chapter 5 provides conclusions and future perspectives. Part II presents all included articles.

Chapter 2

Background

This chapter provides background information on the relevant topics related to the research conducted in the thesis.

2.1 2.5D printing

2.1.1 Introduction to 2.5D printing

Multilayer structures that can be printed with radiation curable inks can have a specific range of height (0.1 mm - 10 mm) and that height gives extra dimension to a 2D image and such print is called 2.5D print [3]. The process for fabricating 2.5D prints is called 2.5D printing, relief printing, or elevated printing.

2.5D printing is a combination of technologies from 2D and 3D printing [4]. In practice, the surface layer can be printed with inks of a surface color while inks of other colors can be used for printing elevation thereby improving efficiency of the process of the 2.5D printing [5]. It is possible to create relatively stable gloss and matt appearance in spite of the ink amount [6].

Various companies work with 2.5D printing technology such as Canon, Mimaki, swissQPrint, Agfa, to name a few. In order to print images in 2.5D, the workflow is that one needs to send images with their height maps to a printer (Figure 2.1) and set the desired elevation level.

Various content images can be fabricated with 2.5D printing depending on application field. 2.5D printing can be used in graphic art applications (e.g., interior design (Figure 2.2)), applications for visually impaired people, art facsimiles, signage, packaging, elevated maps (Figure 2.3), architecture, and many more. It can also be used to create prototypes and molds.

To conclude, 2.5D printing has an important role in various application fields and it has promising future.

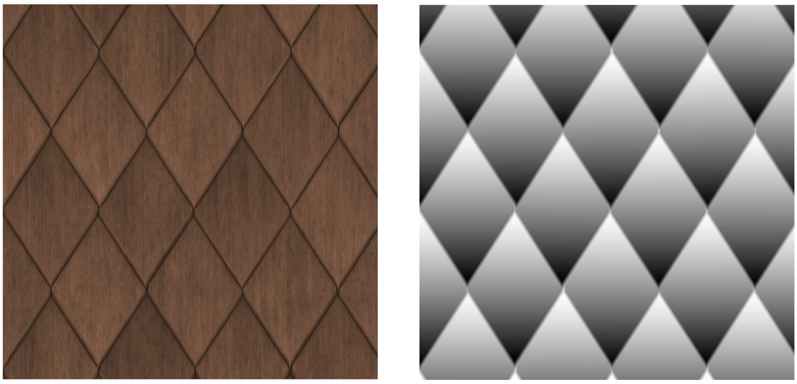

Figure 2.1: Example of input images to a printer. Color image is on the left and height map is on the right. In height map, black areas have maximum elevation set by user, white areas are not elevated, and gray areas have in between elevation levels. The images were used in **Articles C** and **D**.

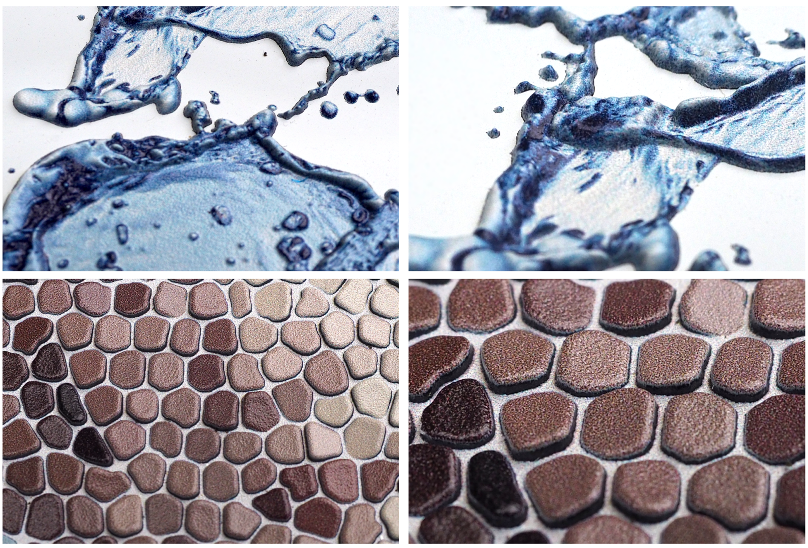

Figure 2.2: Examples of 2.5D prints for interior design application. Images: Canon Production Printing Netherlands.

Figure 2.3: Example of 2.5D print for elevated map application. Image: Canon Production Printing Netherlands.

2.1.2 2.5D printing challenges in industry

There exist challenges with 2.5D printing in terms of design, printing, and naturalness aspects. Some of them are discussed below.

As design of 2.5D print is highly dependent on image content and objects it contains, many possible artefacts may occur. Also, it is important to keep in mind that various rendering tools may not give perfect view on how exactly the design will look like after printing. Therefore, trial and error printing processes may help to make right decision or develop a common sense of how a rendered image possibly will look like after printing.

The possible challenges during the design stage of 2.5D prints can be appearance of various artefacts. If the intended elevation for an object in an image is quite high, then color quality of steep edges might be an issue. More specifically, color of the elevation (i.e., a wall from background to the elevated edge) becomes black at a higher elevation (Figure 2.4). According to industry, the black wall becomes visible at around 0.2 mm and gets worse at 1 mm or higher. In order to solve the black edge problem, different filtering tools can be an option. Also, if the slope is 60° or less, it can become less of a problem. Weijkamp and Valade [5] described a method to prevent dark edges. More specifically, when the elevation is printed, it tends to become dark gray, and therefore white ink is used to cover dark edge and afterwards transparent and color inks are used to create a colored edge on the surface layer.

Another problem may occur after solving black edges such as the transition can be distributed partly in object and partly in background sides and the edge color may come from background color while it is expected to be from the object color (Figure 2.5). To get around this problem, different morphological operations performed on gray scale height map images can help. If the black edges continue to be present, then the intended elevation level can be reduced to deal with the artefacts. Furthermore, industry experts recommend working with large size images and then reduce designs on preferred image size for printing. In other words, larger size images tend to be relatively robust to design flaws. It is important to note that printing at higher elevation can be slow.

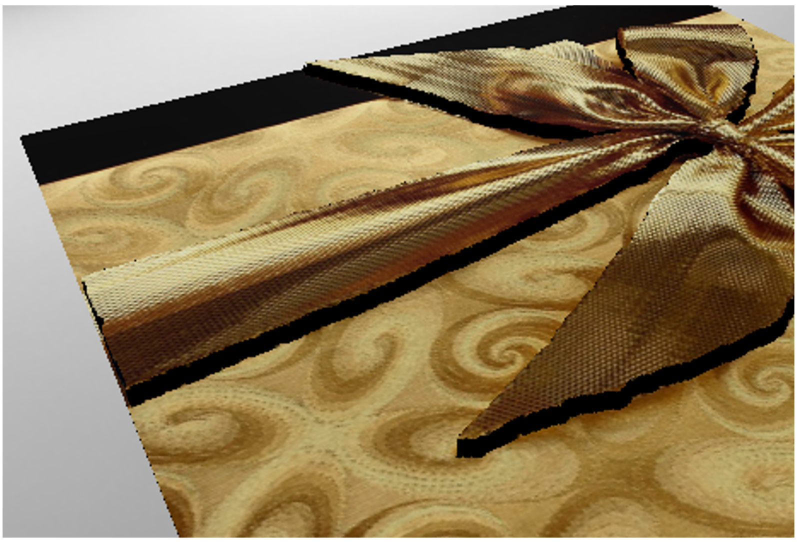

Figure 2.4: Example of black wall appearance in height, rendered in Canon Touchstone software. The image was used in **Articles A** and **E**.

In addition, it can be a good practice to have few elevation levels in images (preferably, maximum elevation (e.g., object) and no elevation (e.g., background)) (Figure 2.6) to save materials and reduce printing time in order to be more ecologically friendly. It can also help to save costs. However, images that need to be printed with 2.5D effect may widely vary in terms of content and number of objects they contain, and, for creativity sake, several elevation levels per image might be needed (Figure 2.7). Thus, trade-

Figure 2.5: Example of edge colors coming from background and object, rendered in Canon Touchstone software at 1 mm elevation and zoomed for illustration purpose. The height map of the image was filtered to reduce effect of black wall and intensity adjusted to reach the intended maximum elevation, and morphological operation was applied to the height map (right side version). The image was used in **Article B**.

offs between the intended effect and technical set-ups should be decided beforehand.

Overall, the design stage can influence the quality of 2.5D prints to some extent. In addition to the design aspects, the printing process can also be one of the contributors to the quality of 2.5D prints.

Another challenge with 2.5D prints is related to naturalness aspect. More specifically, it is important to produce prints that have a natural look. Sometimes 2.5D prints might have a plastic look. Therefore, it is reasonable to start with investigating what parameters determine naturalness of 2.5D prints. The substrate type (e.g., plastic, glass, metal, etc.) may impact the final quality of 2.5D prints. However, the substrate effect on quality is not considered in the thesis. It is critical but not unduly so.

When printing 2.5D print, which is a stack of several ink layers, the final ink tends to be brittle and consequently can be cracked [7]. There is also issue with colors of cutting edges during cutting of 2.5D prints [7]. Without using support ink, 2.5D technology based on inkjet printers cannot print overhangs. Available 2D printers that work with 2.5D technology, for example Canon Arizona series flatbed printers with Touchstone software, do not offer support ink and cannot print overhangs.

To conclude, 2.5D printing comes with its challenges and they need to be addressed to provide high-quality and natural-looking 2.5D prints.

2.2 Quality assessment

2.2.1 What is quality?

In different fields, quality can have different definitions. Thus, it is a complex concept [8]. A common understanding of quality is that it can be con-

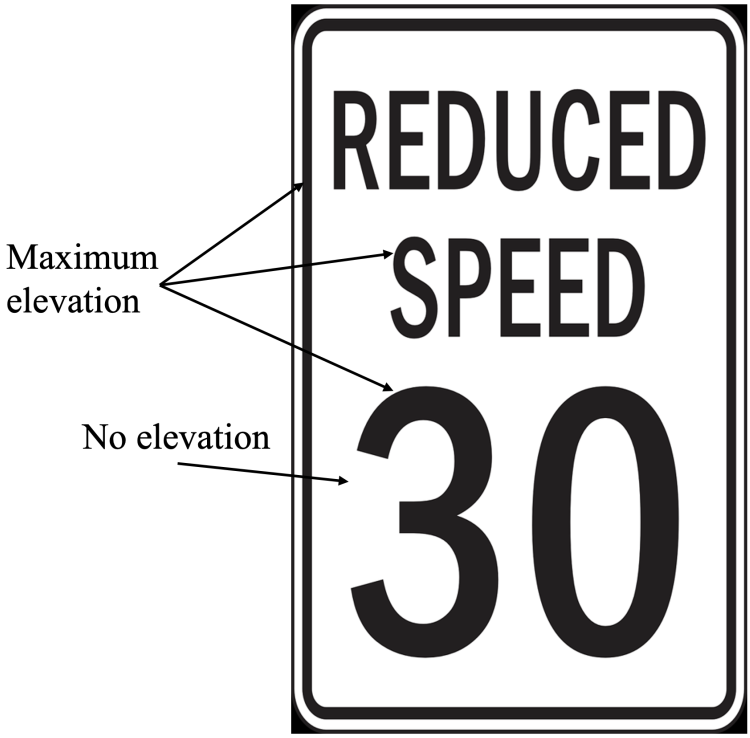

Figure 2.6: Example of height map with two elevation levels. Black areas have maximum elevation set by user while white areas are not elevated. The image was used in **Articles A** and **E**.

sidered as a correspondence to a set of criteria set by industry or organization. The quality can be related to images, prints, objects, and other products.

According to Silverstein and Farrell [9], image quality can be deduced from preference of one image compared to another one. The image quality can also be described as an image excellence perceived by an observer where the observer did not take part in anything related to the image [10, 11].

Kipphan [12] mentioned several set of influences and specifications that define a print's quality. More specifically, print (e.g., technology, material and ink flow), prepress (e.g., scanning, calibration), postpress (e.g., coating, folding), and material (e.g., ink, paper) can influence quality of the print while color (e.g., color coordinates, optical density), resolution

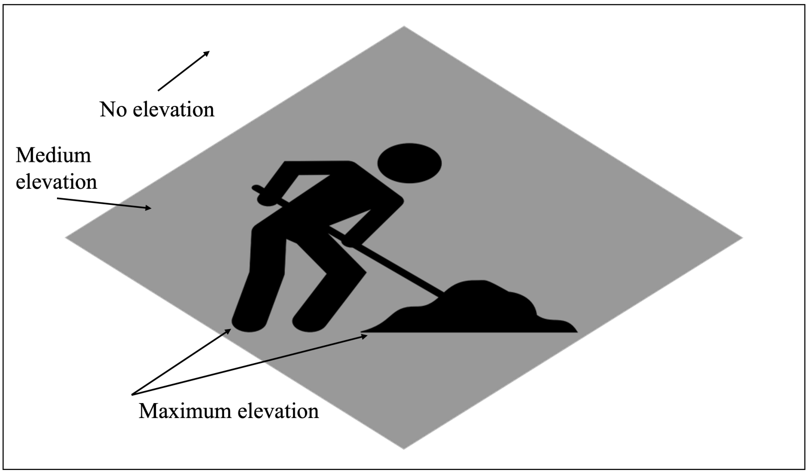

Figure 2.7: Example of height map with three elevation levels. Black areas have maximum elevation set by user, white areas are not elevated, and gray areas have in between elevation levels. The image was used in **Article A**.

(e.g., sharpness, tone value range), register (e.g., dot/color separation position, printed image position), and surface (e.g., evenness, gloss) can specify quality of the print. Marcu [13] highlighted parameters that affect the quality of color prints. They were reproduction technology, colorant media interaction, geometric resolution, color resolution, color separation with black generation mechanism, and tone reproduction. The reproduction technology (e.g., ink jet, laser, offset printing) was described as a key parameter in quality of the print. Colorant media interaction (e.g., coated, plain paper, glossy) was described as a parameter that defines actual color gamut achievable on a specific device.

Either image quality or print quality leads to product quality that will be used by customers. Garvin [8] suggested to have different approaches to quality when products shift from design to market and provided five approaches to determine quality. The five approaches were transcendent, product based, user based, manufacturing based, and value based. Further, he identified eight dimensions of quality: performance, features, reliability, conformance, durability, serviceability, aesthetics, and perceived quality.

Because different definitions of quality can have both strengths and shortcomings in terms of customer relevance, generalizability, and managerial usefulness, there is no best definition of quality that can work in every situation and one needs to explore trade-offs between definitions of quality to select relevant definitions [14].

To conclude, quality is an important term and one needs to define it in advance to produce high-quality products.

2.2.2 Quality attributes

According to International Organization for Standardization (ISO) [10], attribute is a term that can be used to mean component or aspect of image quality.

The quality attributes have utmost importance in a variety of fields. Depending on application field, they can be called as, for instance, quality attributes, perceptual attributes, appearance attributes, or just attributes. There can be differences between these terms, however, they can also be considered as synonymous to certain degree. The reasoning for this is that appearance attributes can be perceivable and therefore can be called as perceptual attributes. Once perceived, attributes can lead to certain quality impression or judgement and hence can also be called as quality attributes. Because this research was focused on QA, the term quality attributes was used unless specified otherwise (e.g., in **Article B**, the term perceptual attributes was used).

A diverse set of quality attributes exist in literature. They can define quality and help in quantification of the quality. The quality attributes need to be optimized to find the most relevant ones for specific applications. Otherwise, it is challenging to create Image Quality Metrics (IQMs) that can estimate quality of all attributes [11]. As a result, combined effects of different quality attributes have been studied [15, 16].

The quality perception of a printed material or object is affected by numerous attributes. The top five relevant quality attributes for QA of 2.5D prints were found to be color, sharpness, elevation, lightness, and naturalness [17] (**Article A**). Lightness, contrast, color, sharpness, and artefacts were found to be important quality attributes for 2D color prints [18]. It is worth to give a brief description about color - the most common quality attribute that is used in everyday life.

Color is usually a dominant quality attribute in many application fields. It has its own set of attributes (i.e., components) such as hue, brightness, lightness, colorfulness, chroma, saturation, and related and unrelated colors [19]. The color combinations can create a diverse set of effects on perception. The color combinations that can create emotional response produce warm colors, cool colors, hot colors, and cold colors, to name a few [20]. Additionally, micro structures in combination with the light can affect the color thereby producing structural color [4]. The structural color consists of colloidal crystals that have unique color and gloss effects, and it is not easy to reproduce printed colloidal crystals by conventional printing technologies [4].

To conclude, the quality attributes have a vital importance in various application fields. The quality attributes and their combinations can be used in both subjective and objective QA.

2.2.3 Subjective quality assessment

Subjective QA involves human observers and can be performed either visually or with tactility, or both through psychophysical experiment. Although it can be time consuming and costly, its main advantage is that its data represent (i.e., subjective data) the HVS responses. During experiments, consent of observers has to be acquired following General Data Protection Regulation (GDPR) [21]. There are other human subject protection declarations such as the Declaration of Helsinki [22] and Massachusetts Conduct of Human Subject Research Protocol [23] that should be followed in experiments with human observers.

Subjective visual QA has to be inline with ISO 3664 [24] requirements for viewing and illumination conditions. Following ISO 20462-1 [10], information on experimental conditions must be reported. Some of them are number of observers, criteria for observer selection, viewing distance, sample size, psychophysical method used, and others. They must be reported to aid the interpretation of results acquired from experiments [10].

There can be controlled (e.g., laboratory) and uncontrolled (e.g., online, field) experiment types. The main difference between the controlled and the uncontrolled experiments is that, for example, illumination and viewing conditions are not standardized in the latter. The viewing conditions can be critical in QA, and thus need to be carefully treated. For example, the viewing distance from an observer to a print can reveal how micro and macro structural textures and details are perceived [4].

It is important to check visual acuity and color vision of observers before the experiment. Various visual acuity and color vision tests can be used for the experiment. A Snellen chart for visual acuity test and Ishihara plates for color vision test were used in **Articles A, B, C,** and **D**. In Figure 2.8, several color vision and visual acuity tests that can be used for the experiment are listed.

The choice of the right psychophysical method for specific application area has a vital importance. There are number of psychophysical methods to choose from. Visual experiments can be grouped into threshold, matching, and scaling experiments [19]. The threshold methods contain method of adjustment, method of limits, and method of constant stimuli. The matching methods can be divided as asymmetric and memory matching. The scaling methods can be classified as one dimensional and multidimensional. The multidimensional scaling methods can have metric and non-metric forms. The one-dimensional scaling methods are rank order,

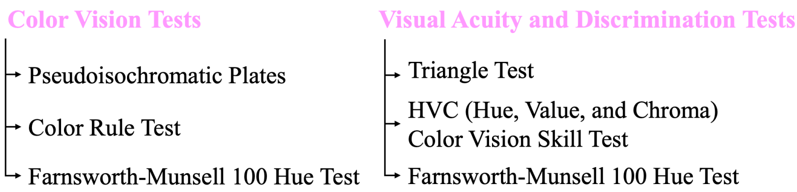

Figure 2.8: Recommended color vision and visual acuity tests for the selection and evaluation of observers, summarized from American Society for Testing and Materials (ASTM) [25].

graphical rating, category scaling, paired comparison, partition scaling, magnitude estimation (or production), and ratio estimation (or production). There are other methods that can be used for QA. For example, mean opinion score [26], double stimulus impairment scale [27], double stimulus continuous quality scale [27], triplet comparison (ISO 20462-2) [28], and quality ruler (ISO 20462-3) [29]. Although the mean opinion score was used first in speech quality, it has also been used in other domains [30].

The paired comparison method is simple and little knowledge from observers is needed [31]. The rank order method can be viewed as a fairly simple method for an observer and this leads to fairly high observer repeatability [32]. Also, it can be viewed as doing paired comparison method of all images at once [31]. Due to the subjectivity of the judgements, results of the category judgement experiment tend to be dependent on observers [31]. Unlike with paired comparison method, experiments with the category judgement method are often faster and therefore it is suitable for experiments with large number of samples [31]. Nussbaum [32] mentioned that rank order and paired comparison are the most important psychophysical methods for image reproduction field.

Observers tend to be stricter during assessment of 2.5D or 3D prints while they can tolerate on quality of 2D prints [4]. A brief summary of some studies on subjective QA of 2.5D prints is provided below.

The quality of 2.5D prints was assessed by observers in a visual experiment to derive the relevant quality attributes for QA of 2.5D prints [17] (**Article A**). Baar et al. [33] studied impact of gloss on perceived texture and texture on perceived gloss of 2.5D printed surfaces. They found a slight impact of texture on gloss perception.

To conclude, subjective QA is inevitable part of QA workflow. Therefore, it is still used and will continue to be used in subjective QA of 2.5D prints and other products.

2.2.4 Objective quality assessment

Objective QA can be performed through metrics or measurements. It is somewhat faster than subjective QA though some measurements might be time consuming.

There are variety of 2D IQMs and they can be grouped into various categories based on application area, approaches used, attributes, availability of reference images, to name a few. The IQMs can have three types depending on reference image availability: Full Reference (FR) where reference image is available, Reduced Reference (RR) where some information on reference image is available, and No Reference (NR) where reference image is not available. Seshadrinathan and Bovik [34] classified QA techniques into three groups: HVS modeling based approaches, structural approaches, and information theoretic approaches. The IQMs were classified as mathematically based, low level, high level, and other groups [35]. Depending on what information IQMs use, they were also classified as pixel difference based, correlation based, edge based, spectral distance based, context based, and HVS based [36].

To the best of our knowledge, there is no standardized IQM for QA of 2.5D prints. To assess quality of 2.5D prints objectively, existing 2D IQMs were tested and it was found that the tested IQMs can work with 2.5D prints under certain conditions [37] (**Article E**). Baar et al. [38] mentioned multiple approaches that can be considered to make an IQM for QA of 2.5D prints. Objective QA of 2.5D printing process focusing on height information was studied in terms of different aspects such as fidelity and surface finish by Liu et al. [39].

In the case when one is working with physical prints (e.g., 2D, 2.5D), they need to be captured (i.e., digitized) or measured through specific instruments or devices. This is needed to objectively assess quality of the prints. Measurement instruments can output value(s) which can give indication to a user about quality of the prints or can give information on various attributes' level in the prints. Another alternative is to input captured digital data into IQMs.

Measurement science is called metrology [40]. There are important terms related to measurements such as accuracy and precision. In simple words, accuracy can be explained as a match of a measured value with a defined standard value. Precision can be described by repeatability and reproducibility [41]. The repeatability means that the same instrument produces the same value on the same sample when measured at different times. The reproducibility means to acquire the same results with the same instrument at different locations.

The instruments or capture set-ups vary depending what needs (e.g., 2D prints, 2.5D prints, 3D objects or color, gloss) to be measured or captured. For example, spectral reflectance of object can be measured by a spectrophotometer while optical radiation power as a function of wavelength can be measured by a spectroradiometer [42]. Glossmeters can be used to measure gloss. Gloss measuring procedure has to follow various standards, for example, ISO 2813 [43] for paints and varnishes and ISO 8254-1 [44], ISO 8254-2 [45], ISO 8254-3 [46] for paper and board. Reflectance Transformation Imaging (RTI) can be used to capture color and surface texture of objects under multiple illumination angles [4]. There are many techniques that can be considered for capturing and measuring such as laser triangulation scanning, goniospectrophotometry, photogrammetry, and multifocal microscopy, to name a few.

Different techniques can be combined and hybrid set-ups can measure or capture more complex surfaces and objects. For instance, a camera and a linear array of light-emitting diode were used to capture spatially-varying gloss of hand-painted samples [47]. Three 3D scanning techniques were compared to capture painting's surface topology [48]. They were 3D digital microscopy, multi-scale optical coherence tomography, and 3D scanning based on fringe-encoded stereo imaging.

Higher dimensional objects need somewhat complex capture set-ups to be able to capture specific attributes of objects with a high accuracy to digitally reproduce objects. Presently, there is no tool that reproduces (i.e., captures or fabricates) full appearance (i.e., all attributes) of object [49] and complex surface [4].

To conclude, objective QA can save time and resources, and it can provide consistent results. Therefore, integrating objective QA into QA workflow is demanded. Objective QA techniques (e.g., IQMs, capture set-ups) are application and product dependant.

2.3 Naturalness and its perception

Naturalness has a tendency to be multidimensional [50]. It can hold variety of definitions depending on the usage field. Hence, no standard definition of a natural product exists [51]. Overall, the term natural can be described as an attribute that enhances object's perception [52].

Naturalness perception may vary considerably from user to user because different users might interpret naturalness perception of specific content of images or prints, or objects in their own way based on their experience and knowledge. Thus, it is important to follow one or another definition of naturalness because there can be content dependency with regard to naturalness perception.

In general, naturalness can be viewed as a positive attribute [53]. It was found that perceptual image quality and naturalness have connection with each other [54]. Heynderickx [55] stated that naturalness can explain both perceived depth and perceived image quality. According to Halonen et al. [56], observers tend to take into account naturalness during overall image QA and naturalness can be considered as a high-level preferential quality attribute. Naturalness was one of the most used quality attributes during QA of 2.5D prints [17] (**Article A**).

Naturalness can incorporate various other attributes. Yoshida et al. [57] suggested that there can be dependence between naturalness and combination of other attributes. For example, naturalness measure was defined as the joint probability density function of brightness and contrast [58]. There can be a link between naturalness and changes in other quality attributes, namely color and lightness [18]. Furthermore, various attributes (e.g., naturalness, unnaturalness, real, unreal) can be used to describe image naturalness [56]. Based on this, it can be observed that naturalness perception needs to be studied by taking into consideration effects of other attributes regardless of application area. The experience and knowledge of users, content, cognitive understanding, aesthetic appearance, how samples (e.g., images or prints) were produced and processed, dimensions of samples, and other factors can impact naturalness perception. As an example, processing history affects naturalness of foods [52]. Higher the dimensionality of images or prints, higher the naturalness complexity becomes [59].

Customers might select products based on perception of naturalness of products. Therefore, naturalness and its perception seem inevitable to consider in various application areas. For instance, naturalness and its perception were explored in terms of food [52, 60], wood [50, 61], textile [62], water [63], 2.5D prints [59, 64, 65] (**Articles B, C,** and **D**), among others.

To conclude, the naturalness tends to be a complex concept and one needs to define it by considering effects of various parameters in order to acquire consistent and reliable data on naturalness perception.

Chapter 3

Summary of the included articles

This chapter provides a brief summary of the included articles. The articles can be found in Part II for more information.

3.1 Article A: A study on attributes for 2.5D print quality assessment

3.1.1 Objectives

The objective of this work was to find a set of relevant quality attributes for QA of 2.5D prints. Further, it was interesting to check whether observers use the same quality attributes when reference images were provided and not provided. In addition, collection of observations (i.e., feedback) on QA of 2.5D prints from observers was aimed.

3.1.2 Methodology

Fifteen images for this work were reproduced from Pixabay (copyright free web site) [66] and they represented a wide range of image quality aspects such as memory colors, large area of the same color, to name a few. Canon Touchstone software was used to design height maps and five sets of quality variations were introduced. A preliminary visual experiment was conducted to observe how observers assess quality of 2.5D prints. Based on preliminary experiment results, the main visual experiment was conducted on QA of 2.5D prints by using a rank order method and it had two parts. In the first part, the observers were not provided with reference images while in the second part - the reference images were provided. There were forty-two 2.5D prints in total. Fifteen observers with various levels of experience, ethnicity, backgrounds, and both genders participated in the experiment and their consents were acquired before starting the experiment. The experiment was recorded and the observers were allowed to describe their rankings by their own words. The experiment was visual and tactility was not considered in this work. The experiment was carried out in English.

Frequency analysis was used to find the most used distinct quality attributes by observers. Sub-attributes were combined into quality attribute groups. In addition, inter-observer variability was checked and observations were collected from the recorded data on QA of 2.5D prints.

3.1.3 Results

The frequency analysis showed the most used distinct quality attributes with and without reference image cases. The top five were color, sharpness, elevation, lightness, and naturalness for both cases (i.e., with and without reference image). The inter-observer variability showed consistent usage of quality attributes between the observers. When the observers described their rankings with their own vocabulary, several observations were collected. The observations revealed that aesthetic appearance seems to be connected with high-quality perception. It was also found that the observers tend to pay attention on content and material type presented in the prints. The outcome of this work was a proposal to consider the most used quality attributes as the relevant ones for QA of 2.5D prints. Furthermore, dataset was another outcome of this work.

3.2 Article B: Effect of elevation and surface roughness on naturalness perception of 2.5D decor prints

3.2.1 Objective

The objective of this work was to investigate impact of elevation and surface roughness on naturalness perception of 2.5D decor prints.

3.2.2 Methodology

In this work, decor content was selected and the definition of naturalness by Drago et al. [67] was used which states that naturalness is the extent of similarity between image and realistic scene. Thus, the realistic term to mean naturalness was used. Four material category images were selected to work with, namely wood, stone, metal, and glass. There were five images in each material category and the images were reproduced from 3D textures (copyright free web site) [68]. Figure 3.1 illustrates the image processing steps used in this work.

Based on multiple test printings and assessments, various levels of elevation and surface roughness were created. A ranking experiment was car-

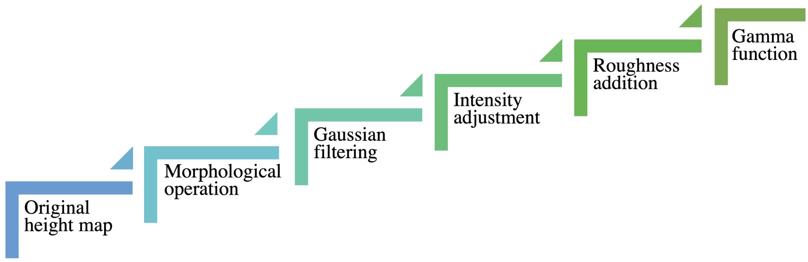

Figure 3.1: Steps (from left to right) of image processing used.

ried out on naturalness perception of 2.5D prints with twenty observers. There were one hundred and eighty 2.5D prints in total. For each set of prints for every image, a keyword as a reference was given which was the name of the material category. Similar to the experiment in **Article A**, the observers gave their consent before starting the experiment and tactility was not considered in this work. The observers were asked to provide an explanation to their rankings and the experiment was recorded. English was the used language for the experiment.

For analysis of results, both quantitative and qualitative methods were used. For quantitative analysis, Z-scores, binomial sign test with Bonferroni correction, and inter-observer variability check by using Spearman correlation coefficient were used. For qualitative analysis, collected audio data of the observers were transcribed and extracted attributes were grouped. Frequency analysis was used to identify the most used perceptual attributes.

3.2.3 Results

There was impact by elevation on naturalness perception of 2.5D decor prints and it was found to be linked with content. Moreover, lower elevation was found to be perceived as natural for 2.5D prints of wood and glass images by the observers whereas no clear tendency was found for 2.5D prints of stone and metal images. The perceptual attributes that one tends to use during naturalness assessment of 2.5D decor prints were identified. Color, roughness, gloss, elevation, and lightness were the top five ones. In addition, it was observed that change in one or more attributes can create perception of variation of other attributes connected with naturalness of the prints.

3.3 Article C: What elevation makes 2.5D prints perceptually natural?

3.3.1 Objective

Based on the results of **Article B**, it was found that the observers perceived lower elevation as natural for 2.5D prints of wood images. As a result, the objective was to find the exact elevation level that makes 2.5D prints of wood images to be perceived natural in this work.

3.3.2 Methodology

As in **Article B**, the realistic term for naturalness was used in this work. Twenty wood images which were reproduced from 3D textures [68] were selected. Within wood content, images of wooden floor, wall, roof, and wicker were included. Elevation levels were varied between 0 mm and 0.5 mm with a step of 0.1 mm. There were one hundred and twenty 2.5D prints in total. The visual experiment with a rank order method was conducted in the UK and Norway. Twenty-one observers participated in the experiment. The same experiment design used in **Article B** was followed. For a keyword reference, the name of wood content (e.g., wooden floor) was given in this work. Z-scores, binomial sign test with Bonferroni correction, and inter-observer variability check by using Spearman correlation coefficient were used for analysis.

3.3.3 Results

The results showed that the optimal elevation that makes 2.5D prints of wood images to be perceived as natural was content dependent and it was in a range between 0.3 mm and 0.5 mm. However, if to consider wood images regardless of content within wood images, then the optimal elevation was found to be 0.5 mm. It was also found that the observers perceived flat prints as the least natural. In addition, a high correlation was found between majority of observers on their rankings which showed that they were relatively consistent on naturalness perception of 2.5D prints of wood images.

3.4 Article D: Effect of various ink types on naturalness perception of 2.5D prints

3.4.1 Objective

The objective of this work was to investigate the effect of various ink types on naturalness perception of 2.5D prints.

3.4.2 Methodology

Twenty wood images used in **Article C** were used in this work. The realistic term was used for naturalness as in **Articles B** and **C**. Based on the results of **Article C**, it was found that the observers perceived 2.5D prints of wood images as natural at 0.5 mm. Therefore, 0.5 mm for all prints was used and ink types were varied. Commercially available Canon Ink Jet Consumable (IJC) UV-curable inks with various coatings were used. The experiment design used in **Article C** was followed. There were one hundred and twenty 2.5D prints in total and twenty-two observers participated in the experiment.

For analysis, Z-scores, binomial sign test with Bonferroni correction, inter-observer variability check and relation between observer ranking and gloss measurements by using Spearman correlation coefficient, and frequency analysis on each ink type selection by the observers through all images for each rank and on attributes extracted from audio data of the observers were used.

3.4.3 Results

Based on the results, it was found that 2.5D prints with matt inks were perceived as more natural than with glossy inks by the observers. The observers differed in naturalness perception of 2.5D prints based on their level of expertise. More specifically, high-medium-levels and low-level expertise observers had difference in perception of naturalness of 2.5D prints with various inks. Based on the frequency analysis on attributes extracted from the audio data, it was found that gloss and matt were the most used attributes. This shows that 2.5D prints had sufficient levels of glossiness as the observers were able to perceive the levels. The correlation between gloss measurements and observer ranking showed that some observers ranked based on the level of gloss. It also showed that some observers ranked, more likely, based on content.

3.5 Article E: Quality assessment of 2.5D prints using 2D image quality metrics

3.5.1 Objectives

The objective of this work was to test a set of IQMs whether they can assess quality of 2.5D prints. Because metrics require digital images as input, 2.5D prints needed to be digitized (i.e., captured). Therefore, addressing how 2.5D prints can be captured for QA was also another objective of this work.

3.5.2 Methodology

The 2.5D prints used in **Article A** were used in this work. Digital data of 2.5D prints were acquired through camera-based capture set-ups. Two set-ups were defined, namely single-shot and multiple-shot. The single-shot set-up captured 2.5D prints at one illumination angle while the multiple-shot set-up at multiple illumination angles. RTI was used for the multiple-shot set-up capture.

After capture, preprocessing of the captured images was performed. Relevant FR metrics (Structural Similarity Index (SSIM) [69], Multi-Scale SSIM (MS-SSIM) [70], improved Color Image Difference (iCID) [71], color difference metric - Commission Internationale de l'Éclairage Delta Empfindung 2000 (CIEDE2000) [72], and Spatial-CIELAB (S-CIELAB) [73]) were selected to apply to the captured data of 2.5D prints. For analysis, quality maps from the metrics, responsivity (i.e., whether the metrics can detect differences) of the metrics, and overall insights from the metrics were used, and the two capture set-ups were compared.

3.5.3 Results

Analysis on the responsivity of the metrics showed that CIEDE2000, S-CIELAB, and iCID were responsive to the quality variations between the prints in both set-up captures. As a result, to obtain insights on QA of 2.5D prints, the above-mentioned metrics were used afterwards. It was found that the tested metrics were informative on difference detection on various areas (e.g., on edges, surfaces of the elevated parts, background) in a similar way and the detection was better visible in the multiple-shot set-up captures than in the single-shot set-up ones.

Comparison of the two capture set-ups showed the dominance of the multiple-shot set-up over the single-shot set-up as the captured differences by the tested metrics were clearly visible in the former set-up than in the latter one. In other words, the multiple-shot set-up captures can better reveal elevation impact on appearance.

Chapter 4

Discussion

In this chapter, the articles are discussed in more detail along with their limitations. The articles are also discussed in terms of their relations with each other.

Based on the literature review, there was no study that clearly identifies a relevant set of quality attributes for QA of 2.5D prints. As a result, this was addressed in **Article A**. In **Article A**, a ranking experiment was carried out at a given illumination and viewing distance (also in **Articles B, C,** and **D**) due to the number of reproductions that were presented to observers. The experiment would have become long for observers which could make observers feel fatigue, if to use other methods (e.g., a paired comparison method). During the experiment, the observers were instructed to describe quality of 2.5D prints by their own lexicon. This approach is not standardized and the reason why the observers were allowed to describe their rankings with their own words was not to influence them to concentrate on a specific quality attribute.

The relevant quality attributes derived from **Article A** might be content dependent and be limited to the images used. However, the images were selected in the way that they cover a wide range of quality aspects (e.g., memory colors, neutral gray, large area of the same color) and present the most used 2.5D printing application areas (e.g., packaging, signage, paintings). The images used along with their height maps were released as a dataset [74]. The relation of the articles focused on subjective QA in the thesis is illustrated in Figure 4.1.

One of the top five relevant quality attributes found from **Article A** was naturalness (Figure 4.1). Moreover, naturalness aspect of 2.5D prints is industry relevant. As a result, naturalness perception of 2.5D prints was considered. Naturalness can be a complex concept and therefore the term realistic was used to mean naturalness in the following research.

To the best of our knowledge, naturalness perception of physical 2.5D prints has not been studied. Therefore, it was addressed in **Articles B, C,**

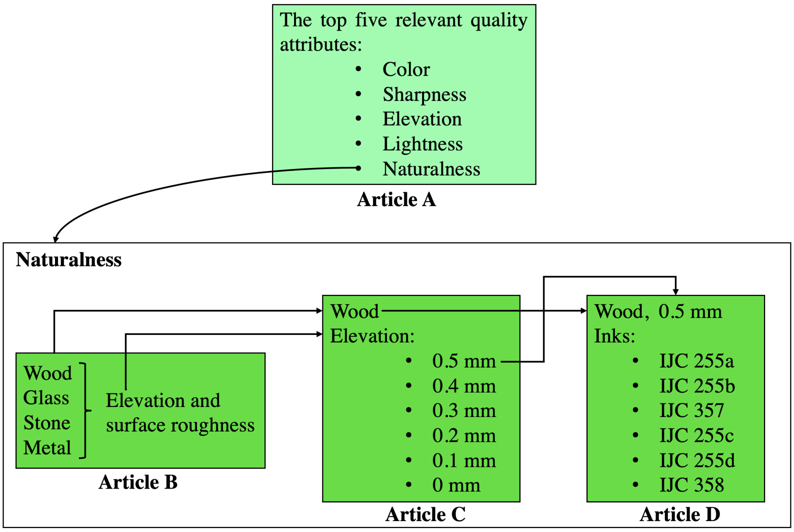

Figure 4.1: Schematic illustration of the relation of the articles focused on subjective QA.

and **D** and the effect of various parameters was considered. In **Article B**, the effect of elevation and surface roughness on naturalness perception of 2.5D decor prints of four material category images was considered (Figure 4.1). Elevation was selected because it is the main feature of 2.5D prints and it was in the top five most used quality attributes during QA of 2.5D prints (**Article A**). Surface roughness was selected because it can give a realistic appearance to prints. Wood, glass, stone, and metal materials were selected because of assumption that they, more likely, represent the most used decor materials.

The elevation levels used in **Article B** were 0.4 mm, 0.6 mm, and 0.8 mm, and the surface roughness was considered as the height difference within a local neighborhood. Number of variations per attribute was three and this can be a limitation. However, its increase could make the experiment duration longer which in turn could decrease performance of observers. During the ranking experiment, material category name (e.g., wood, glass) as a keyword reference was provided (also in **Articles C** and **D**, wooden content name (e.g., wooden floor, wicker) was provided). This was done to help observers to make easy judgements. In other words, it was assumed that observers can better assess naturalness of the prints if they know the material category.

Based on the results of **Article B**, it was found that elevation rather than surface roughness had impact on naturalness perception of 2.5D prints. More specifically, lower elevation was found to be perceived as natural for 2.5D prints of wood and glass images. However, what exact elevation that gives naturalness perception was not defined. This was addressed in the follow-up work in **Article C**.

In **Article C**, the effect of elevation on naturalness perception of 2.5D prints was considered. Wood images with content such as wooden floor, wall, roof, and wicker were used and elevation levels from 0 mm to 0.5 mm were varied (Figure 4.1) because the results of **Article B** showed that elevation had impact on naturalness perception of 2.5D prints and 2.5D prints of wood images were perceived as natural at lower elevation. Because lower elevation in **Article B** was 0.4 mm, it was interesting to investigate if it is 0.4 mm or elevation level around this number (i.e., 0.3 mm or 0.5 mm) that makes the prints to be perceived as natural. In addition, 0 mm, 0.1 mm, and 0.2 mm were included. It was found that the optimal elevation that makes perception of naturalness of 2.5D prints of wood images is content dependent and in a range between 0.3 mm and 0.5 mm. However, if to consider wood images regardless of content within wood images (i.e., wooden floor, wall, roof, wicker), 0.5 mm was found to be the optimal elevation. Because wood images only were used (also in **Article D**), this makes findings to be limited to this application. Nevertheless, **Articles C** and **D** supply a workflow that can be used for other applications.

Article D was a collaborative work with industry and it shows how acquired results can be applied in industry to some degree. In **Article D**, the effect of various ink types on naturalness perception of 2.5D prints was considered. Based on the results of **Article C**, 0.5 mm for all the prints and wood images used in **Article C** were selected to work with in **Article D** (Figure 4.1). Due to the industrial need, ink types (i.e., glossiness) with various coatings were varied. The ink types used in **Article D** are shown in Table 4.1.

The main outcome of **Article D** was that the observers perceived 2.5D prints of wood images with glossy inks as less natural while with matt inks as more natural. The same observation was acquired in **Article B** where the observers preferred 2.5D prints of wood images to be less glossy in order to be perceived as natural.

Definitions were not provided during the experiments in **Articles A, B, C,** and **D** because of the following reasons - not to influence observers with definitions on their preferences and it can take sometime for observers to understand definitions. Also, observers can be more consistent in their preferences if they can give explanations to their preferences [75]. Furthermore, physical 2.5D prints were used in **Articles A, B, C,** and **D** because digital images restrict interactivity.

It is worth to mention that, in **Articles A, B, C,** and **D**, all printing as-

Table 4.1: Ink types used in **Article D**. RTAAV - Royal Talens Amsterdam Acrylic Varnish. Both industrial information and gloss measurements were used to make glossy/matt distribution for the inks.

Inks	**Coating**	**Glossy/matt**
IJC 255a	Gloss coating (RTAAV) applied with 2 layers with brush	Glossy
IJC 255b	Gloss coating (RTAAV) applied with 1 layer with spray can	Glossy
IJC 357	Without coating	Glossy
IJC 255c	Canon IJC 257 varnish printed on top	Matt
IJC 255d	Matt coating (RTAAV) applied with 1 layer with spray can	Matt
IJC 358	Without coating	Matt

pects (e.g., calibration) were set to a printing company. Thus, the printing process was not controlled and there could be variances in the prints because of this. This is a drawback and it was done this way to focus on the subjective side of QA. It should be noted that this scenario is also more likely to happen from a customer side. The visual experiments in **Articles A, B, C,** and **D** were carried out in English and majority of the participated observers were non-native English speakers. Thus, there could be influence from the language on the performance of the observers when they were giving feedback on their rankings. Furthermore, number of observers and number of images in **Articles A, B, C,** and **D** varied between 15 - 22 and 15 - 20, respectively. The number for observers is according to the standards [27, 76] and sufficient number of observers was recruited and reasonable number of images was used. If to increase either number of observers or number of images, or both, then the experiment would have become long in duration. Thus, there is a trade-off between number of observers and number of images. Considering the number of images and number of reproductions per image, there were in total 2.5D prints between 42 - 180 in **Articles A, B, C,** and **D**. In the experiments conducted in **Articles A, B,** and **C**, both naive and experienced observers (e.g., holding computer science background, having knowledge of 2.5D printing to some degree) were involved. The observers who participated in the experiment conducted in **Article D** were experienced type of observers who had a good knowledge of 2.5D printing.

Both Asian and European observers were involved in the experiments conducted in **Articles A** and **C** whereas mostly Europeans were involved in the experiments conducted in **Articles B** and **D**. There was a relatively equal number of both genders in **Articles A** and **B**. There were more female observers in **Article C** while more male observers in **Article D**. Thus, it would be interesting to investigate how different gender and different ethnicity perceive naturalness of 2.5D prints and do QA in further works. The experiments were conducted in laboratories in **Articles A, B, C,** and **D**. Hence, they were controlled type of experiments. This could impact the results to some degree because it is common that customers see the prints outside of laboratories. Therefore, it would also be interesting to investigate the performance of observers on QA and naturalness perception of 2.5D prints in controlled versus uncontrolled environments. Although the first action observers would do when they see 2.5D prints can be to touch the surface of the prints, tactility was not considered in all conducted experiments because it tends to go more towards user interaction which is out of scope of the thesis. In practice, quality of the print tends to be assessed visually. Nevertheless, constraining to the visual assessment could impact the results and therefore tactility can be considered in further work. In addition, D65 simulator was used during the visual experiments in **Articles A, B,** and **C** while D50 simulator was used in **Article D** due to what was available at the company. It is necessary to select specific illumination, experiment type, observer type, whether it should be visual assessment or with tactility, and others to narrow down the scope of the research in order to acquire reliable results as it may not be feasible to test all experiment aspects (e.g., to conduct both controlled and uncontrolled experiment, to test several illuminations, to involve diverse ethnicity, and others) in just one experiment.

The above-mentioned articles focused on subjective QA. **Article E**, which focused on objective QA, investigated capability of existing metrics to work with 2.5D prints. In **Article E**, 2.5D prints from **Article A** were used. The aim was not to test all the metrics but relevant ones as a start. FR metrics were used because of availability of the reference images. FR metrics used in **Article E** were SSIM, MS-SSIM, iCID, CIEDE2000, and S-CIELAB. Color difference metric (CIEDE2000) was included to see how a pixel-based metric performs on QA of 2.5D prints and to compare CIEDE2000 (i.e., without spatial filtering of the HVS) with S-CIELAB (i.e., with spatial filtering). S-CIELAB operates on its spatial filtering to simulate spatial blurring of the HVS [73]. SSIM and MS-SSIM are structure-based metrics and most of the quality variations that the used images in **Article E** have impact the structure. iCID considers structure along with contrast and color of the reference image [71]. Furthermore, the prints used in **Article E** were fabricated with the same printer and material (i.e., substrate) and therefore it can be expected to see more structural differences between the prints, not much on the color aspect unless structural aspect leads to color difference. It was observed that more shadows appear due to the higher elevation. Larger color difference with higher elevation can be expected because of the shadow. It is not because color changes in the prints, it is because interaction with light casts larger shadow. Based on the specifics of the selected metrics, it was assumed that choice of the metrics was reasonable. In addition, NR metrics could also be tested because they work without a reference image. But FR metrics were used as a start and NR metrics can be considered in future work.

Since metrics work with digital images, physical 2.5D prints were captured with camera-based set-ups because it is cost effective and fairly fast. Because existing metrics which work mainly with 2D images were tested, the prints were first captured at one illumination angle and it was named as the single-shot set-up. As 2.5D prints were used and because their appearance can vary under various illumination angles due to the elevated surface (Figure 4.2), the prints were further captured with RTI based on a robotic arm [77] at multiple illumination angles and it was named as the multiple-shot set-up.

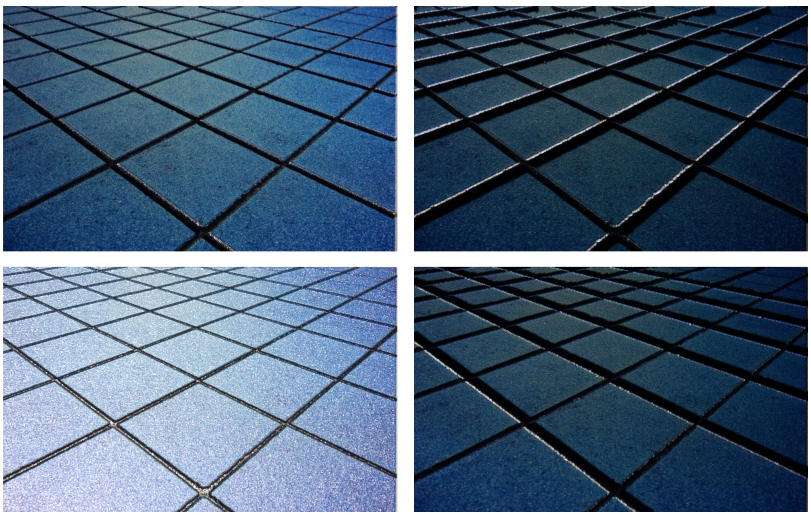

Figure 4.2: RTI capture example of the same 2.5D print under multiple illumination angles. The image was used in **Articles A** and **E**.

RTI was chosen to capture the prints at multiple illumination angles because it is a relevant tool for surface features' identification and can be used for various materials [78]. Although there can be relevant industrial solutions (e.g., [79–81]) to capture some appearance aspects of objects similar to 2.5D prints, they might be costly. Furthermore, RTI can have advantages compared to 3D scanning methods in terms of cost, accuracy, and range of size of objects and materials for capturing [4]. As a result, cost-effective tool was used which was available at the facility during the implementation time of **Article E**.

Since FR metrics were used, the captured data needed to be registered with the reference images. Image registration of the captured 2.5D images (i.e., captured images of 2.5D prints) with 2D reference images was found to be challenging due to these reasons: there was a dimension difference (i.e., 2.5D versus 2D); a small content loss and extra white space of the substrate were present in some prints due to cutting process and this introduced difficulties to register images, to name a few. To tackle the image registration challenge, various image registration methods were used and eventually desirable results via manual image registration were acquired.

When RTI captures were analyzed, it was challenging to find a single or a group of illumination angles that could provide insights on quality of 2.5D prints. As a result, the mean of various illumination angles was used. The results of **Article E** showed that iCID, CIEDE2000, and S-CIELAB were able to detect differences between the prints in both set-up captures. The responsiveness of the metrics to the differences was better visible in the multiple-shot set-up captures than in the single-shot set-up ones.

In other words, certain metrics and certain types of prints combined with one or another capture set-up can reveal interesting observations on QA of 2.5D prints. One might not be able to use just one metric and one acquisition set-up to measure everything. Thus, somewhat complex quality inspection system is needed to address all the quality aspects. But if one is interested only in one quality aspect, then one metric and one acquisition set-up might be sufficient.

Industry might prefer relatively simple over complex solutions. However, if application is high-end reproduction (e.g., paintings), then industry can use complex solutions. One of the limitations of **Article E** is that a generic solution for objective QA of 2.5D prints was not proposed. Nevertheless, contribution is that industry can use this research as a meaningful quality measure and knowledge on how metrics and acquisition set-ups can be done for objective QA of 2.5D prints was also added. In addition, other limitations of **Article E** can be related to, for example, camera calibration and accurate color acquisition because primary focus was on QA rather than on capture side. These limitations should be addressed in further works.

To conclude, it was shown that the tested metrics can be responsive to the differences between 2.5D prints for objective QA (**Article E**) and due to importance of subjective assessment at current stage of QA of 2.5D prints, there was focus more on the subjective side (**Articles A, B, C,** and **D**) than on the objective side. In the research conducted in the thesis, focus was set on specific content thereby generating workflows that can be used for other contents. It is understandable that data which can be generalized to diverse set of contents are preferable but in practice it is challenging to tackle generalization issue in every field. Furthermore, one can acquire results which have significance and usefulness in a particular area by focusing on specific content.

The contribution was an attempt to understand and find what is important for observers when they assess quality and naturalness of 2.5D prints. If one does not know what is important for observers when they judge quality, it will be difficult to make a metric that could work and optimize quality. The results of the research presented in the thesis can be seen as a step towards being able to do better reproduction and better QA.

Chapter 5

Conclusions and future perspectives

This chapter provides conclusions and future perspectives.

5.1 Conclusions

A visual experiment was conducted to investigate what are the relevant quality attributes for QA of 2.5D prints (**Article A**). There are many quality attributes and therefore it is desirable to reduce and optimize attribute space and find the most relevant ones. The relevant quality attributes can help to do QA consistently. It was found that color, sharpness, elevation, lightness, and naturalness were the top five most used distinct quality attributes.

Furthermore, naturalness aspect can affect QA of 2.5D prints. As a result, naturalness perception of 2.5D prints was investigated by considering effects of various parameters through a series of visual experiments (**Articles B, C,** and **D**). First, the effect of elevation and surface roughness on naturalness perception of 2.5D prints of various material images (i.e., wood, glass, stone, and metal) was considered (**Article B**). It was found that elevation had impact on naturalness perception of 2.5D prints. The elevation effect was found to be content dependent. The observers perceived 2.5D prints of wood and glass images as natural at lower elevation whereas 2.5D prints of stone and metal images did not reveal a clear trend.

Second, the effect of elevation on naturalness perception of 2.5D prints of wood images with various wooden content (i.e., wooden floor, roof, wall, and wicker) was considered (**Article C**). It was found that the optimal elevation that makes 2.5D prints of wood images to be perceived as natural was content dependent and two cases regarding the optimal elevation were observed. The first case was when wood images overall were considered and the second case was when content within wood images (e.g., wooden floor) was considered. The optimal elevation that makes 2.5D prints of wood images to be perceived as natural was found to be 0.5 mm in the first case and in a range between 0.3 mm and 0.5 mm in the second case.

Third, the effect of various ink types on naturalness perception of 2.5D prints of wood images with various wooden content (i.e., wooden floor, roof, wall, and wicker) was considered (**Article D**). It was found that the observers perceived 2.5D prints of wood images with matt inks as more natural than with glossy inks.

In addition, existing metrics, which are mostly designed to work with 2D images, were investigated whether they are capable of working with 2.5D prints to provide objective QA (**Article E**). It was found that the tested metrics can work with 2.5D prints under limited conditions.

Table 5.1 provides a brief information on answers to the research questions and on industrial value of the results.

In conclusion, acquired findings can create a basis for developing both subjective and objective guidelines for QA of 2.5D prints.

5.2 Future perspectives

Potential future perspectives on QA and naturalness perception of 2.5D prints can be as follows:

- Test different viewing distances
- Test various sizes of prints
- Use real substrate instead of paper or plastic based one
- Conduct experiment with tactility
- Conduct experiment (either visually or with tactility, or both) with real materials versus with reproductions
- Consider effect of various illuminations
- Consider various shapes (e.g., rectangular, circle, square)
- Recruit more observers
- Test more capture methods
- Test more existing metrics
- Consider virtual reality headsets for experiment with digital samples

To conclude, above-mentioned potential future works can help in generalizing findings and, consequently, in developing reliable industrial protocol(s) for both subjective and objective QA of 2.5D prints.

Table 5.1: The main outcomes of the research conducted in the thesis are presented with research questions and suggestions on industrial value.

Research questions	**Main results**	**Industrial value**
What are the relevant quality attributes for QA of 2.5D prints? (**Article A**)	The top five relevant quality attributes are color, sharpness, elevation, lightness, and naturalness	The results can help to focus on the relevant quality attributes for QA of 2.5D prints
What is the effect of elevation and surface roughness on naturalness perception of 2.5D prints? (**Article B**)	Elevation has impact and it is content dependent and 2.5D prints of wood and glass images at lower elevation are perceived as natural	Industry can consider to fabricate 2.5D prints of wood and glass images at lower elevation to make them perceived as natural looking for customers
What is the effect of elevation on naturalness perception of 2.5D prints? (**Article C**)	The optimal elevation for 2.5D prints within wood content images: 0.3 mm - 0.5 mm and overall for wood images: 0.5 mm	The optimal elevation can be used to reproduce 2.5D prints of wood images that will be perceived as natural for customers
What is the effect of various ink types on naturalness perception of 2.5D prints? (**Article D**)	2.5D prints of wood images with matt inks are perceived as more natural than with glossy inks	Industry can use specific type of ink to reproduce natural-looking wooden content
Can existing metrics assess quality of 2.5D prints? (**Article E**)	Tested metrics can work under limited conditions	Tested metrics can be used to detect differences between 2.5D prints during serial print production

Bibliography

[1] M. Page, G. Obein, M. V. O. Segovia, C. Boust, and A. Razet, “Adapted modulation transfer function method for characterization and improvement of 3D printed surfaces,” *Electronic Imaging*, vol. 2017, pp. 92–100, San Francisco, CA, USA, January, 2017.

[2] J. Wang, J. P. Allebach, and M. V. Ortiz-Segovia, “Semantically-based 2.5D texture printing,” in *2014 IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 2634–2638, Paris, France, October, 2014.

[3] C. T. Weijkamp, *Printing system and method of printing a multilayer structure using radiation curable ink*, US Patent 9,469,135, October, 2016.

[4] C. Parraman and M. V. O. Segovia, *2.5D Printing: Bridging the gap between 2D and 3D applications*. John Wiley & Sons, 2018.

[5] C. Weijkamp and C. Valade, *Method of elevated printing*, US Patent 9,834,006, December, 2017.

[6] T. Baar, S. Samadzadegan, H. Brettel, P. Urban, and M. V. O. Segovia, “Printing gloss effects in a 2.5D system,” in *Measuring, Modeling, and Reproducing Material Appearance*, SPIE, vol. 9018, pp. 160–167, San Francisco, CA, USA, February, 2014.

[7] C. T. Weijkamp and A. M. Gerrits, *Method of relief printing and a relief printer*, US Patent 10,471,730, November, 2019.

[8] D. A. Garvin, “What does "product quality" really mean?” *Sloan Management Review (pre-1986)*, vol. 26, no. 1, pp. 25–43, 1984.

[9] D. A. Silverstein and J. E. Farrell, “The relationship between image fidelity and image quality,” in *Proceedings of 3rd IEEE International Conference on Image Processing*, IEEE, vol. 1, pp. 881–884, Lausanne, Switzerland, September, 1996.

[10] International Organization for Standardization (ISO), ISO 20462-1, Photography — Psychophysical experimental methods for estimating image quality — Part 1: Overview of psychophysical elements, Geneva, Switzerland, 2005.

[11] B. W. Keelan, *Handbook of image quality: Characterization and prediction*. CRC Press, 2002.

[12] H. Kipphan, *Handbook of print media: Technologies and production methods*. Springer Science & Business Media, 2001.

[13] G. Marcu, “Factors governing print quality in color prints,” in *PICS: Image Processing, Image Quality, Image Capture, Systems Conference*, pp. 221–225, Portland, Oregon, USA, May, 1998.

[14] C. A. Reeves and D. A. Bednar, “Defining quality: Alternatives and implications,” *Academy of Management Review*, vol. 19, no. 3, pp. 419–445, 1994.

[15] J. F. Sawyer, “Effect of graininess and sharpness on perceived print quality,” in *Photographic Image Quality Symposium*, Royal Photographic Society, pp. 222–231, Oxford, UK, September, 1980.

[16] C. J. Bartleson, “The combined influence of sharpness and graininess on the quality of colour prints,” *The Journal of Photographic Science*, vol. 30, no. 2, pp. 33–38, 1982.

[17] A. Kadyrova, V. Kitanovski, and M. Pedersen, “A study on attributes for 2.5D print quality assessment,” in *Color and Imaging Conference*, Society for Imaging Science and Technology, vol. 2020, no. 28, pp. 19–24, Online, 4-19 November, 2020.

[18] M. Pedersen, N. Bonnier, J. Y. Hardeberg, and F. Albregtsen, “Attributes of image quality for color prints,” *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011 016, 2010.

[19] M. D. Fairchild, *Color appearance models*. John Wiley & Sons, 2005.

[20] B. M. Whelan, *Color Harmony, 2: A Guide to Creative Color Combinations*. Rockport publishers, 1994.

[21] P. Voigt and A. Von dem Bussche, “The EU General Data Protection Regulation (GDPR): A practical guide,” *1st Ed., Cham: Springer International Publishing*, vol. 10, no. 3152676, pp. 10–5555, 2017.

[22] World Medical Association, "Declaration of Helsinki: Ethical principles for medical research involving human subjects," *Journal of the American Medical Association (JAMA)*, vol. 310, no. 20, pp. 2191–2194, 2013.

[23] Massachusetts Executive Office of Health and Human Services, Department of Public Health, Conduct of Human Subject Research Protocol, Available online: https://www.mass.gov/service-details/policy-for-conduct-of-human-subject-research, (Accessed on 20 August 2022).

[24] International Organization for Standardization (ISO), ISO 3664, Graphic technology and photography — Viewing conditions, Geneva, Switzerland, 2009.

[25] American Society for Testing and Materials (ASTM), Standard guide for selection, evaluation, and training of observers, E 1499 – 97, Reprinted from the Annual Book of ASTM Standards, Copyright ASTM, 1997.

[26] International Telecommunication Union (ITU), ITU-T P. 800: Methods for subjective determination of transmission quality, 1996.

[27] International Telecommunication Union (ITU), Recommendation ITU-R BT. 500-13, Methodology for the subjective assessment of the quality of television pictures, Geneva, Switzerland, 2012.

[28] International Organization for Standardization (ISO), ISO 20462-2, Photography — Psychophysical experimental methods for estimating image quality — Part 2: Triplet comparison method, Geneva, Switzerland, 2005.

[29] International Organization for Standardization (ISO), ISO 20462-3, Photography — Psychophysical experimental methods for estimating image quality — Part 3: Quality ruler method, Geneva, Switzerland, 2012.

[30] R. C. Streijl, S. Winkler, and D. S. Hands, "Mean opinion score (MOS) revisited: Methods and applications, limitations and alternatives," *Multimedia Systems*, vol. 22, no. 2, pp. 213–227, 2016.

[31] M. Pedersen, "Image quality metrics for the evaluation of printing workflows," Ph.D. Thesis, University of Oslo, Oslo, Norway, 2011.

[32] P. Nussbaum, "Colour measurement and print quality assessment in a colour managed printing workflow," Ph.D. Thesis, University of Oslo, Oslo, Norway, 2011.

[33] T. Baar, S. Samadzadegan, P. Urban, and M. V. Ortiz Segovia, "Interrelation between gloss and texture perception of 2.5D-printed surfaces," *Electronic Imaging*, vol. 2016, no. 9, pp. 1–6, San Francisco, California, USA, February, 2016.

[34] K. Seshadrinathan and A. C. Bovik, "Encyclopedia of multimedia, Chapter: Advances in image and video quality assessment," *Springer*, pp. 8–17, 2008.

[35] M. Pedersen and J. Y. Hardeberg, "Full-reference image quality metrics: Classification and evaluation," *Foundations and Trends® in Computer Graphics and Vision*, vol. 7, no. 1, pp. 1–80, 2012.

[36] İ. Avcıbaş, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," *Journal of Electronic Imaging*, vol. 11, no. 2, pp. 206–223, 2002.

[37] A. Kadyrova, V. Kitanovski, and M. Pedersen, "Quality assessment of 2.5D prints using 2D image quality metrics," *Applied Sciences*, vol. 11, no. 16, p. 7470, 2021.

[38] T. Baar, H. Brettel, and M. V. O. Segovia, "A survey of 3D image quality metrics for relief print evaluation," in *2013 Colour and Visual Computing Symposium (CVCS)*, IEEE, pp. 1–6, Gjøvik, Norway, September, 2013.

[39] X. Liu, L. Chen, M.-V. Ortiz-Segovia, J. Ferwerda, and J. Allebach, "Characterization of relief printing," in *Measuring, Modeling, and Reproducing Material Appearance*, SPIE, vol. 9018, pp. 183–196, San Francisco, CA, USA, March, 2014.

[40] D. J. Whitehouse, *Handbook of surface metrology*. CRC Press, 1994.

[41] R. S. Hunter and R. W. Harold, *The measurement of appearance*. John Wiley & Sons, 1987.

[42] G. Sharma, *Digital Color Imaging Handbook*. CRC Press, 2003.

[43] International Organization for Standardization (ISO), ISO 2813, Paints and varnishes — Determination of gloss value at 20°, 60° and 85°, Geneva, Switzerland, 2014.

[44] International Organization for Standardization (ISO), ISO 8254-1, Paper and board — Measurement of specular gloss — Part 1: 75° gloss with a converging beam, TAPPI method, Geneva, Switzerland, 2009.

[45] International Organization for Standardization (ISO), ISO 8254-2, Paper and board — Measurement of specular gloss — Part 2: 75° gloss with a parallel beam, DIN method, Geneva, Switzerland, 2016.

[46] International Organization for Standardization (ISO), ISO 8254-3, Paper and board — Measurement of specular gloss — Part 3: 20° gloss with a converging beam, TAPPI method, Geneva, Switzerland, 2016.

[47] W. S. Elkhuizen, B. A. J. Lenseigne, T. Baar, W. Verhofstad, E. Tempelman, J. M. P. Geraedts, and J. Dik, "Reproducing oil paint gloss in print for the purpose of creating reproductions of old masters," in *Measuring, Modeling, and Reproducing Material Appearance 2015*, SPIE, vol. 9398, pp. 293–306, San Francisco, CA, USA, April, 2015.

[48] W. S. Elkhuizen, T. W. J. Callewaert, E. Leonhardt, A. Vandivere, Y. Song, S. C. Pont, J. M. P. Geraedts, and J. Dik, "Comparison of three 3D scanning techniques for paintings, as applied to Vermeer's 'Girl with a pearl earring'," *Heritage Science*, vol. 7, no. 1, pp. 1–22, 2019.

[49] W. S. Elkhuizen, "This is not a painting: Scanning and printing a painting's appearance," Ph.D. Thesis, Delft University of Technology, The Netherlands, 2019.

[50] K. E. Overvliet and S. Soto-Faraco, "I can't believe this isn't wood! An investigation in the perception of naturalness," *Acta Psychologica*, vol. 136, no. 1, pp. 95–111, 2011.

[51] A.-S. Binninger, "Perception of naturalness of food packaging and its role in consumer product evaluation," *Journal of Food Products Marketing*, vol. 23, no. 3, pp. 251–266, 2017.

[52] P. Rozin, "Naturalness judgments by lay Americans: Process dominates content in judgments of food or water acceptability and naturalness," *Judgment and Decision Making*, vol. 1, no. 2, pp. 91–97, 2006.

[53] P. Rozin, C. Fischler, and C. Shields-Argelès, "Additivity dominance: Additives are more potent and more often lexicalized across languages than are "subtractives"," *Judgment and Decision Making*, vol. 4, no. 5, pp. 475–478, August, 2009.

[54] E. A. Fedorovskaya, F. J. J. Blommaert, and H. d. Ridder, “Perceptual quality of color images of natural scenes transformed in CIELUV color space,” in *Color and Imaging Conference*, Society for Imaging Science and Technology, vol. 1993, pp. 37–40, Scottsdale, Arizona, USA, November, 1993.

[55] I. Heynderickx, “Performance evaluation of 2D and 3D-TV systems,” in *ICIS’06: 30th International Congress of Imaging Science*, pp. 374–377, Rochester, New York, USA, May, 2006.

[56] R. Halonen, S. Westman, and P. Oittinen, “Naturalness and interestingness of test images for visual quality evaluation,” in *Image Quality and System Performance VIII*, SPIE, vol. 7867, pp. 349–360, San Francisco, California, USA, January, 2011.

[57] A. Yoshida, V. Blanz, K. Myszkowski, and H.-P. Seidel, “Perceptual evaluation of tone mapping operators with real-world scenes,” in *Human Vision and Electronic Imaging X*, SPIE, vol. 5666, pp. 192–203, San Jose, CA, USA, 2005.

[58] H. Yeganeh and Z. Wang, “Structural fidelity vs. naturalness - objective assessment of tone mapped images,” in *International Conference on Image Analysis and Recognition*, Springer, pp. 111–121, Burnaby, BC, Canada, June, 2011.

[59] A. Kadyrova, M. Pedersen, and S. Westland, “Effect of elevation and surface roughness on naturalness perception of 2.5D decor prints,” *Materials*, vol. 15, no. 9, p. 3372, 2022.

[60] A. Etale and M. Siegrist, “Food processing and perceived naturalness: Is it more natural or just more traditional?” *Food Quality and Preference*, vol. 94, p. 104 323, 2021.

[61] M. D. Burnard, A. Q. Nyrud, K. Bysheim, A. Kutnar, K. Vahtikari, and M. Hughes, “Building material naturalness: Perceptions from Finland, Norway and Slovenia,” *Indoor and Built Environment*, vol. 26, no. 1, pp. 92–107, 2017.

[62] K. E. Overvliet, E. Karana, and S. Soto-Faraco, “Perception of naturalness in textiles,” *Materials & Design*, vol. 90, pp. 1192–1199, 2016.

[63] A. Etale and M. Siegrist, “Perceived naturalness of water: The effect of biological agents and beneficial human action,” *Food Quality and Preference*, vol. 68, pp. 245–249, 2018.

[64] A. Kadyrova, M. Pedersen, and S. Westland, “What elevation makes 2.5D prints perceptually natural?” *Materials*, vol. 15, no. 10, p. 3573, 2022.

[65] A. Kadyrova, M. Pedersen, S. Westland, and C. Weijkamp, “Effect of various ink types on naturalness perception of 2.5D prints,” Under review in a journal, 18 pages.

[66] Pixabay, Available online: https://pixabay.com, (Accessed on 17 December 2019).

[67] F. Drago, W. L. Martens, K. Myszkowski, and H.-P. Seidel, "Perceptual evaluation of tone mapping operators," in *ACM SIGGRAPH 2003 Sketches & Applications*, pp. 1–1, San Diego, CA, USA, July, 2003.

[68] 3D textures, Available online: https://3dtextures.me, (Accessed on 6 July 2021).

[69] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[70] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, IEEE, vol. 2, pp. 1398–1402, Pacific Grove, CA, USA, November, 2003.

[71] J. Preiss, F. Fernandes, and P. Urban, "Color-image quality assessment: From prediction to optimization," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1366–1378, 2014.

[72] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Research & Application*, vol. 30, no. 1, pp. 21–30, 2005.

[73] X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," in *SID International Symposium Digest of Technical Papers*, Citeseer, vol. 27, pp. 731–734, Princeton, NJ, USA, 1996.

[74] A. Kadyrova, V. Kitanovski, and M. Pedersen, *The Colourlab Image Dataset for 2.5D Quality Assessment*, Zenodo, 2021. DOI: 10.5281/zenodo.4609031.

[75] T. Leisti and J. Häkkinen, "The effect of introspection on judgment and decision making is dependent on the quality of conscious thinking," *Consciousness and Cognition*, vol. 42, pp. 340–351, 2016.

[76] Commission Internationale de l'Éclairage (CIE), Guidelines for the evaluation of gamut mapping algorithms, Austria, 2003.

[77] V. Kitanovski and J. Y. Hardeberg, "Objective evaluation of relighting models on translucent materials from multispectral RTI images," *Electronic Imaging*, vol. 2021, no. 5, pp. 133–1, Online, January, 2021.

[78] H. Mytum and J. R. Peterson, "The application of reflectance transformation imaging (RTI) in historical archaeology," *Historical Archaeology*, vol. 52, no. 2, pp. 489–503, 2018.

[79] CRUSE, Available online: https://crusescanner.com, (Accessed on 3 July 2022).

[80] Factum Arte, Available online: https://www.factum-arte.com/, (Accessed on 3 July 2022).

[81] Arius Technology, Available online: https://www.ariustechnology.com, (Accessed on 4 July 2022).

Part II: Included articles

Article A

A study on attributes for 2.5D print quality assessment

Altynay Kadyrova, Vlado Kitanovski, Marius Pedersen,
Color and Imaging Conference,
Society for Imaging Science and Technology,
Vol. 2020, No. 28, pp. 19-24,
November, 2020.

This article is not included due to copyright restrictions.

Article B

Effect of elevation and surface roughness on naturalness perception of 2.5D decor prints

Altynay Kadyrova, Marius Pedersen, Stephen Westland,
Materials,
Vol. 15, No. 9, p. 3372,
May, 2022.

materials

Article

Effect of Elevation and Surface Roughness on Naturalness Perception of 2.5D Decor Prints

Altynay Kadyrova [1,*], Marius Pedersen [1] and Stephen Westland [2]

[1] Department of Computer Science, Norwegian University of Science and Technology, 2802 Gjøvik, Norway; marius.pedersen@ntnu.no
[2] School of Design, University of Leeds, Leeds LS2 9JT, UK; s.westland@leeds.ac.uk
* Correspondence: altynay.kadyrova@ntnu.no

Abstract: Naturalness is a complex concept. It can involve a variety of attributes. In this work, we considered the effect of elevation and surface roughness on naturalness perception of 2.5D decor prints for four material categories. We found that elevation has an impact on the naturalness perception of 2.5D decor prints and that it is linked with content. The observers found lower elevation to be more natural for wood and glass 2.5D prints while there was no clear tendency for stone and metal 2.5D prints. We also found the perceptual attributes used for naturalness assessment of 2.5D decor prints. The top five ones are color, roughness, gloss, elevation, and lightness. The obtained findings can be useful for companies that produce 2.5D prints.

Keywords: decor; 2.5D printing; naturalness

check for updates

Citation: Kadyrova, A.; Pedersen, M.; Westland, S. Effect of Elevation and Surface Roughness on Naturalness Perception of 2.5D Decor Prints. *Materials* **2022**, *15*, 3372. https://doi.org/10.3390/ma15093372

Academic Editor: Pawel Pawlus

Received: 30 March 2022
Accepted: 4 May 2022
Published: 8 May 2022

1. Introduction

Decor is one of the active interest areas in 2.5D printing based on industry feedback. Therefore, the production of 2.5D decor prints that look natural is demanded. A variety of aspects might affect the naturalness perception of 2.5D decor prints: the presence or perception of various quality attributes, illumination, and viewers' perspectives on the quality depending on their experience and preferences, to name a few. If decor prints look natural to the viewers, then they will be considered as high quality, and consequently, will be the most demanded by the customers. As a result, it is important to investigate what parameters impact the naturalness perception of 2.5D decor prints and to what degree. In this work, we consider the effect of various quality attributes on the naturalness perception of 2.5D decor prints at a given illumination and viewing distance. To date, no study has looked specifically at 2.5D decor prints' naturalness perception.

Elevation and naturalness were found to be in the top five most used distinct attributes during quality assessment of 2.5D prints [1]. Moreover, they are relevant from an industrial point of view as industry is investigating how elevated prints (i.e., 2.5D prints) look natural. The main feature of 2.5D prints is elevation, and it should look natural to be of high quality perceptually. The surface roughness might help to provide a realistic appearance for the prints, and it is content and material dependent. Hence, our goal is to investigate how the elevation and the surface roughness affect the naturalness perception of 2.5D decor prints. The relevance of this work is that it can provide insights on how people define the naturalness of 2.5D decor prints. Furthermore, it can be a source (or a motivation) for developing (industrial) protocols or guidelines for creating 2.5D decor prints with a natural look and finding out what level of elevation (e.g., 0.4 mm or 0.6 mm) makes a perceptually natural appearance for 2.5D decor prints. We limit to two (elevation and surface roughness) quality attributes because looking at three or more quality attributes will make the experiment long, which in turn might affect observers' performance (i.e., leads to observer fatigue). For simplicity, by prints we mean 2.5D decor prints, by roughness we mean surface roughness, and by wood/glass/stone/metal prints we mean 2.5D prints of wood/glass/stone/metal images hereafter in the text unless specified otherwise.

This paper is organized as follows: first, we give background information about the naturalness concept (e.g., in 2D and 3D images, 2D and 2.5D prints) followed by our methodology description; afterwards, the results and discussion are given; last, we provide our conclusions and future works.

2. Background

We give brief background information on naturalness in images (2D, 3D) and prints (2D, 2.5D) to show that naturalness is a complex concept and combines various quality attributes.

Naturalness can be defined as a close matching between an image's visual presentation and the understanding of the reality that is in memory [2], and it usually arises during overall image quality assessments [3]. During quality assessments, observers might use words such as natural, real, unnatural, and unreal (2D images [4], 2D prints [5,6]) and most of the time they are used to express image naturalness [3]. These attributes along with words such as edited, photoshopped, aged photo, and others were grouped into the naturalness category by Virtanen et al. [7] in their proposed image quality wheel.

Naturalness was stated as a preferential quality attribute of high level [3]. Generally, studying how, for example, chroma or sharpness variations impact the image naturalness perception is the typical approach of exploring naturalness [3]. Dependence of naturalness on a combination of various attributes was mentioned by Yoshida et al. [8] for tone-mapped 2D images. Yeganeh et al. [9] defined the naturalness of 2D images with two attributes' (brightness and contrast) joint probability density function. Pedersen et al. [10] mentioned that naturalness can be related to, for example, color and lightness changes for 2D color prints.

Halonen et al. [3] stated that naturalness and interestingness need to be balanced when creating test images for visual quality assessments, and they worked with 2D prints. Fedorovskaya et al. [11] found that naturalness and perceptual quality have a close relationship in the context of 2D images. More specifically, they found that an image becomes unnatural due to an increase in colorfulness, which decreases the image quality. Naturalness along with details were found to be the most important/salient perceptual attributes that describe perceptual differences of 2D images [12]. There are also works dedicated to model naturalness of 2D images [13–16]. For instance, Choi et al. [13] used the sharpness and colorfulness of images, shadow-detail reproduction, and lack of washed-out appearance factors along with memory colors for 2D image naturalness modeling. The image sharpness was represented by averaged pixel-based color difference because the authors assumed that neighboring pixels' color difference might become larger when the sharpness is increased. They used lightness to represent the shadow detail and the washed-out appearance reproduction and chroma to represent the colorfulness. They worked with both CIECAM02 and CAM02-UCS spaces.

According to Seuntiëns [17], people tolerate image distortions when rating the naturalness of both 2D and 3D images. Additionally, naturalness was found to be among the top five most used distinct attributes during quality assessment of 2.5D prints [1].

To conclude, the complexity of naturalness increases with the increase in image/print dimensionality. To our knowledge, there is no study on naturalness of higher dimensional physical prints (i.e., 2.5D).

3. Methodology

Naturalness can be multidimensional [18] and can have various meanings depending on content. Thus, we focus on one type of content, which is decor prints. Our workflow is illustrated in Figure 1.

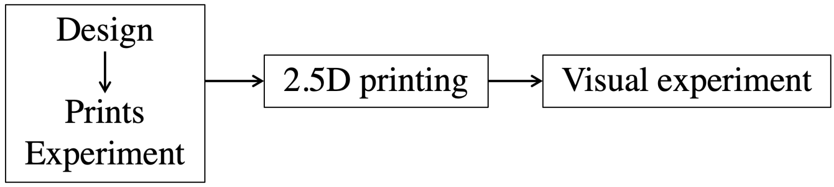

Figure 1. Our workflow. We started with designing the experiment and prints followed by 2.5D printing stage. Afterwards, we conducted the visual experiment.

From the literature in the previous section, we can see that naturalness is a complex concept. In our context, realistic can mean naturalness. Naturalness was grouped together with the word real as a synonym for 2.5D prints [1]. Virtanen et al. [7] classified the word real into the naturalness category based on their data consisting of sixty-two scenes presented to the observers either via 2D print images or images on display. Thus, we define naturalness by substituting it with the term realistic representation of a print. We follow the definition of Drago et al. [12] where naturalness is considered as the extent to which an image is similar to a realistic scene. We do not refer to material properties with naturalness in this work.

3.1. Images

Sharan et al. [19] defined ten material categories. We worked with four material categories: wood, stone, metal, and glass. For each material category, we had a variety of content. As an example, for the wood category, we had images of wooden decor, wooden walls, and more. These four categories were selected because we believe that they represent the most used decor materials.

For each material category, we had five color images, resulting in a total of 20 images. The images and their height maps (both are in 782×782 pixels) were reproduced from 3D textures (copyright free web site) [20] under the Creative Commons license. They contain various levels of spatial information and colorfulness [21]. The original height maps underwent the processing illustrated in Figure 2 in order to avoid printing artifacts, such as black edges due to high elevation, and to obtain visually nice prints. To reduce black edges, the height maps were processed with a Gaussian filter with a standard deviation of four, and to ensure visually nice prints, a morphological operation was applied to some images. An intensity adjustment was done to reach the intended maximum elevation. The roughness was added by direct binary search halftone blue noise with a zero mean generated by software [22] (input image was a flat grayscale at 128 with zero-mean uniform noise added, and the output image was a halftone noise image). According to Kitanovski and Pedersen [22], the direct binary search algorithm provides high-quality prints. The halftone noise image was further resized with nearest-neighbor interpolation with a resize factor of two and then cropped to the intended size. This was done to get low-frequency noise. We did not use high-frequency noise because the roughness was not visible with it during our initial tests. We applied a gamma function so that the roughness would be reproducible. We used a gamma value of 1/1.4. It was chosen to get visually nice prints via test printing of various gamma values.

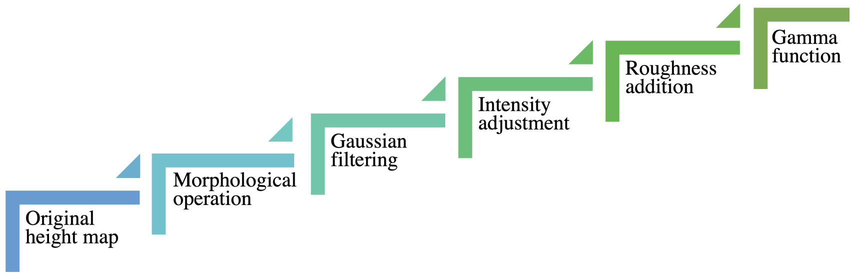

Figure 2. Image processing steps used for design of prints. The original height map went through several processing in the following order: morphological operation (optional, depends on image content), Gaussian filtering, intensity adjustment, roughness addition, and gamma function application.

We used an outdoor paper substrate. An OCE Arizona 2280GT 2.5D printer was used for the fabrication of prints. We used Alto printer mode, meaning that the elevation was opaque. The print size was 6.62 × 6.62 cm. We also added an additional 0.3 cm on each side of the substrate paper so that observers could hold the prints without touching the actual edges.

3.2. Elevation and Surface Roughness Levels

The selected maximum elevation levels and roughness constants (further referred to as R_c) to make the roughness levels and the approximate maximum roughness amounts (further referred to as R_a) are presented in Figure 3. We found that prints with very low elevations look perceptually towards flat through test printing at various elevations. Moreover, it is important to consider that 2.5D prints are elevated prints. As a result, we chose the maximum elevation levels to be 0.4, 0.6, and 0.8 mm.

We used three R_c to acquire three levels of roughness. They were multiplied with the noise image to get a height map with the roughness. These R_c were chosen based on observations from test printing with various R_c at various maximum elevations. With an $R_c < 6$, the roughness looks less visible to the naked eye, especially at lower elevations. If $R_c > 10$, the roughness does not look visually nice, especially at higher elevations. Based on these, three values of R_c between 6 and 10 with a step of 2 were chosen.

The R_a was calculated based on K-values (can be seen with a color picker in Adobe Photoshop) from the processed height maps. For example, if the K-values on two neighboring pixels are 100% and 89% and the maximum elevation is set to 0.4 mm, then the R_a in that part is approximately $(1 - 0.89) \times 0.4$ mm $= 0.044$ mm or 44 µm. Depending on content, the processed height maps have many pixels or few pixels with the maximum R_a. In our work, the roughness is the height difference within a local neighborhood. There were nine reproductions per image considering the three levels of elevation and roughness. This resulted in 4 categories × 5 images × 9 levels, which made a total of 180 2.5D prints for the experiment.

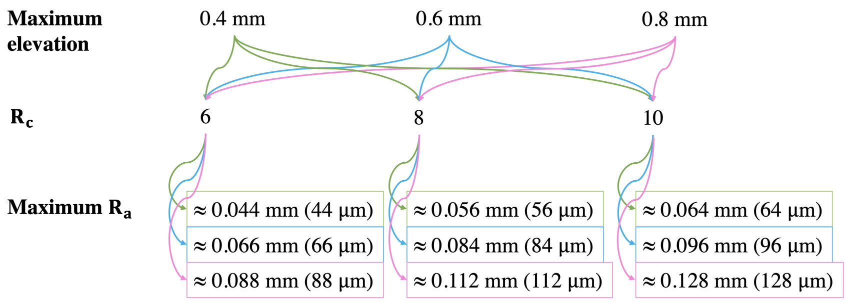

Figure 3. Values of maximum elevation and roughness constants (denoted as R_c) and calculated maximum roughness amount (denoted as R_a). The maximum R_a is an approximation of what we would have physically in the prints. Three elevation levels at three roughness constants (R_c) gave nine reproductions.

3.3. Visual Experiment

We did a ranking experiment because it is fast and easy for observers. Given the number of reproductions, a technique such as pair comparison would be very time-consuming. Our experimental design is illustrated in Figure 4. The observers provided their consent for participation in the experiment and for audio recording, and had a 2–3 min adaptation period to the illumination prior to starting the experiment. The 2.5D prints were presented to the observers in random order inside a light booth cabinet (Verivide CAC 60-5, illumination was 1400 lux) with D65 illumination. The prints were placed onto a 3D-printed 45° holder. As recommended by ITU [23], we did a training session so that observers could better understand the experiment's objective and task. We used one 2.5D print (not from the total 180 prints) for a training session. After the training session, the observers also had the opportunity to ask questions before continuing. The instruction given to the observers was to rank the prints from the most to the least realistic representation of wood/stone/metal/glass decor and explain why. We did not give any physical reference to avoid observers doing fidelity matching. Instead, we provided the material category name for each print. Hence, we gave keywords as a reference. It is easier for observers to judge the realistic representation of prints when they know the material category. The distance between the prints and observers' eyes was around 50 cm. The observers were allowed, with provided gloves, to take the prints from the holder and rotate and move them. However, as we did not consider tactility, they were not allowed to touch the prints' surfaces. They were informed that there was no time restriction. The average duration of the experiment was 1 h and 16 min per observer. All observers finished the experiment in one session except one, who did it in two sessions.

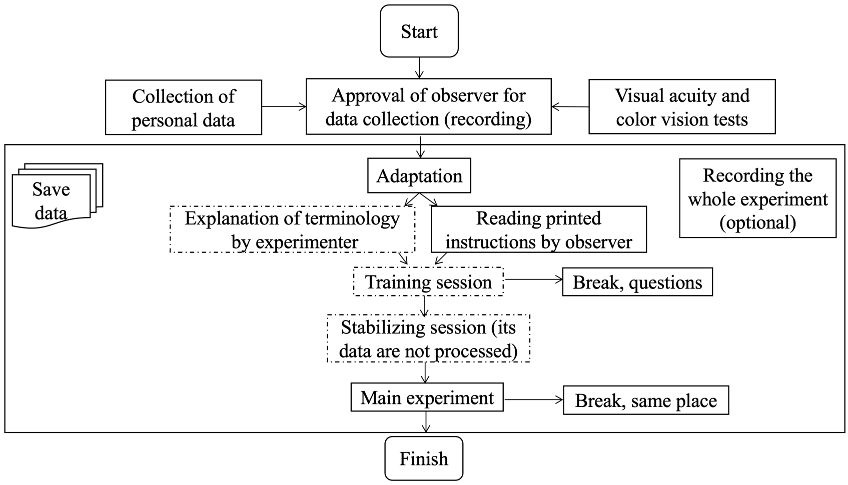

Figure 4. Our proposed framework for subjective quality assessment of 2.5D prints. Parts with dashed dots are optional. We included optional training and stabilizing sessions following the ITU [23] recommendation.

Twenty observers (8 females and 12 males; average age: around 36 years, standard deviation: around 12 years) with normal color vision participated in the experiment, except one observer who was color deficient. Ishihara plates and a Snellen chart were used to test color vision and visual acuity, respectively. The observers were mostly Europeans, and both naive and experienced (having a background in computer science or color imaging) people were involved. It is helpful to have both naive and experienced observers to find out how they define naturalness of prints. They might assess the prints differently [24]. The experiment was run in English. The language might have impacted observers' descriptions that they used to describe how they ranked the prints.

4. Results and Discussion

We analyzed the collected data both qualitatively and quantitatively to determine how people define the naturalness of 2.5D prints and what levels of elevation and surface roughness make 2.5D prints perceptually natural. We also present the limitations of our work in this section.

4.1. How People Define the Naturalness of 2.5D Prints?

To explore how people define the naturalness of 2.5D prints, we first present an analysis of the qualitative data. It provides perceptual attributes used by the observers as a strategy to decide on the naturalness of the 2.5D prints. Moreover, it also provides the most used perceptual attributes for examined material categories. In addition, we studied how elevation and roughness variations can affect the perception of other attributes with regard to the naturalness of 2.5D prints.

4.1.1. What Are the Perceptual Attributes Linked to the Naturalness of 2.5D Prints?

The steps of qualitative data processing were: first, audio data of observers were transcribed; second, the attributes used by the observers during the experiment were extracted; third, the extracted attributes were combined into groups. We followed Virtanen et al.'s [7] approach to grouping some of the sub-attributes and in terms of visual presentation of the attributes.

Figure 5 shows the perceptual attributes used for the naturalness assessment of 2.5D prints. We defined attribute groups at three levels. In total, we found twelve level 1 attribute groups (inner circle in Figure 5). They were color, roughness, gloss, elevation, lightness, sharpness, contrast, transparency, shape, softness, artifacts, and others. For example, level 2 attribute groups of sharpness were details and sharpness. Further, level 3 attribute groups of details were visibility and details. The color-related group included chromatic color, uniform color, artificial color, and similar expressions. The texture-related group included descriptions such as even texture, visible texture, rough texture, and similar expressions whereas the roughness-related group included brown spots, noisy rough, granular, and similar expressions. The reflection group included specular reflections, diffuse specular highlights, scattering effect, and similar. The shiny group included shiny, sparkling, glittery, and similar expressions. The elevation-related group included height, altitude, 2.5D, and similar expressions. The lightness-related group included lightness and dynamic range. The others group included descriptions such as substrate, weight, clean, variability, rusty, old, and intuition (some observers ranked based on their intuitions and were not able to explain why they ranked in a specific way). The shape group included shape, size, width, and geometry. Noise and graininess were combined into the artifacts group whereas softness and hardness were combined into the softness group.

From Figure 5, we can see that the top five most used perceptual attributes for naturalness assessment of 2.5D prints by the observers were color, roughness, gloss, elevation, and lightness. They all were among the top seven most used distinct attributes during quality assessment of 2.5D prints [1].

4.1.2. What Are the Most Used Perceptual Attributes Linked to the Naturalness of 2.5D Prints for Examined Material Categories?

The most used level 1 perceptual attributes were identified by frequency analysis for four material categories (Figure 6). Color, roughness, and gloss were the most used perceptual attributes for four material categories. Most of the observers preferred all wood prints to be less rough, more brown, and less glossy; all glass and all metal prints to be smoother and glossier; and all stone prints to be grayer, both rougher and smoother but more towards rougher, and less glossy to be more realistic based on their explanations provided for their rankings. We can see that the transparency attribute was used only for glass prints as expected and just one time for a stone print. The observer’s criterion for that stone print was translucency in the sense that a more stone-like print should have more translucency. Artifacts were not used for stone and glass prints. Additionally, softness and attributes grouped as others were not used for glass prints.

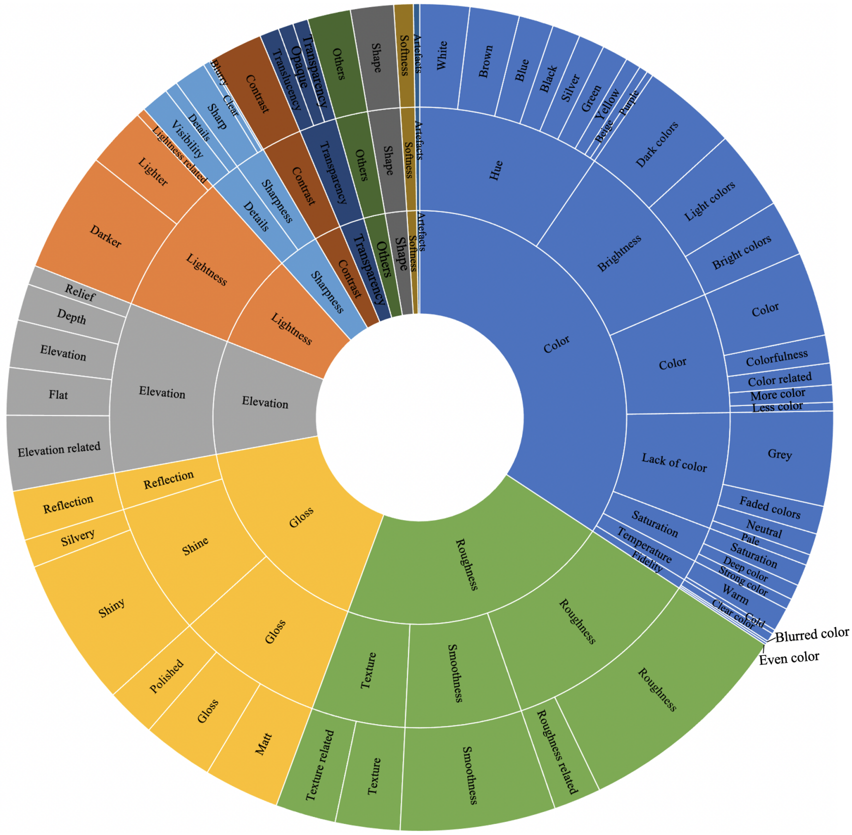

Figure 5. Perceptual attributes used for naturalness assessment of 2.5D prints. The most used attributes have larger areas.

Figure 6. The most used level 1 perceptual attributes for wood, glass, metal, and stone material categories. The size of the attribute's text is the frequency of its usage.

4.1.3. How Variations in Elevation and Surface Roughness Can Be Linked with the Used Perceptual Attributes for the Naturalness of 2.5D Prints?

From audio data of our observers (i.e., the explanations they provided after ranking), we can observe that various levels of elevation and roughness can impact the perception of other attributes' presence and their variations that affect the naturalness aspect. For instance, a combination of various levels of elevation and roughness can change the color appearance and glossiness aspect. Additionally, content can impact on other attributes' variation perception as well with regard to naturalness. We assume that higher elevation can make the print surface rougher, and similarly, lower elevation can make the print surface smoother. Moreover, we assume that more roughness can influence prints' surfaces to appear lighter due to inter-reflections, and higher elevation can cause more contrast in prints' surfaces. As a result, one can experience that, for example, the color of the prints varied even when color was not altered at all. Hence, we can make perception of variations of various attributes by changing just one or two attributes which in turn can impact on the naturalness perception. For 2.5D printing, it could be useful to investigate this observation in further work as it could help to produce eye-catching 2.5D products just by varying, for example, elevation levels.

Furthermore, the observers mentioned a set of factors that impact the naturalness assessment of 2.5D prints. They were grouped and named as others in Figure 5. In particular, it is worth mentioning the weight aspect that three observers mentioned. Two observers were consistent that the stone prints with more elevation should be heavier in weight, while one observer considered weight of the stone print but found the reproductions to be soft. We measured the weights of all 180 prints, and we found that the prints with more elevation had more weight than the prints with lower elevation. This is expected because there are more layers of ink in prints with more elevation and some of the observers were able to feel that.

To conclude, our finding of twelve level 1 perceptual attribute groups with which to judge the naturalness of 2.5D prints could be useful for modeling the naturalness of 2.5D prints objectively. Choi et al. [13] found, through their naturalness model, the attributes that most impact the naturalness perception of 2D images which were image sharpness and colorfulness. Thus, an objective metric for naturalness assessment of 2.5D prints can be a combination of existing models on 2D images and new models that consider the attributes found (Figure 5) in our work.

4.2. What Elevation and Surface Roughness Levels Make 2.5D Prints Perceptually Natural for Examined Material Categories?

In the previous subsection we found that the naturalness of 2.5D prints is linked with both elevation and roughness (Figure 5) along with other attributes, and as we changed the elevation and roughness in our prints, it is interesting to find what levels of elevation and roughness make 2.5D prints perceptually most natural. For this, we analyzed the collected data quantitatively. The raw data from the ranking experiment were converted into Z-scores. We analyzed the Z-scores image by image because, if we were to look at the combined Z-scores for all images, some effects might cancel out. For example, if one preferred a stone print to be rougher whereas another preferred a wood print to be smoother, then they would cancel out when the Z-scores for all images are combined.

When considering all images in each material category, we observed inverse proportionality between elevation and naturalness for all wood prints according to Z-scores (Figure 7). In other words, the observers found that wood prints should be less elevated to look natural. The same can be said of glass prints (Figure 7). No clear tendency for all stone and all metal prints (Figure 8) was found. We visualize Z-scores in error bar plots. Mean Z-scores are given by circles at the centers of the vertical lines. Confidence Interval (*CI*) was calculated as shown in Equation (1) [25].

$$CI = 1.96 \cdot \frac{\sigma}{\sqrt{N}}, \quad (1)$$

where N is the number of observations, and σ is the standard deviation which in the case of Z-score can be computed as $1/\sqrt{2}$ [26]. 95% *CI* is the mean Z-score $\pm$ *CI*. There is a statistically significant difference between the reproductions with 95% confidence, if two *CI*s do not overlap.

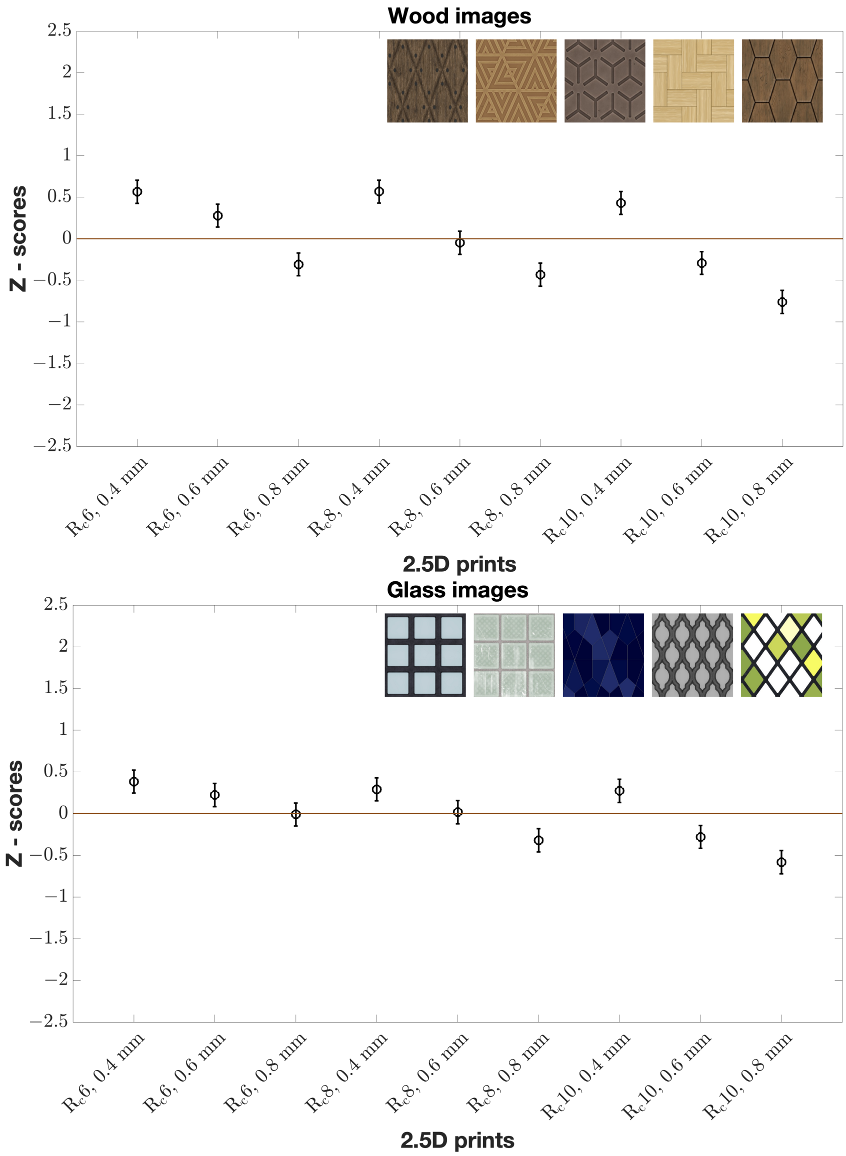

Figure 7. Z-scores of all images of wood and glass material categories by all observers. Mean Z-score values for nine reproductions (x-axis) are given with 95% *CI*s (represented by error bars). Z-scores have a small range. Each material category has five images.

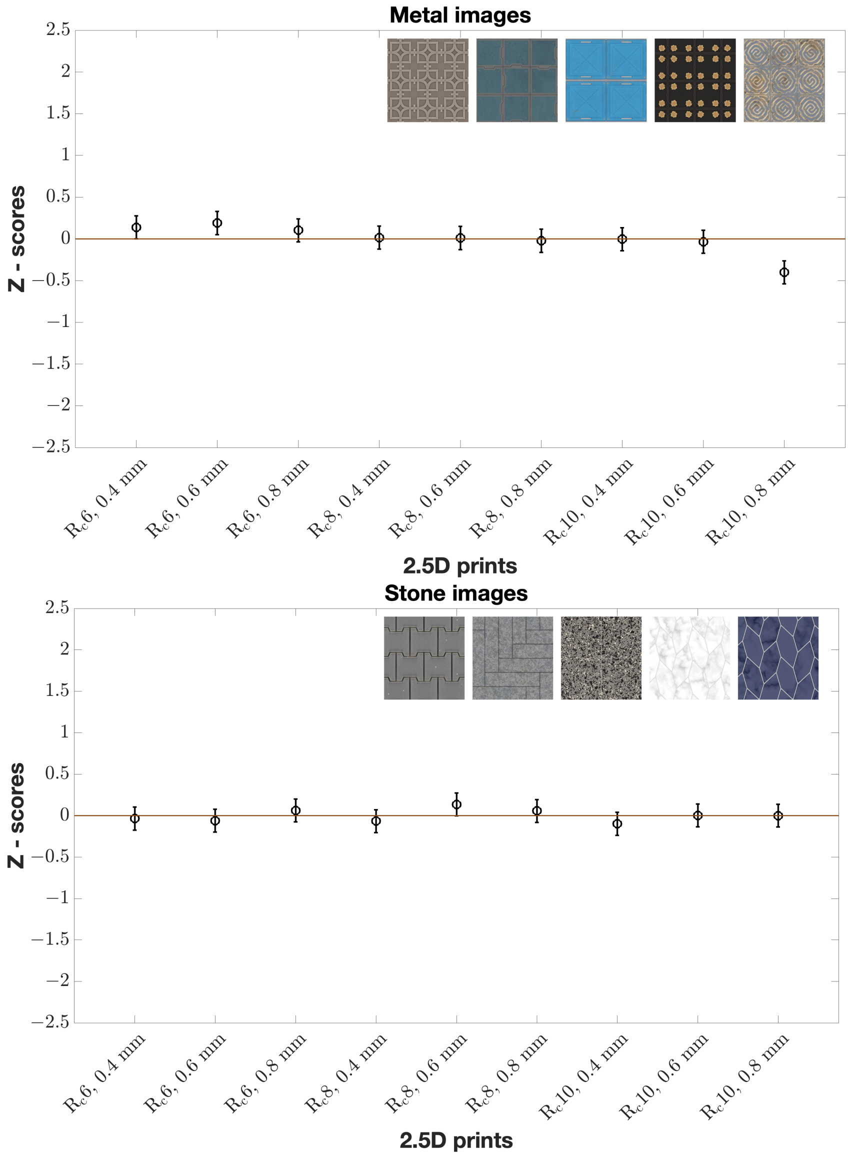

Figure 8. Z-scores of all images of metal and stone material categories by all observers. Mean Z-score values for nine reproductions (x-axis) are given with 95% *CI*s (represented by error bars). Z-scores have a small range. Each material category has five images.

We further analyzed the correlation of elevation with Z-scores and the correlation of roughness with Z-scores for all images in each material category. This showed that elevation had a correlation with Z-scores for all wood (Figure 9) and all glass images unlike all stone and all metal images. There was no clear correlation pattern of roughness with respect to the Z-scores for four material category images. It is worth mentioning that we did not find significant differences in Z-scores between naive and experienced observers and between genders.

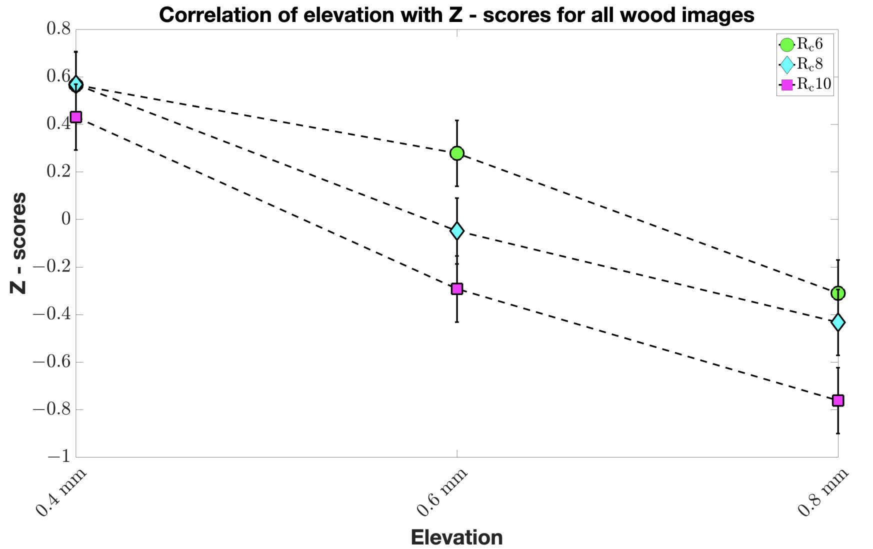

Figure 9. Correlation of elevation with Z-scores for all wood images. The x-axis represents 2.5D prints at various elevation levels. We can observe that the observers found lower elevation natural for all wood images, regardless of the roughness levels, as the *CI*s overlap.

Additionally, we used a binomial sign test on the raw data with Bonferroni correction (with a significance level of α/n, where $\alpha = 0.05$ is the desired alpha value and n is the number of comparisons: 0.05/36) [27]. Table 1 presents p-values obtained from the sign test for all wood images. We can observe that, at any roughness level, 0.4 mm had statistically significant difference in comparison with the other two elevation levels, and 0.6 mm had statistically significant difference in comparison with 0.8 mm. Considering both p-values (Table 1) and Z-scores (Figure 7), we can assume that the observers found 0.4 mm to be more natural than other two elevation levels regardless of the roughness levels for all wood prints. For all glass images (Table 2), lower roughness level resulted in a statistically significant difference compared to higher roughness levels at 0.8 mm. In other words, the observers found it more natural when all glass prints had less roughness at 0.8 mm. None of the reproductions resulted in a statistically significant difference for all stone images. The majority of reproductions resulted in no statistically significant difference for all metal images either. In addition, we looked into inter-observer variability using the Spearman correlation coefficient and found that, on average for all images, the correlation varied between observers. This shows the complexity of assessing the naturalness of 2.5D prints and the variability of the perception of overall print appearance from person to person.

Table 1. *p*-Values obtained by sign test for all wood images. Green cells are those that have a statistically significant difference while red ones are those that have no statistically significant difference. Threshold used in Bonferroni correction is 0.05/36 = 0.0014.

	R6, 0.4 mm	R6, 0.6 mm	R6, 0.8 mm	R8, 0.4 mm	R8, 0.6 mm	R8, 0.8 mm	R10, 0.4 mm	R10, 0.6 mm	R10, 0.8 mm
R6, 0.4 mm	-	9.6685×10^{-4}	3.7979×10^{-8}	0.4839	6.7953×10^{-6}	3.6350×10^{-9}	0.2713	3.7979×10^{-8}	3.6350×10^{-9}
R6, 0.6 mm		-	1.7080×10^{-5}	0.0069	0.0019	9.5837×10^{-7}	0.3681	2.6016×10^{-6}	3.6350×10^{-9}
R6, 0.8 mm			-	3.7979×10^{-8}	0.0124	0.1936	3.3965×10^{-7}	0.4839	6.7953×10^{-6}
R8, 0.4 mm				-	1.1580×10^{-7}	3.6350×10^{-9}	0.1936	3.6350×10^{-9}	2.9765×10^{-10}
R8, 0.6 mm					-	2.1560×10^{-4}	1.7080×10^{-5}	0.0124	1.1981×10^{-8}
R8, 0.8 mm						-	1.1981×10^{-8}	0.2713	9.6193×10^{-5}
R10, 0.4 mm							-	9.5837×10^{-7}	1.0607×10^{-9}
R10, 0.6 mm								-	9.5837×10^{-7}
R10, 0.8 mm									-

Table 2. *p*-Values obtained by sign test for all glass images. Green cells are those that have a statistically significant difference while red ones are those that have no statistically significant difference. Threshold used in Bonferroni correction is 0.05/36 = 0.0014.

	R6, 0.4 mm	R6, 0.6 mm	R6, 0.8 mm	R8, 0.4 mm	R8, 0.6 mm	R8, 0.8 mm	R10, 0.4 mm	R10, 0.6 mm	R10, 0.8 mm
R6, 0.4 mm	-	0.1936	0.0019	0.4839	2.1560×10^{-4}	9.6193×10^{-5}	0.1336	1.7080×10^{-5}	1.7080×10^{-5}
R6, 0.6 mm		-	0.0124	0.3681	0.1336	9.6193×10^{-5}	0.7642	4.6526×10^{-4}	6.7953×10^{-6}
R6, 0.8 mm			-	0.0214	0.6171	4.6526×10^{-4}	0.0019	0.2713	3.6350×10^{-9}
R8, 0.4 mm				-	0.0124	2.1560×10^{-4}	0.3681	2.1560×10^{-4}	4.1315×10^{-5}
R8, 0.6 mm					-	0.0019	0.0124	0.0019	6.7953×10^{-6}
R8, 0.8 mm						-	9.6193×10^{-5}	0.6171	4.6526×10^{-4}
R10, 0.4 mm							-	6.7953×10^{-6}	1.7080×10^{-5}
R10, 0.6 mm								-	0.0019
R10, 0.8 mm									-

To conclude, the observers preferred wood and glass 2.5D prints to have lower elevation to look perceptually natural. Furthermore, a lower elevation can make a print look smoother. In other words, we assume that the observers preferred wood and glass 2.5D prints to be less elevated and smoother.

4.3. Limitations

We focused on one type of content (decor) and worked with three variations of the selected quality attributes (i.e., elevation and roughness). If we were to increase the number of variations per attribute, then the experiment would have become long which would have affected observers' performance. We sampled sparsely (i.e., 3 × 3 grid) to find an area of interest that could be investigated further as a future work. We chose one content to narrow down our scope; otherwise, it would have become difficult to differentiate the results for different contents. By focusing on one content, we generated a workflow that can be followed to study the naturalness perception of 2.5D prints in other contents. It is important to mention that the results can vary depending on the content selected. In addition, our work is useful in the selected application area—decor—which is the most active area in 2.5D printing presently.

5. Conclusions and Future Works

According to the literature, there have been studies where the naturalness was involved in 2D images [2,4,8,9,11–16,28], 3D images [17,29], and 2D prints [3,5,10]. Naturalness as an attribute was mentioned in Kadyrova et al.'s [1] work on attributes for the quality assessment of 2.5D prints. To our knowledge, this work is the first which studied the naturalness perception of physical 2.5D prints. Thus, our work is unique. We investigated the effect of elevation and surface roughness on the naturalness perception of 2.5D prints. We found that the observers define the naturalness of 2.5D prints with various attributes (Figure 5, Section 4.1.1). The top five attributes that the observers prefer to look at when assessing the naturalness of 2.5D prints are color, roughness, gloss, elevation, and lightness. Moreover, we found that color, roughness, and gloss are the most used attributes for four examined material categories (Section 4.1.2). Based on the results, lower the elevation, more natural the wood and glass 2.5D prints to observers (Section 4.2). We also found that the naturalness of 2.5D prints is content dependent. Thus, it is important to consider what type of content one needs to reproduce to decide on what elevation level needs to be used. Additionally, we found that a change in one or more attributes can make perception of other attributes' variation with regard to the naturalness of 2.5D prints (Section 4.1.3).

Future work will be to explore what exact lower elevation makes 2.5D prints look perceptually natural, particularly wood prints. Additionally, it would be interesting to repeat the experiment with tactility.

Author Contributions: Conceptualization, A.K., M.P., and S.W.; methodology, A.K., M.P. and S.W.; data acquisition and processing, A.K.; analysis, A.K.; writing—original draft, A.K.; writing—review and editing, A.K., M.P. and S.W.; visualization, A.K.; supervision, M.P. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the ApPEARS project from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska–Curie grant, agreement number 814158.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original images and their height maps used in this work can be found at https://3dtextures.me, accessed on 6 July 2021.

Acknowledgments: We would like to thank Vlado Kitanovski and Clemens Weijkamp for fruitful discussions and observers for participation in the experiment.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Kadyrova, A.; Kitanovski, V.; Pedersen, M. A study on attributes for 2.5D print quality assessment. In Proceedings of the Color and Imaging Conference, Online, 4–19 November 2020; Volume 2020, pp. 19–24.
2. Janssen, T.; Blommaert, F. Predicting the usefulness and naturalness of color reproductions. *J. Imaging Sci. Technol.* **2000**, *44*, 93–104.
3. Halonen, R.; Westman, S.; Oittinen, P. Naturalness and interestingness of test images for visual quality evaluation. In *Image Quality and System Performance VIII*; International Society for Optics and Photonics: San Francisco, CA, USA, 2011; Volume 7867, p. 78670Z.
4. Nyman, G.; Häkkinen, J.; Koivisto, E.M.; Leisti, T.; Lindroos, P.; Orenius, O.; Virtanen, T.; Vuori, T. Evaluation of the visual performance of image processing pipes: Information value of subjective image attributes. In *Image Quality and System Performance VII*; International Society for Optics and Photonics: San Jose, CA, USA, 2010; Volume 7529, p. 752905.
5. Leisti, T.; Radun, J.; Virtanen, T.; Halonen, R.; Nyman, G. Subjective experience of image quality: Attributes, definitions, and decision making of subjective image quality. In *Image Quality and System Performance VI*; International Society for Optics and Photonics: San Jose, CA, USA, 2009; Volume 7242, p. 72420D.

6. Radun, J.; Leisti, T.; Häkkinen, J.; Ojanen, H.; Olives, J.L.; Vuori, T.; Nyman, G. Content and quality: Interpretation-based estimation of image quality. *ACM Trans. Appl. Percept. (TAP)* **2008**, *4*, 1–15. [CrossRef]
7. Virtanen, T.; Nuutinen, M.; Häkkinen, J. Image quality wheel. *J. Electron. Imaging* **2019**, *28*, 013015. [CrossRef]
8. Yoshida, A.; Blanz, V.; Myszkowski, K.; Seidel, H.P. Perceptual evaluation of tone mapping operators with real-world scenes. In *Human Vision and Electronic Imaging X*; International Society for Optics and Photonics: San Jose, CA, USA, 2005; Volume 5666, pp. 192–203.
9. Yeganeh, H.; Wang, Z. Structural fidelity vs. naturalness-objective assessment of tone mapped images. In *Proceedings of the International Conference Image Analysis and Recognition*; Springer: Burnaby, BC, Canada, 22–24 June 2011; pp. 111–121.
10. Pedersen, M.; Bonnier, N.; Hardeberg, J.Y.; Albregtsen, F. Attributes of image quality for color prints. *J. Electron. Imaging* **2010**, *19*, 011016.
11. Fedorovskaya, E.A.; Blommaert, F.J.; Ridder, H.d. Perceptual quality of color images of natural scenes transformed in CIELUV color space. In *Color and Imaging Conference*; Society for Imaging Science and Technology: Kilworth Lane Springfield, VA, USA 1993; Volume 1993, pp. 37–40.
12. Drago, F.; Martens, W.L.; Myszkowski, K.; Seidel, H.P. Perceptual evaluation of tone mapping operators. In Proceedings of the ACM SIGGRAPH 2003 Sketches & Applications, San Diego, CA, USA, 27–31 July 2003; p. 1.
13. Choi, S.; Luo, M.; Pointer, M. Modelling image naturalness. In Proceedings of the Conference on Colour in Graphics, Imaging, and Vision, Terrassa, Spain, 9–13 June 2008; pp. 126–131.
14. Yendrikhovskij, S.N.; Blommaert, F.; De Ridder, H. Color reproduction and the naturalness constraint. *Color Res. Appl.* **1999**, *24*, 52–67. [CrossRef]
15. Tian, D.; Peng, R.; Chang, C.; Luo, M.R. Models to Predict Naturalness of Images Including Skin Colours. In *Advances in Graphic Communication, Printing and Packaging Technology and Materials*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 30–35.
16. Tian, D.; Ji, C.; Luo, M.R. Image naturalness model based on prototypical memory colors (sky, grass, and skin). *Color Res. Appl.* **2022**, 1–14. [CrossRef]
17. Seuntiens, P. Visual Experience of 3D TV. Ph.D. Thesis, Eindhoven University of Technology, Eindhoven, The Netherlands, 2006.
18. Overvliet, K.E.; Soto-Faraco, S. I cannot believe this is not wood! An investigation in the perception of naturalness. *Acta Psychol.* **2011**, *136*, 95–111. [CrossRef] [PubMed]
19. Sharan, L.; Rosenholtz, R.; Adelson, E. Material perception: What can you see in a brief glance? *J. Vision* **2009**, *9*, 784. [CrossRef]
20. 3D Textures. Available online: https://3dtextures.me (accessed on 6 July 2021).
21. Winkler, S. Analysis of public image and video databases for quality assessment. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 616–625. [CrossRef]
22. Kitanovski, V.; Pedersen, M. Orientation modulation for data hiding in chrominance channels of direct binary search halftone prints. In Proceedings of the Color and Imaging Conference, San Diego, CA, USA, 7–11 November 2016; Volume 2016, pp. 32–40.
23. Series, B.T. Methodology for the subjective assessment of the quality of television pictures. In *Recommendation ITU-R BT*; International Telecommunication Union: Switzerland, Geneva, 2012; pp. 500–513.
24. Cui, L.C. Do experts and naive observers judge printing quality differently? In *Image Quality and System Performance*; International Society for Optics and Photonics: San Jose, CA, USA, 2003; Volume 5294, pp. 132–145.
25. Green, P.; MacDonald, L. *Colour Engineering: Achieving Device Independent Colour*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
26. Bertalmío, M. *Vision Models for High Dynamic Range and Wide Colour Gamut Imaging: Techniques and Applications*; Academic Press: Cambridge, MA, USA, 2019.
27. Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilita. *Pubbl. R. Ist. Super. Sci. Econ. Commericiali Firenze* **1936**, *8*, 3–62.
28. Krasula, L.; Fliegel, K.; Le Callet, P.; Klíma, M. Objective evaluation of naturalness, contrast, and colorfulness of tone-mapped images. In *Applications of Digital Image Processing XXXVII*; International Society for Optics and Photonics: San Diego, CA, USA, 2014; Volume 9217, p. 92172D.
29. Seuntiëns, P.J.; Heynderickx, I.E.; IJsselsteijn, W.A.; van den Avoort, P.M.; Berentsen, J.; Dalm, I.J.; Lambooij, M.T.; Oosting, W. Viewing experience and naturalness of 3D images. In *Three-Dimensional TV, Video, and Display IV*; International Society for Optics and Photonics: Boston, MA, USA, 2005; Volume 6016, p. 601605.

Article C

What elevation makes 2.5D prints perceptually natural?

Altynay Kadyrova, Marius Pedersen, Stephen Westland,
Materials,
Vol. 15, No. 10, p. 3573,
May, 2022.

materials

Article

What Elevation Makes 2.5D Prints Perceptually Natural?

Altynay Kadyrova [1,*], Marius Pedersen [1] and Stephen Westland [2]

[1] Department of Computer Science, Norwegian University of Science and Technology, 2802 Gjøvik, Norway; marius.pedersen@ntnu.no
[2] School of Design, University of Leeds, Leeds LS2 9JT, UK; s.westland@leeds.ac.uk
* Correspondence: altynay.kadyrova@ntnu.no

Abstract: Elevation plays a considerable role in naturalness perception of 2.5D prints. The necessary level of elevation to make 2.5D prints look perceptually natural may vary from application to application. Therefore, one needs to know the right elevation for specific applications to make the prints look perceptually natural. In this work, we investigated what elevation makes 2.5D prints of wood images perceptually natural. We worked with various wood content images such as wooden wicker, wall, roof, and floor. We found that the optimal elevation that makes 2.5D prints of wood images perceptually natural is content-dependent and in a range between 0.3 mm and 0.5 mm. Moreover, we found that the optimal elevation becomes 0.5 mm if we consider images of wood regardless of the wood content. In addition, there was a high correlation between majority of observers on naturalness perception of 2.5D prints of wood images.

Keywords: naturalness; 2.5D printing; elevation; wood images

check for updates

Citation: Kadyrova, A.; Pedersen, M.; Westland, S. What Elevation Makes 2.5D Prints Perceptually Natural? *Materials* **2022**, *15*, 3573. https://doi.org/10.3390/ma15103573

Academic Editor: Young-Hag Koh

Received: 1 May 2022
Accepted: 16 May 2022
Published: 17 May 2022

1. Introduction

Naturalness of 2.5D prints is important for industry because 2.5D prints need to provide realistic representation of the content they depict. Naturalness perception of 2.5D prints can be affected by various factors (e.g., illumination, viewing angle, user experience, ink types, etc.) and the quality attributes of the prints (e.g., elevation, color, gloss, etc.). For example, Kadyrova et al. [1] found that elevation affects the naturalness perception of 2.5D prints.

Elevation level tends to change perceived appearance aspects such as the naturalness of 2.5D prints. We found from our previous work that observers perceive 2.5D prints of wood images to be natural at lower elevation levels [1]. As a result, it is relevant to investigate the exact elevation level that is perceived as natural for 2.5D prints of wood images. Hence, this work seeks to provide detailed information on this. We focus on various elevation levels at a given viewing distance and illumination. Images of wood material were selected because wood is a familiar material for most people and has a variety of forms [2]. Our work may provide industry with valuable insights on what elevation levels to use for 2.5D prints of wood images' content so that they are perceived as natural. Wood is also a commonly used reproduced material in decor printing and is therefore also industry relevant. It is important to mention that, in the case of 2.5D prints, elevation can be synonymous with height, relief, depth, and similar attributes and expressions [1,3]. Thus, we do not differentiate definitions of, for example, elevation and height as in geography or other fields. By elevation we mean a height that is a distance from the bottom (i.e., substrate) to the top (i.e., raised point).

This paper is organized as follows: background information about the naturalness perception of 2.5D prints is given first; methodology is given afterwards followed by the results and discussion; last, conclusions and suggestions for future work are given.

2. Background

Naturalness perception tends to be multidimensional [2]. In other words, it is complex and varies depending on what is examined as natural. For example, considerable research has been carried out to explore wood's naturalness perception [2,4,5] and the impact of wood on human well-being [6–8]. Overvliet and Soto-Faraco [2] studied the impact of vision and touch, and their combination on naturalness perception of wood samples with varying treatment levels through four psychophysical methods (ranked ordering, binary decision, magnitude estimation, and labelled scaling). They found consistent results across the four methods. They also found that visual and tactile assessments show a high correlation with visuo-tactile assessment. However, they mentioned that it is challenging to separate the effects of vision and tactile assessment on material's naturalness perception. Strobel et al. [9] investigated the link between wood's physical properties and general knowledge on interior wood products. They found several physical properties that their observers used for wood assessment (e.g., grain, color, chemical composition, etc.) as well as other properties that had impact on wood use in interior such as noise, scent, flammability, warmth, and feeling. They further stated that scent and grain are the two main properties that impact on naturalness perception of wood. Their results are important for the reproduction of wooden materials. In particular, feeling wood is the key property for the reproduction of wooden materials that should be considered by industry.

Some wood types can be expensive and therefore their realistic reproduction can be helpful in different applications, especially in decor or in education for architecture students. The reproduction can be achieved in different ways. For decor, the relevant technique for reproduction can be 2.5D printing. It allows superimposition of successive ink layers to create surface relief (i.e., elevation) and can achieve better fine detail reproduction than 3D printing [10]. In other words, 2.5D printing creates slightly elevated prints. To reproduce wood via 2.5D printing, one needs to input wood images and their height maps, and desired elevation level (additionally, one can define settings for gloss and color parameters). The output is 2.5D prints of wood images fabricated at the desired elevation level. The substrate for printing can be real wood or other substitutes depending on the need. One can use images of any material (e.g., stone, metal) to reproduce those materials via 2.5D printing. It is reasonable to assess naturalness perception of the 2.5D reproductions/prints of different material images because one might be curious if 2.5D printing was able to realistically reproduce, for example, wood images on either wood or another substrate.

There is limited research that has investigated the naturalness perception of 2.5D prints. The work of Kadyrova et al. [1] demonstrates the effect of elevation and surface roughness on the naturalness perception of 2.5D decor prints. The elevation (i.e., height) levels they used were 0.4 mm, 0.6 mm, and 0.8 mm while the surface roughness was considered as a height difference within a local neighborhood. They found that there is an impact by elevation rather than surface roughness on the naturalness perception of 2.5D decor prints. Moreover, they found that the naturalness of 2.5D decor prints is content-dependent with regard to the elevation effect. Images of four material categories were examined in their work. They were wood, glass, stone, and metal, and 2.5D prints were fabricated on an outdoor paper substrate. Based on their results, the observers found lower elevation as natural for wood and glass 2.5D prints. For the stone and metal categories, they did not find a clear tendency. They also provided perceptual attributes that one tends to use during naturalness assessment of 2.5D prints. They defined three levels of perceptual attributes. There were twelve main attributes in total. Color, roughness, gloss, elevation, and lightness were the top five. Furthermore, they found that the most used perceptual attributes among the four examined material categories were color, roughness, and gloss.

The quality attributes most commonly used for the quality assessment of 2.5D prints were studied by Kadyrova et al. [3]. The top five were color, sharpness, elevation, lightness, and naturalness. This shows that observers tend to look at elevation and naturalness aspects of 2.5D prints during quality assessment. Hence, this supports the point that it is important to investigate the naturalness perception of 2.5D prints with respect to the elevation. Furthermore, they observed the following factors that might affect the quality assessment of 2.5D prints: content, aesthetic appearance, and previous knowledge and experience.

A model for assessing 3D display visual performance proposed by Heynderickx [11] shows that naturalness spans both perceived depth and image quality. In the case of 2.5D prints, observers tend to use depth to mean elevation [1,3]. Hence, we can view depth as elevation in this model for 2.5D prints' case or we can slightly modify the model for 2.5D prints' case as illustrated in Figure 1. The elevation is the main feature of 2.5D prints and it impacts the naturalness perception of 2.5D prints [1]. As a result, it can be considered as the representative attribute that describes the naturalness perception of 2.5D prints.

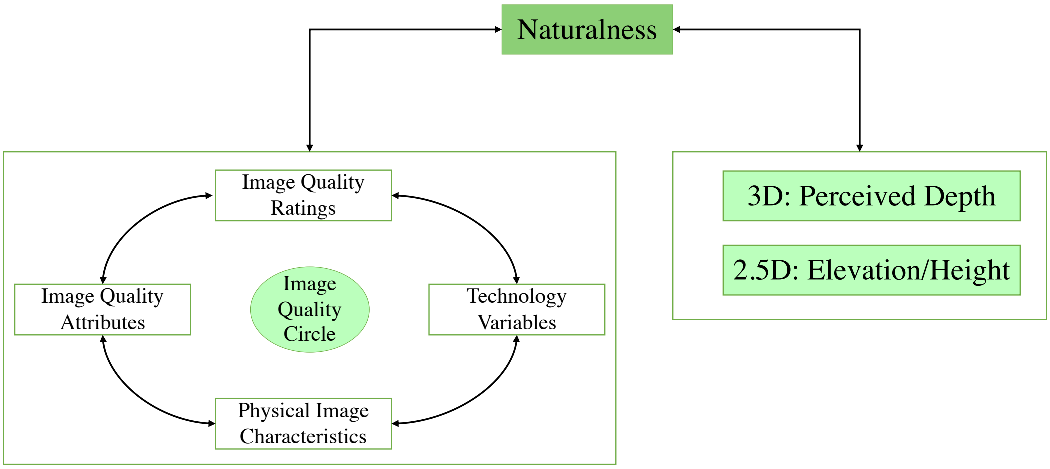

Figure 1. Slight modification proposal of Heynderickx [11] model with regard to 2.5D prints. We propose to consider elevation (or height) as the representative attribute that affects the naturalness perception of 2.5D prints in this model.

To conclude, the literature supports the need to explore the effect of elevation on the naturalness perception of 2.5D prints.

3. Methodology

Following our previous work [1], we refer to the term realistic representation of a print for the definition of naturalness. We considered 20 wood images with various wood content. We included images of wooden wicker (8 images), floor (4 images), roof (2 images), and wall (6 images) (Figure 2). The original color images and their height maps (both are 782 × 782 pixels) were reproduced from 3D textures (copyright free web site) [12]. The height maps undergone processing such as Gaussian filtering with a standard deviation of four and intensity adjustment. The reasons for this processing were to reduce black edges and reach the intended maximum elevation.

Based on the results of Kadyrova et al. [1], 2.5D prints of wood images were found to be perceived as more natural at 0.4 mm than at 0.6 mm and 0.8 mm. It is interesting to check whether 0.4 mm or elevation level around this number (i.e., 0.5 mm or 0.3 mm) provides a natural look to 2.5D prints of wood images. As a result, we varied elevation levels between 0 mm and 0.5 mm, meaning that each image had 6 reproductions. The reason for including flat prints (i.e., 0 mm) was to check if observers still prefer flat prints over elevated ones (i.e., 2.5D prints) or whether flat prints should be eliminated from the focus. The prints were fabricated with the OCE Arizona 2280GT 2.5D printer on an outdoor paper substrate. We used Alto printer mode (i.e., opaque elevation) and made print size 6.62 × 6.62 cm with an additional 0.3 cm on each side of the substrate paper for observers to hold the prints without touching their surface.

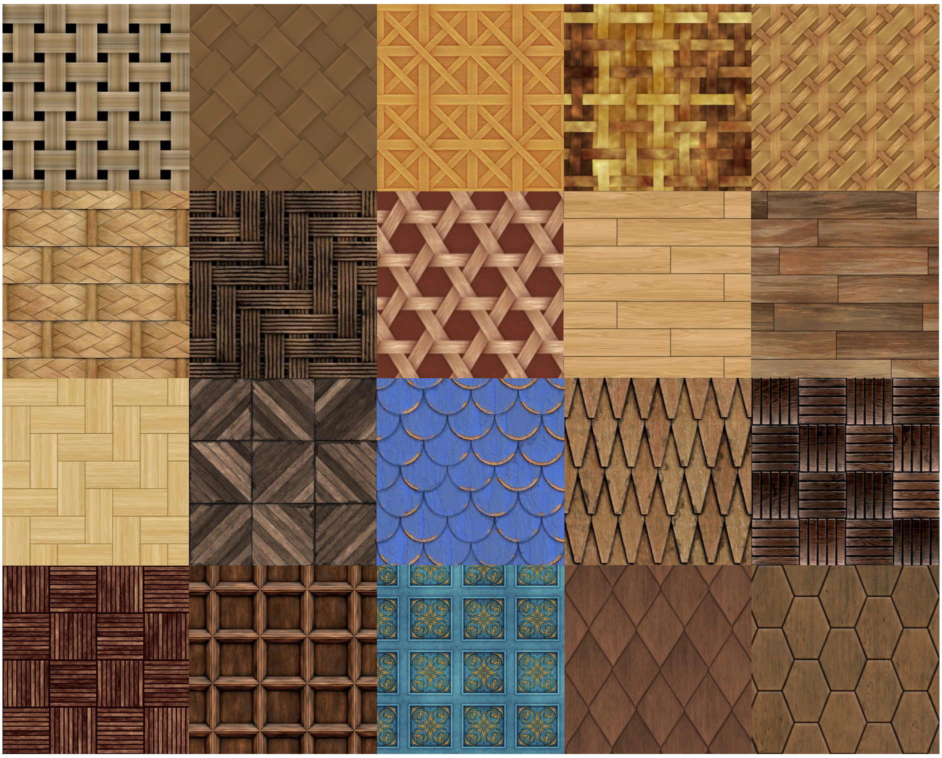

Figure 2. Wood images used in our work. From top left: 8 wicker, 4 floor, 2 roof, and 6 wall images.

The visual experiment was conducted at two locations, UK and Norway, in English. We followed the same experiment design we used in our previous work [1]—first, we acquired consent from the observers followed by adaptation to illumination while the observers were reading instruction; afterwards, a training session was performed. A 3D-printed 45° holder was used to place the prints which were given in random order inside the light booth cabinet (Verivide CAC 60-5, illuminations were 1400 lux (Norway) and 1364 lux (UK)) with D65 illumination. The ranking experiment was conducted with instructions similar to those of our previous work [1]. We asked the observers to rank the 2.5D prints from the most to the least realistic representation of wooden wicker/floor/roof/wall and explain why. Thus, we gave keywords as a reference, and the observers were allowed to move and tilt the prints with provided gloves but not touch the surface. The distance between the eyes of the observers and the prints was approximately 50 cm. The observers were informed that there was no time restriction. The average duration of the experiment was 38 min per observer excluding the training session time (which was approximately 3 min per observer on average). With the exception of one observer, all other observers finished the experiment in one session. The audio responses of the observers were recorded for analysis purposes.

We had 21 observers (15 female and 6 male with an average age around 35 years and a standard deviation around 11 years) with normal color vision. We checked their color vision and visual acuity with Ishihara plates and a Snellen chart, respectively. There were 10 Asians and 11 Europeans who represented both naive and experienced (from computer science background) observers.

4. Results and Discussion

To determine what elevation makes 2.5D prints perceptually natural, we provide Z-scores (acquired from raw ranked data) of observers on the naturalness perception of the 2.5D prints at various elevation levels. We use an error bar plot to visualize Z-scores. Mean Z-scores are shown by a circle in the centre of the vertical lines. The Confidence Interval (*CI*) was derived using Equation (1) [13].

$$CI = 1.96 \cdot \frac{\sigma}{\sqrt{N}}, \tag{1}$$

where N is the number of observations, and σ is the standard deviation which in the case of Z-score can be computed as $1/\sqrt{2}$ [14]. 95% *CI* is the mean Z-scores $\pm$ *CI*. If two *CI* do not overlap, then there is a statistically significant difference between reproductions with 95% confidence.

From Figure 3 for the results for all wood images, we can see a clear trend that observers found higher elevation prints more natural than flatter ones. A binomial sign test was used on the raw data with Bonferroni correction (with a significance level of α/n, where $\alpha = 0.05$ is the desired alpha value and n is the number of comparisons: 0.05/15) [15] to check statistically significant differences between elevation levels. According to the p-values (Table 1), each elevation level had a statistically significant difference between each other. Based on the results, the optimal elevation that makes 2.5D prints of wood images look perceptually natural was 0.5 mm. However, if we consider the Z-scores of all the images in each wood content, then there was a slight variation regarding the optimal elevation in the case of floor and wall images while wicker and roof images showed that 0.5 mm is the optimal elevation. However, the overall trend that higher elevation prints are more natural than flatter ones stays the same whether we consider all wood images or images in each wood content. We present the results of each wood content image below.

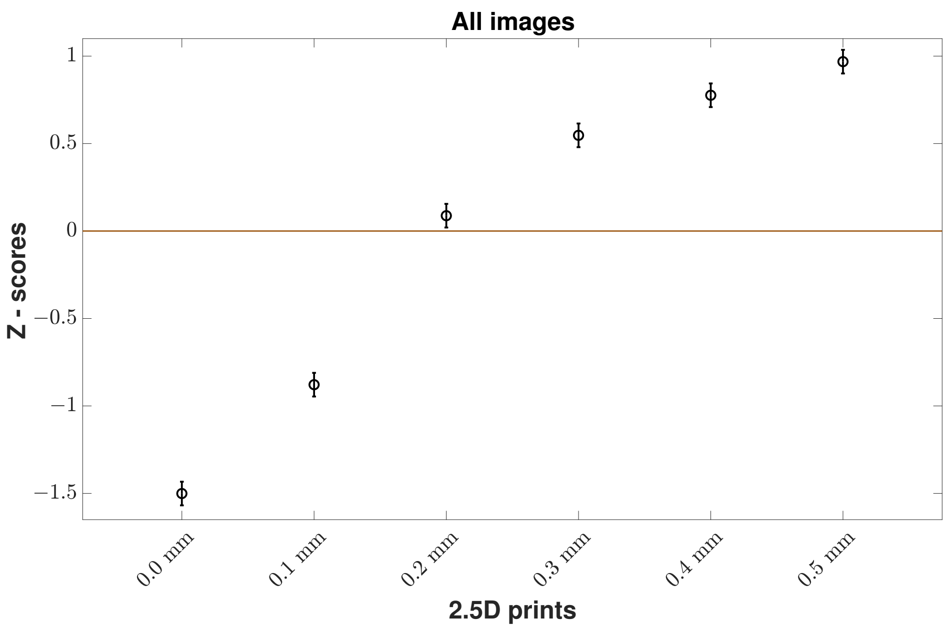

Figure 3. Z-scores of all wood images by all observers. Mean Z-score values for 2.5D prints at various elevation levels (x-axis) are given with 95% *CI*s. It shows that the observers found 2.5D prints with 0.5 mm as more natural looking than with no elevation.

Table 1. *p*-values obtained by a sign test for all wood images. Green cells are those that have a statistically significant difference. The threshold used in the Bonferroni correction is 0.05/15 = 0.0033.

	0.0 mm	0.1 mm	0.2 mm	0.3 mm	0.4 mm	0.5 mm
0.0 mm	-	1.1602×10^{-52}	6.1431×10^{-78}	5.1105×10^{-74}	1.0257×10^{-71}	3.3530×10^{-70}
0.1 mm		-	1.8900×10^{-69}	5.8364×10^{-68}	7.0823×10^{-63}	2.2802×10^{-59}
0.2 mm			-	4.7035×10^{-41}	1.0702×10^{-37}	2.3412×10^{-39}
0.3 mm				-	1.7632×10^{-19}	2.0207×10^{-24}
0.4 mm					-	8.2898×10^{-14}
0.5 mm						-

In the case of the *p*-values of all wooden floor images (Table 2), each elevation level had a statistically significant difference between each other except 0.4 mm and 0.5 mm. From Figure 4 for the results for all wooden floor images, we can see that 2.5D prints with higher elevation were found to be more natural than those at 0 mm by the observers.

Table 2. *p*-values obtained by a sign test for all wooden floor images. Green cells are those that have a statistically significant difference whereas red cells are those that do not have a statistically significant difference. The threshold used in the Bonferroni correction is 0.05/15 = 0.0033.

	0.0 mm	0.1 mm	0.2 mm	0.3 mm	0.4 mm	0.5 mm
0.0 mm	-	7.8639×10^{-11}	2.0973×10^{-19}	3.4018×10^{-18}	3.4018×10^{-18}	5.1376×10^{-16}
0.1 mm		-	4.5431×10^{-17}	5.1376×10^{-16}	4.3086×10^{-14}	4.3086×10^{-14}
0.2 mm			-	1.9382×10^{-9}	1.3329×10^{-7}	8.5423×10^{-9}
0.3 mm				-	2.6645×10^{-4}	1.0715×10^{-4}
0.4 mm					-	0.2299
0.5 mm						-

In the case of the *p*-values of all wooden wall images, there was no statistically significant difference in elevation levels between 0.3 mm and 0.5 mm. From Figure 4 for the results for all wooden wall images, we can see that observers preferred 2.5D prints with higher elevation as more natural than those at 0 mm.

In the case of the *p*-values of all wooden wicker images, each elevation level had a statistically significant difference between each other. From Figure 5 for the results for all wooden wicker images, we can see that 2.5D prints with 0.5 mm were found to be more natural than those at 0 mm by the observers.

In the case of the *p*-values of all wooden roof images, each elevation level had a statistically significant difference between each other except 0.3 mm and 0.4 mm. From Figure 5 for the results for all wooden roof images, we can see that 2.5D prints with higher elevation were found to be more natural than those at 0 mm by observers. It is important to mention that we had a low number of images in wooden roof content.

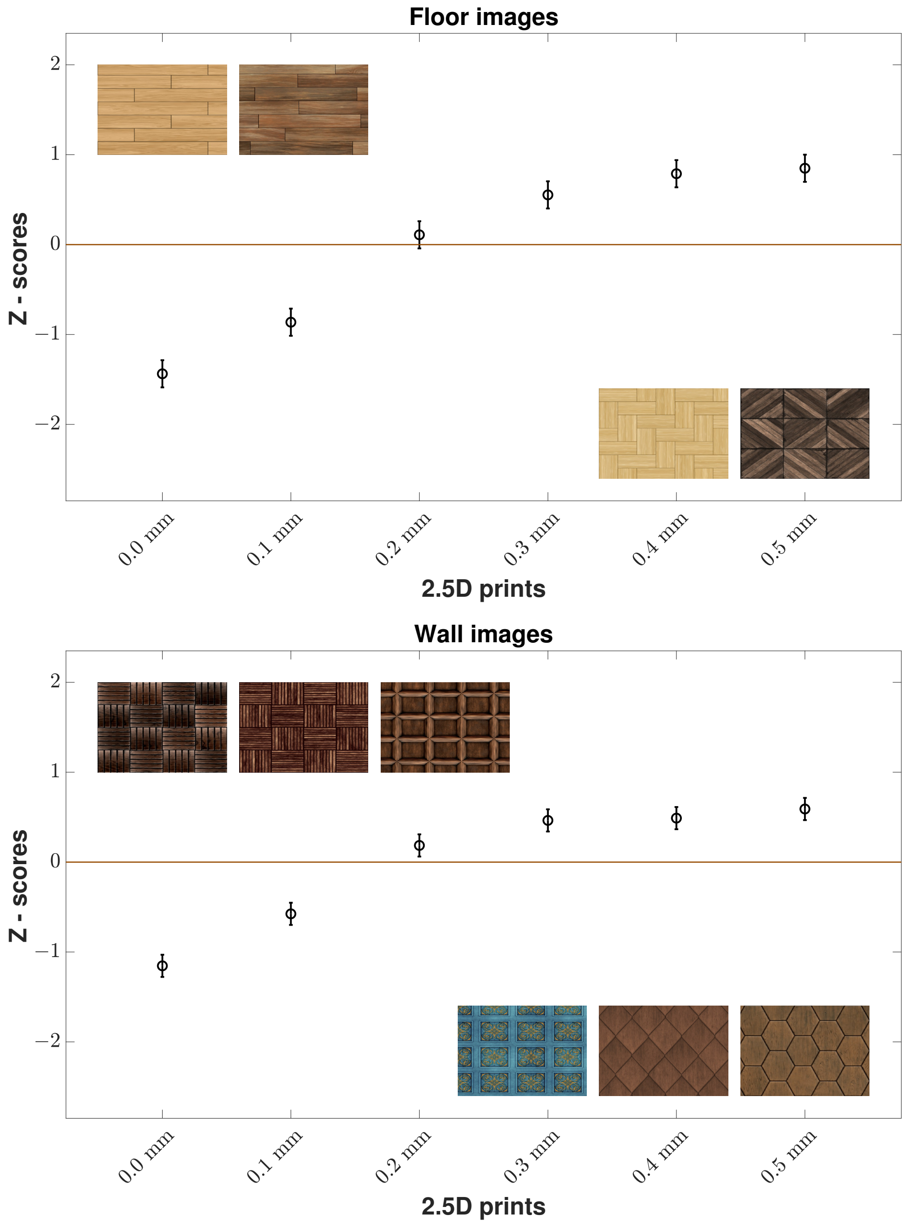

Figure 4. *Z*-scores of all wooden floor and all wooden wall images by all observers. Mean *Z*-score values for 2.5D prints at various elevation levels (x-axis) are given with 95% *CI*s. We can observe that for both floor and wall images, the observers found flat prints as the least natural. There are 4 floor and 6 wall images.

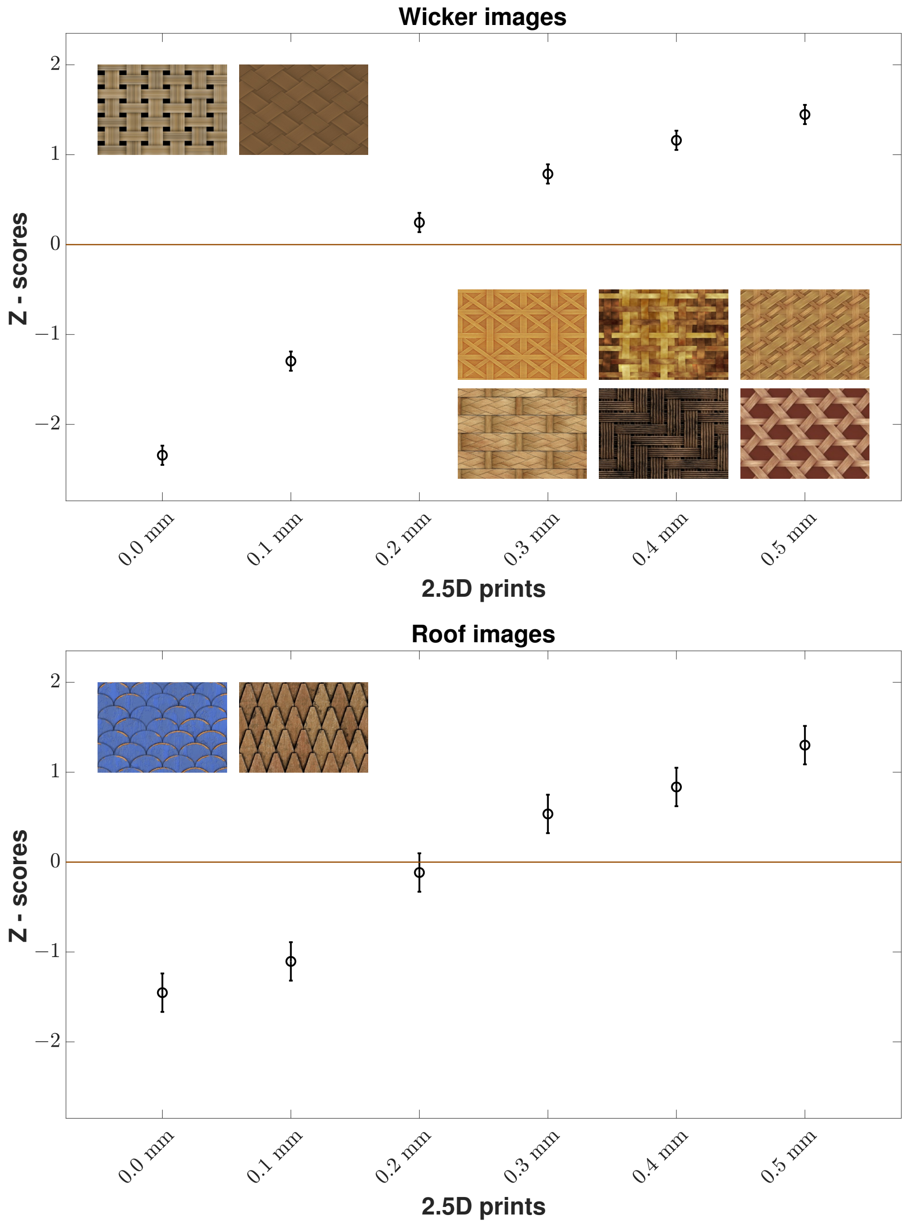

Figure 5. *Z*-scores of all wooden wicker and all wooden roof images by all observers. Mean Z-score values for 2.5D prints at various elevation levels (x-axis) are given with 95% *CIs*. We can observe that for both wicker and roof images, the observers found 2.5D prints with 0.5 mm as the most natural. There are 8 wicker and 2 roof images.

As a result, we can observe that the optimal elevation that makes the 2.5D prints of wood images look perceptually natural is somewhat content-dependent. More specifically, when a certain elevation is reached for some content such as images of wooden floor or wall, that elevation is perceived as natural by the observers. In the case of images of wooden wicker and roof, higher elevation makes the most natural looking 2.5D prints.

Furthermore, we analyzed each observer's Z-score in the case of all wood images and found that the majority preferred 0.5 mm as the one that makes the most natural looking 2.5D prints. We also checked inter-observer variability by Spearman correlation coefficient. On average for all images, majority of observers showed an agreement between each other on their rankings. This is different from the inter-observer variability results in Kadyrova et al.'s [1] work. We find this difference reasonable due to the following reasons: first, we used in our work only wood images while they used images of four material categories (wood, stone, glass, and metal); second, we varied only elevation while they varied elevation and surface roughness in the prints; last, the task in our case was somewhat easy for the observers to perceive differences between elevation levels than in their work. Additionally, we found similar performance when comparing the results for UK and Norway observers as well as between Asians and Europeans. The recorded audio data showed that most observers were able to find that the varying parameter was the elevation. They used a wide range of attributes and words to describe the elevation such as elevation, height, relief, depth, coming out, etc.

The limitation of our work is that it is based on images of wood only. Nevertheless, we provide a workflow that can be followed for other types of application where the output results may vary from application to application. In other words, the optimal elevation that makes 2.5D prints, for example, of stone images perceptually natural, might be different from the one found for wood images.

To conclude, it was clear that the flat prints do not look perceptually natural to observers. There should be elevation to make the prints look perceptually natural. In addition, we found that the content plays a role in finding the optimal elevation for the specific content of 2.5D prints of wood images to look perceptually natural. Moreover, there was a high correlation between majority of observers on their rankings which shows that the observers are rather consistent when assessing naturalness of 2.5D prints of wood images.

5. Conclusions and Future Work

The naturalness of 2.5D prints is affected by the level of elevation and is application-dependent. Therefore, we studied what elevation makes 2.5D prints of wood images to be perceived as natural. Various wood content images were considered such as wooden floor, wall, roof, and wicker. Based on the results, the optimal elevation is found to be content-dependent and in a range between 0.3 mm and 0.5 mm. More specifically, one can find the optimal elevation to use when certain elevation is reached for certain content of wood images. However, if one is interested in wood images regardless of content within wood images, then we found that 0.5 mm is the optimal elevation to use. Moreover, it was clear that observers found flat prints to be the least natural. Majority of observers showed a high correlation on their rankings, meaning that they are fairly consistent when it comes to 2.5D prints' naturalness perception of wood images. Future work will be to study the naturalness perception of 2.5D prints further in terms of the effect of various types of ink because we hypothesize that the core effect on the naturalness perception of 2.5D prints might come from the ink itself that is used to fabricate the prints.

Author Contributions: Conceptualization, A.K., M.P. and S.W.; methodology, A.K., M.P. and S.W.; data acquisition and processing, A.K.; analysis, A.K.; writing—original draft, A.K.; writing—review and editing, A.K., M.P. and S.W.; visualization, A.K.; supervision, M.P. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the ApPEARS project from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 814158.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original images and their height maps used in this work can be found at https://3dtextures.me, accessed on 6 July 2021.

Acknowledgments: We would like to thank observers for participation in the experiment.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Kadyrova, A.; Pedersen, M.; Westland, S. Effect of Elevation and Surface Roughness on Naturalness Perception of 2.5D Decor Prints. *Materials* **2022**, *15*, 3372. [CrossRef]
2. Overvliet, K.E.; Soto-Faraco, S. I cannot believe this is not wood! An investigation in the perception of naturalness. *Acta Psychol.* **2011**, *136*, 95–111. [CrossRef] [PubMed]
3. Kadyrova, A.; Kitanovski, V.; Pedersen, M. A study on attributes for 2.5D print quality assessment. In Proceedings of the Color and Imaging Conference, Online, 4–19 November 2020; Society for Imaging Science and Technology: Springfield, VA, USA, 2020; Volume 2020, pp. 19–24.
4. Burnard, M.D.; Nyrud, A.Q.; Bysheim, K.; Kutnar, A.; Vahtikari, K.; Hughes, M. Building material naturalness: perceptions from Finland, Norway and Slovenia. *Indoor Built Environ.* **2017**, *26*, 92–107. [CrossRef]
5. Whitaker, T.A.; Simões-Franklin, C.; Newell, F.N. The natural truth: The contribution of vision and touch in the categorisation of "naturalness". In Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, Madrid, Spain, 10–13 June 2008; pp. 319–324.
6. Rice, J.; Kozak, R.A.; Meitner, M.J.; Cohen, D.H. Appearance wood products and psychological well-being. *Wood Fiber Sci.* **2006**, *38*, 644–659.
7. Poirier, G.; Demers, C.M.; Potvin, A. Wood perception in daylit interior spaces: An experimental study using scale models and questionnaires. *BioResources* **2019**, *14*, 1941–1968.
8. Li, J.; Wu, J.; Lam, F.; Zhang, C.; Kang, J.; Xu, H. Effect of the degree of wood use on the visual psychological response of wooden indoor spaces. *Wood Sci. Technol.* **2021**, *55*, 1485–1508. [CrossRef]
9. Strobel, K.; Nyrud, A.Q.; Bysheim, K. Interior wood use: linking user perceptions to physical properties. *Scand. J. For. Res.* **2017**, *32*, 798–806. [CrossRef]
10. Page, M.; Obein, G.; Boust, C.; Razet, A. Adapted modulation transfer function method for characterization and improvement of 3D printed surfaces. In Proceedings of the Electronic Imaging, San Francisco, CA, USA, 29 January–2 February 2017; pp. 92–100.
11. Heynderickx, I. Performance evaluation of 2D and 3D-TV systems. In Proceedings of the ICIS'06: 30th International Congress of Imaging Science, Rochester, NY, USA, 7–11 May 2006; pp. 374–377.
12. 3D Textures. Available online: https://3dtextures.me (accessed on 6 July 2021).
13. Green, P.; MacDonald, L. *Colour Engineering: Achieving Device Independent Colour*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
14. Bertalmío, M. *Vision Models for High Dynamic Range and Wide Colour Gamut Imaging: Techniques and Applications*; Academic Press: Cambridge, MA, USA, 2019.
15. Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilita. *Pubbl. Del Ist. Super. Sci. Econ. Commericiali Firenze* **1936**, *8*, 3–62.

Article D

Effect of various ink types on naturalness perception of 2.5D prints

Altynay Kadyrova, Marius Pedersen, Stephen Westland, Clemens Weijkamp,
Under review in a journal,
18 pages.

This article is under review for publication and is therefore not included.

Article E

Quality assessment of 2.5D prints using 2D image quality metrics

Altynay Kadyrova, Vlado Kitanovski, Marius Pedersen,
Applied Sciences,
Vol. 11, No. 16, p. 7470,
August, 2021.

applied sciences

MDPI

Article

Quality Assessment of 2.5D Prints Using 2D Image Quality Metrics

Altynay Kadyrova *, Vlado Kitanovski and Marius Pedersen

Department of Computer Science, Norwegian University of Science and Technology, 2802 Gjøvik, Norway; vlado.kitanovski@ntnu.no (V.K.); marius.pedersen@ntnu.no (M.P.)
* Correspondence: altynay.kadyrova@ntnu.no

Abstract: Quality assessment is an important aspect in a variety of application areas. In this work, the objective quality assessment of 2.5D prints was performed. The work is done on camera captures under both diffuse (single-shot) and directional (multiple-shot) illumination. Current state-of-the-art 2D full-reference image quality metrics were used to predict the quality of 2.5D prints. The results showed that the selected metrics can detect differences between the prints as well as between a print and its 2D reference image. Moreover, the metrics better detected differences in the multiple-shot set-up captures than in the single-shot set-up ones. Although the results are based on a limited number of images, they show existing metrics' ability to work with 2.5D prints under limited conditions.

Keywords: image quality metrics; quality assessment; 2.5D printing

check for updates

Citation: Kadyrova, A.; Kitanovski, V.; Pedersen, M. Quality Assessment of 2.5D Prints Using 2D Image Quality Metrics. *Appl. Sci.* **2021**, *11*, 7470. https://doi.org/10.3390/app11167470

Academic Editor: Enrico Vezzetti

Received: 10 July 2021
Accepted: 12 August 2021
Published: 14 August 2021

1. Introduction

Objective image Quality Assessment (QA) has a significant demand because it is automatic, consistent, and less resource demanding compared to subjective image QA. There are studies that assess the quality of 2D print images objectively [1,2]. However, 2.5D reproduction QA in an objective way is less studied. Therefore, the goal of the current work is to investigate if existing 2D Image Quality Metrics (IQMs) are suitable to assess the perceptual quality of 2.5D reproductions. At this moment, to our knowledge, no IQM is standardized for 2.5D reproduction QA. As a result, existing IQMs should be tested even though it is expected that it will be a difficult scenario for them to assess the quality of 2.5D prints. Moreover, different options to digitize prints should be tested to determine which data representation is more appropriate for IQMs. Consequently, we also address the question of how to digitize 2.5D prints for QA. The acquisition set-up and light source might influence the captured data, but information regarding the responsiveness of the IQMs on the quality variations that 2.5D prints hold can still be valuable, especially in (serial) production of prints. Thus, first we check the responsiveness of the selected 2D IQMs on the quality variations that 2.5D prints hold. Next, we analyze the quality maps of IQMs because they can reveal more information about print quality in comparison with just IQMs' values. Last, we compare which of the two set-ups we tested is more suitable for 2.5D prints capture based on IQMs' performance.

This paper is organized as follows. First, we give background information about different dimensional prints' (e.g., 2D, 2.5D prints) QA as well as capture approaches. Afterwards, we describe our methodology followed by the results and discussion. Finally, we provide our conclusions and future works.

2. Background

There have been works in improving quality of 2.5D printing including characterization of relief printers [3], printing gloss effects [4], development of a 2.5D printing machine with a software [5], and proposal of a novel approach for a 2.5D printing based on semantic information [6]. Presently, 2.5D printing is used widely in many applications such as decoration (e.g., interior design), signage, and maps (e.g., for visually impaired people), to name a few. Therefore, 2.5D prints quality should be assessed carefully before releasing them into the market for sale. Many studies have explored the quality of images [7–11] and prints [12–14]. The quality of 2.5D prints in this work can be related to the definitions given by Keelan [15] and ISO [16], where image quality is an observer's perception of image excellence where the observer was not involved with anything related to the given image. In addition, 2.5D prints tend to have surface elevation. This can impact on the overall quality perception by customers. The importance of 2.5D prints is that the variations in surface elevation of 2.5D prints can create many possible options for shapes and texture patterns such as diversification of what printer can print as well as enhancement of appearance of prints [17]. This is useful in many applications such as signage, decorations, maps or reproduction of art works (e.g., to reproduce brush strokes). First, brief QA aspects and then capture aspects will be described for 2D and 2.5D prints.

2.1. Quality Assessment

In general, QA of any print can be either subjective or objective, or a combination of both. For example, Pedersen et al. [18] conducted a subjective experiment to identify meaningful image quality attributes for 2D color prints' QA. Their observers' task was to rate the quality of 2D color prints and state every quality attribute they used even if some quality attributes had little impact on the QA. The rating task considered seven scale levels, where a value of 1 meant that observers found the 2D color prints as the most pleasing, whereas a value of 7 meant the least pleasing. Afterwards, they validated the chosen quality attributes (color, sharpness, lightness, contrast, and artifacts) through another subjective experiment [19].

Regarding 2.5D print subjective QA, an experiment with observers was conducted to study the most used distinct attributes [20]. The observers judged the quality of 2.5D prints that were fabricated with a 2.5D printer. They were asked to rank the quality of 2.5D prints and describe the reasons for their ranking. The experiment comprised of two parts: first, when observers were not provided with the reference images and second, when the reference images were provided. The relevant attributes were proposed to be the top five most used distinct attributes in their experiment. These were color, sharpness, elevation, lightness, and naturalness. Samadzadegan et al. [21] performed a subjective experiment with 2.5D prints and their goal was to find the effect of color on gloss. According to their results, color has no significant effect on gloss. Nevertheless, they suggested that color and gloss need to be taken into account for print QA.

The objective QA aspect can involve a variety of metrics. There are many diverse sets of 2D IQMs depending on the application area, attributes, performance and accuracy level, and availability of reference images. According to the literature [1,22–26], Human Visual System (HVS)-based metrics are better than those metrics without using a model of the HVS. Additionally, visual quality can be better predicted by the HVS-based metrics rather than simple pixel-based difference metrics [27]. There are Full-Reference (FR), reduced-reference, and No-Reference (NR) IQMs. Most of the IQMs are based on detecting distortions and predicting quality based on that. Due to the unavailability of reference images in practice, sometimes NR metrics are preferable. However, FR metrics can be more straightforward to detect distortions because of reference image availability [28]. Thus, there are pros and cons of each type of metric. Moreover, it is important to mention that these days it is becoming a trend to use deep learning or machine learning methods to create IQMs [29–32]. For instance, Akyazi et al. [28] created a FR metric that takes as input-filtered reference and distorted images and is based on deep neural networks.

There are several frameworks for QA of 2D color prints/images using IQMs [1,33–35]. For example, Pedersen and Amirshahi [1] used Spatial CIELAB (SCIELAB) due to its often use as a reference metric, spatial hue angle metric because it combines two state-of-the-art metrics, adaptive bilateral filter due to the ability of bilateral filtering to simulate the HVS, Structural Similarity (SSIM) due to its common use and working scheme on local neighborhood, and other IQMs for QA of 2D color prints. They concluded that their results are both image and metric dependent.

Baar et al. [36] proposed potential approaches with advantages and disadvantages towards an IQM for 2.5D prints based on their reviewed literature for 3D prints' QA. Liu et al. [3] performed objective QA of 2.5D printing process in height dimension in terms of fidelity and surface finish using modulation transfer function, mean absolute difference, and other metrics. The surface finish difference between real and ideal prints that might be perceived by observers was simulated through their created light-reflection model. They concluded that the surface roughness depends on prints' geometry, increase of frequency of fine details reduces reproduction accuracy, and both viewing angle and illumination direction impact the visual experience.

2.2. Capture Techniques

The IQMs require a digitized version of the reference and/or the reproduction to be used. In the case of printed images, the physical print needs to be digitized prior to QA. Scanners have been used for 2D print digitization [1,2,33,34,37]. Moreover, cameras have been used for QA of flat surfaces [38]. However, scanners are not suitable in this work because we work with non-flat prints (i.e., 2.5D prints with a surface elevation). The results of the work by Zhao et al. [39] support that image QA of flat surfaces (i.e., projection displays) captured through a camera is a relevant approach that works.

Different techniques can be used to capture 2.5D prints. It depends on which attributes one is interested to capture. For example, art paintings are known to have relief surface structure and Zaman et al. [40] were able to capture the topography and color of oil paintings by their proposed hybrid set-up. It consisted of two cameras and a projector. They connected fringe projection with stereo imaging which performed well in color and depth information capturing. The paintings' depth and color information perception with respect to our eyes was mimicked using stereo imaging. The image registration process was avoided by capturing topography and color at the same time. Elkhuizen et al. [41] captured the spatially varying gloss of their hand-painted samples. They used high dynamic range images. Their set-up consisted of a camera and a series of light-emitting diode lights. They concluded that the essential attributes for art paintings are translucency and gloss. High accuracy and precision for painting measurement were acquired using multi-scale optical coherence tomography and 3D digital microscopy [42]. However, these measurement tools were slow due to the small field of view. They also tested 3D scanning based on fringe encoded stereo imaging.

Reflectance Transformation Imaging (RTI) is a common tool to capture appearance under different directional light [43]. The object and camera are fixed perpendicular to each other while the light source moves in RTI. It is common to use RTI for texture visualization purposes of paintings [44] and capturing of low-relief surfaces [45]. For example, Pintus et al. [46] performed RTI of cultural heritage data visualization assessment both subjectively and objectively. Recently, Kitanovski et al. [47] assessed the quality of relighting from images acquired through their proposed multispectral RTI system. They captured 3D objects with various colors and translucencies.

These capture techniques can be applied to different 2.5D prints (maps, signage, etc.). However, some of them might be costly, time consuming, and, most importantly, they might involve post-processing to reconstruct color, depth, gloss, and other attributes that might require (costly) software tools. Overall, capturing the whole appearance features of 2.5D prints or similar objects is challenging. Nonetheless, there are some ongoing works in this direction [48].

3. Methodology

We used the physical 2.5D prints from Kadyrova et al. [20]. They fabricated 42 2.5D prints consisting of 12 reference images with three instances (i.e., in terms of quality aspect variations) and 3 images with two instances by Canon Arizona series 2.5D printer.

The quality variations were divided into five sets in their work. In our work, we focus on three sets because the quality issues were clearly visible/distinguishable for the observers in these sets based on data from Kadyrova et al. [20]. They are as follows (Figure 1):

- Naturalness set with natural elevation, unnatural elevation, and surface roughness prints (tiles, wood, brick images);
- Height set with maximum heights of 1 mm, 0.5 mm, and 0.25 mm prints (scissor, speed sign, running track images);
- Printer mode set with Alto (i.e., when elevation is opaque) and Brila (i.e., when elevation is varnish) modes prints (flower, packaging, snowflake images).

Our workflow is illustrated in Figure 2. We start with acquiring digital data of 2.5D prints followed by preprocessing. Afterwards, relevant IQMs are selected and applied, and data analysis is performed. The results are expected to be from responsiveness test of the IQMs, insights from the IQMs, and comparison of the two capture set-ups.

Figure 1. The 2D reference images, reproduced from the dataset of Kadyrova et al. [20]. From top left: tiles, wood, brick, scissor, speed sign, running track, flower, packaging, and snowflake images.

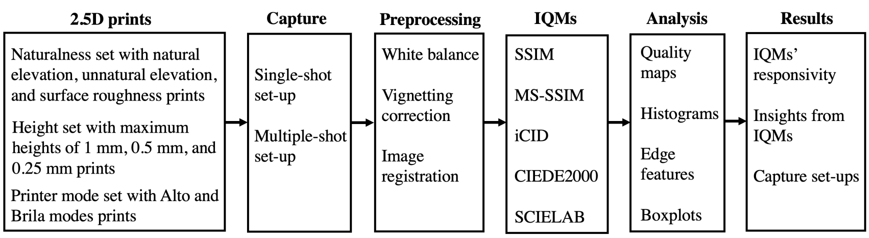

Figure 2. Our workflow. The prints are captured in the single-shot and multiple-shot set-ups, then they are processed, further IQMs are applied, before we do our analysis and report the results.

3.1. Data Capture

There are several options to digitize physical 2.5D prints. In this work, we used camera-based set-ups to acquire digital data of physical 2.5D prints because it is relatively fast and affordable. We included diffuse and multiple angle directional illumination set-ups using a Nikon D610 professional camera alone (single-shot set-up) and RTI with the same camera (multiple-shot set-up), respectively. There was a single viewing angle in both set-ups.

3.1.1. Single-Shot Set-Up

We used a Nikon D610 professional camera with a Sigma 24–105 mm lens placed on a fixed tripod. The distance from the camera lens to the prints was approximately 51 cm. The 2.5D prints and a Macbeth ColorChecker were placed inside the light booth cabinet (VeriVide CAC 60-5, illumination was around 1328 lux) with D65 (diffuse) illumination. We used the following setting parameters for the camera: ISO was 125, the aperture was 4.5, and the shutter speed was 1/80 with manual exposure mode. We worked with the camera jpeg images and the color checker was used for white balancing. The single-shot set-up is further referred to as SS.

3.1.2. Multiple-Shot Set-Up

Because 2.5D prints have elevation and angle dependence appearance due to, for example shadows, RTI was used. The 2.5D prints were digitized by an RTI set-up based on a robotic arm [47]. The camera and prints were fixed (the distance between them was approximately 51 cm) while the light source, mounted on the robotic arm, was moving. We acquired 60 captures based on 60 illumination angles per print. Similar to SS, the camera jpeg images were used and the color checker was used for white balancing (it was captured separately in this set-up). To our knowledge, this is the first attempt to assess the objective quality of 2.5D prints using RTI captures. The multiple-shot set-up is further referred to as MS.

3.2. Preprocessing

We performed vignetting correction (for the SS captures) and image registration (for both set-up captures).

We used the method from Zhao et al. [38] for vignetting correction mask generation for data captured in the SS. We applied the correction mask on the Y channel of YCbCr for all captures. In the MS, the prints were placed relatively in the center and vignetting was less of an issue.

Accurate registration of the captured images to the reference is required for using FR IQMs. Different methods have been used for the registration of 2D print captures [35,38,49] but registration of 2.5D print captures with respect to the 2D reference images was found to be challenging. Nevertheless, manual image registration with homography transformation gave desirable results. The points (coordinates) for the homography were selected manually for each print capture from both set-ups. The algorithm uses bilinear interpolation. The registration was manually checked by the authors to ensure a correct registration.

3.3. Full-Reference Image Quality Metrics

We chose to work with FR IQMs because we have reference images and FR IQMs' universality is higher than that of NR IQMs [50].

The SSIM metric is selected because it is a HVS-based FR metric that considers structural information from the image scene [51]. This metric can be used to assess lightness, contrast, sharpness, and artifact attributes [2]. Additionally, Pedersen and Amirshahi [1] mentioned its potential to detect artifacts. It is used to calculate the perceived difference between the reference and distorted input images and it is independent of both average luminance and contrast [51]. It is appropriate for use with 2.5D prints because it is based on structure and most of the aspects that have been changed in the dataset influence structure.

The Multi-Scale SSIM (MS-SSIM) is selected because it incorporates variations of image resolution and viewing conditions [52]. We used MS-SSIM at scale 2 (scale 1 is the same as SSIM) for both structure and contrast terms because at higher scales it assesses a portion of an image due to downsampling.

The improved Color Image Difference (iCID) is selected because it considers the reference image's color, structure, and contrast [53]. The metric was used with 51 cm viewing distance and downsampling was deactivated. We used its lightness difference, lightness contrast, and lightness structure maps.

We included the color difference metric-CIEDE2000 [54] to see how a pixel-based metric will behave on 2.5D prints' QA. It works with the CIELAB color space.

The SCIELAB is selected because of its spatial filtering to mimic the HVS's spatial blurring [55]. We modified this metric so that it works with the CIEDE2000 formula. The viewing distance in SCIELAB was set to 51 cm. The white point was D65 in both CIEDE2000 and SCIELAB.

For all metrics, the input images were color images of 904 $\times$ 550 pixels. The color images were converted into grayscale following ITU recommendation BT.601-7 for SSIM and MS-SSIM. We used the default parameters for the selected IQMs.

4. Results and Discussion

We present readers with examples of what 2.5D print captures and their height maps look like in Figure 3. White corresponds to flat areas while black corresponds to maximum elevation in the height maps. The height and printer mode sets use one height map where the maximum height and the printer mode were varied during printing, respectively.

Due to the small size of the dataset and that there are only three (two for three images in the printer mode set) quality variations per image, it is less appropriate to look at the correlations between the IQMs and the subjective scores. Nevertheless, the IQMs can still be informative on quality aspects. Investigation of a single value of the IQM can be limiting, therefore an in-depth analysis of their quality maps can be valuable [56]. A quality map from a metric can be a good tool to find the location of errors for 2.5D printing applications. More specifically, the mean values of IQMs for different prints can be similar, but differences can be found in the quality maps.

We found that it is challenging to define a single or group of illumination angles from the MS captures that could be informative on 2.5D prints' quality aspects from the selected IQMs. In this light, we worked with mean quality maps of 60 different illumination angle captures. In other words, the IQMs were applied to each of the 60 illumination angle captures and the mean of 60 quality maps were taken per captured print. This can help to incorporate individual pixel values from all 60 angles. Furthermore, this is useful as the light source positions were symmetrically distributed and all effects, for example from shadows, also will symmetrically impact the mean quality map. Although the effects of the directional light may cancel out because of the symmetric distribution of the light positions, what will remain after can be interesting to analyze.

4.1. Are 2D IQMs Responsive to the Quality Variations in 2.5D Prints?

It is relevant first to check if the selected 2D IQMs are responsive to the quality variations introduced in the dataset of 2.5D prints. We will do this using the FR IQMs and calculating the difference between two prints, instead of between a print and the reference. If the IQMs are not responsive, they will yield no difference between the prints. For this test, we calculate the difference between the unnatural elevation - 1 mm maximum height - Brila mode (used as the test image) and surface roughness - 0.25 mm maximum height - Alto mode (used as the reference image) prints in each three images in naturalness - height - printer mode sets, respectively.

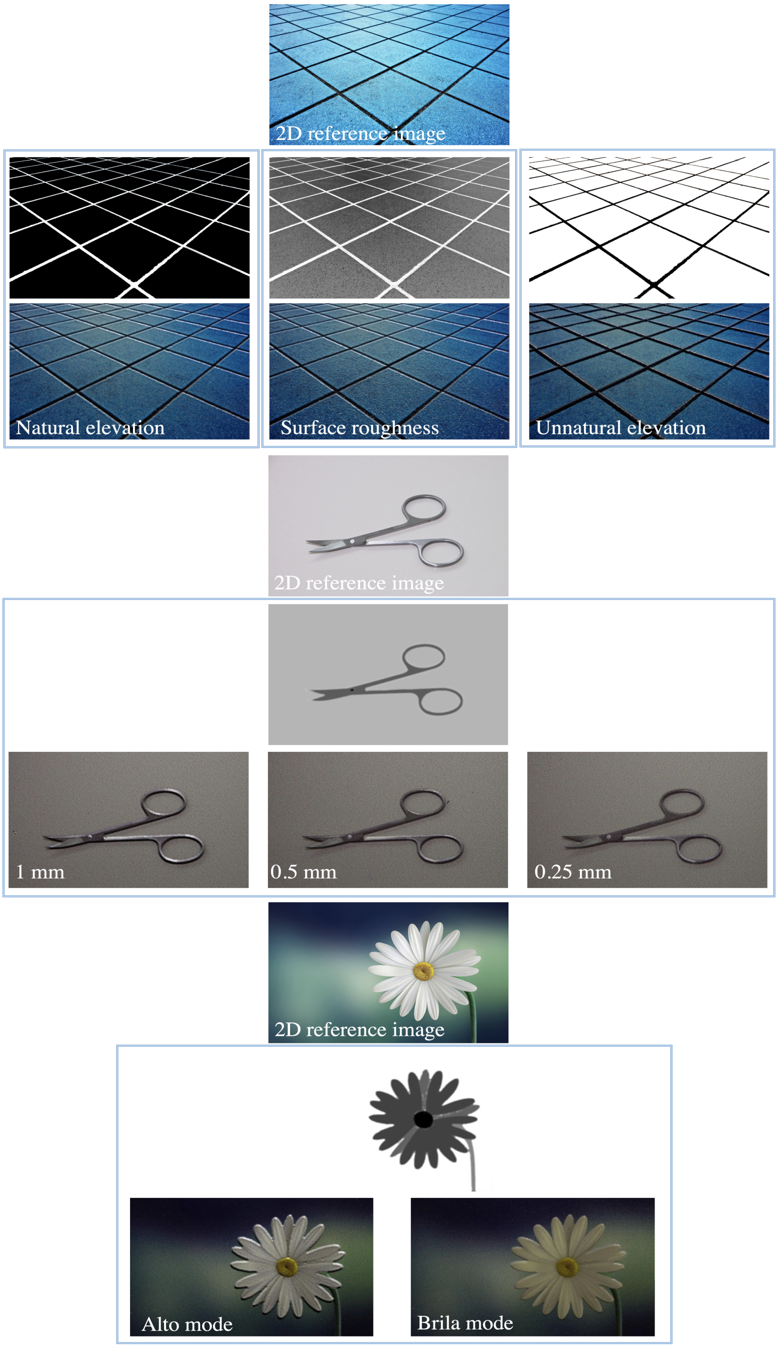

Figure 3. Illustration of 2D reference images (**top**), their height maps (**middle**), and their registered 2.5D print captures (**bottom**) from the MS as an example at elevation angle = 30° and azimuth angle = −20° from naturalness (**top**), height (**middle**), and printer mode (**bottom**) sets. In the height maps, white corresponds to flat areas while black corresponds to maximum elevation. All images and height maps were reproduced from the dataset of Kadyrova et al. [20].

4.1.1. Naturalness Set

The differences are expected to be in elevation and fine details from surface roughness. Shadows might be introduced due to elevation and thus more differences on the edges are expected, especially in the MS captures because it captures the prints under multiple angle directional illumination.

All selected IQMs are responsive to this set's images on the edges, on the background, and on the elevated parts' surfaces in both set-up captures. More specifically, the highest response is found on the edges. This shows that they are responsive to the differences between unnatural elevation and surface roughness prints (i.e., the significant differences between the two were in elevation and shadows induced from it and the secondary difference was in fine details). The differences in the prints are better captured by the IQMs in the MS captures than in the SS ones for the three images. Figure 4 illustrates the above-mentioned observations by iCID's lightness difference map on the tiles image where the difference comes from shadows that can appear on both sides of the grout line, which is better captured in the MS captures than in the SS ones. In addition, iCID's lightness structure map is responsive to both edges and fine details in the three images in this set in both set-up captures.

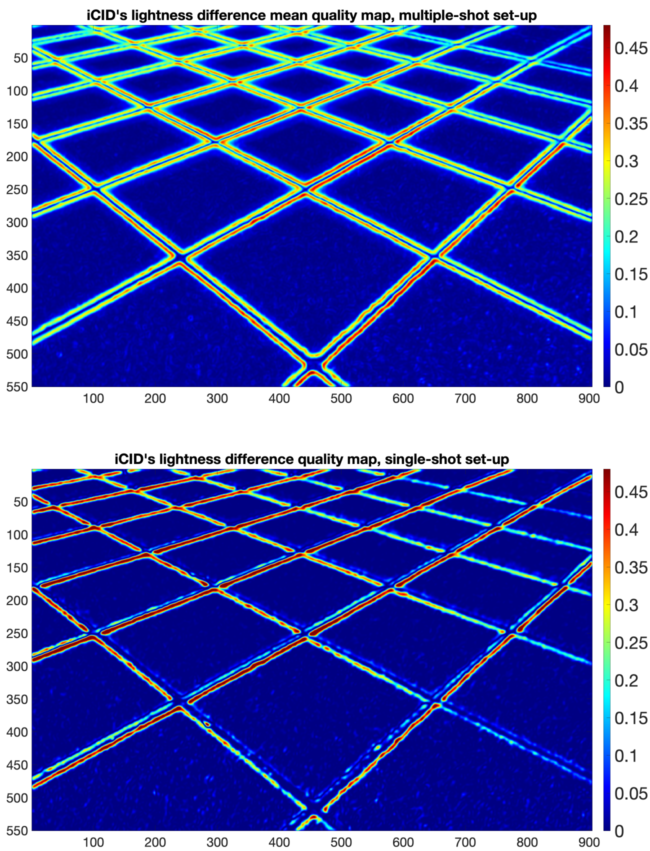

Figure 4. iCID's lightness difference (mean) quality maps for the tiles image with unnatural elevation print used as test image and surface roughness print used as reference image from the MS (**top**) and the SS (**bottom**) captures. It shows that iCID's lightness difference map is the most responsive (i.e., detects larger differences) on the edges in both set-up captures and that the responsiveness is better captured in the MS captures compared to the SS ones. Zero means there is no difference.

4.1.2. Height Set

The differences are expected to be in elevation and shadows because only height was changed between the prints in this set. The differences in the prints are better captured by the IQMs in the MS captures than in the SS ones for the three images, similar to the previous set. All selected IQMs are responsive to the differences (i.e., the highest response was on the edges compared to the background and the surface of the elevated parts) in both set-up captures. SSIM and MS-SSIM (both structure and contrast terms, scale 2) responded to differences in the background in the scissor image in the MS captures, although this area had not been elevated or changed between the prints. These metrics were also responsive to the edges and elevated parts' surfaces. However, this does not indicate that they are performing incorrectly in the scissor image. We assume that they are detecting halftone noise which is perceptually difficult to see at a normal viewing distance. This can explain the performance of SSIM and MS-SSIM (both terms, scale 2) on the scissor image.

Figure 5 shows histograms of the mean quality maps from the MS captures for the scissor image by SSIM and iCID's lightness difference map as a comparison. The histogram of SSIM is more spread than of iCID's lightness difference map because the latter contains Contrast Sensitivity Function (CSF) which filters the halftone noise. Thus, the halftone noise becomes less visible to some degree. SSIM and MS-SSIM do not contain the CSF and thus they are calculating differences that might be not perceptible at a certain distance. As a result, we observe that iCID's lightness difference map is fairly robust to halftone noise and a model of the HVS is useful to avoid the metric calculating differences that are outside the CSF threshold (i.e., invisible to the human eye). We assume that SSIM and MS-SSIM (both terms, scale 2) in combination with the MS can be used in a quality assurance application to detect differences present between two prints/images not necessarily related to the perceptual aspects.

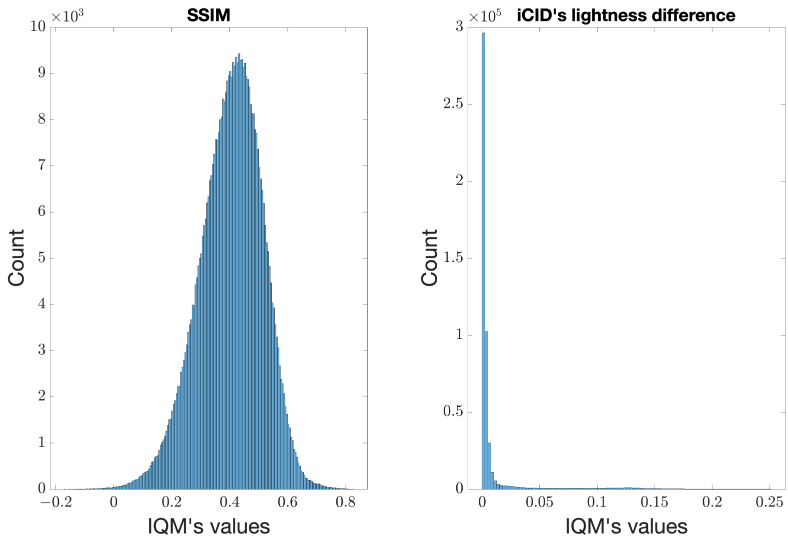

Figure 5. The histograms of SSIM and iCID's lightness difference mean quality maps for the scissor image with 1 mm maximum height print used as test image and 0.25 mm maximum height print used as reference image from the MS captures. The histograms show that iCID's lightness difference map detects more differences between prints around zero whereas SSIM detects more spread differences.

Another reason for the behavior of SSIM and MS-SSIM (both terms, scale 2) can be the directional dependent appearance (i.e., complex) of the 2.5D prints. For instance, the histograms of (mean) quality maps of the scissor image by SSIM in both MS and

SS captures show that the prints are noisy, but the prints are noisier in the MS captures (Figure 6).

Similar to SSIM and MS-SSIM (both terms, scale 2), iCID's lightness structure map also detected halftone noise in the scissor image from the MS captures between the prints (which was also responsive on the edges and elevated parts' surfaces). Its performance can be somewhat expected although it uses Gaussian weight distribution in its formula and the CSF was applied to the input images before calculating different maps in iCID. However, we used a lower pixels-per-degree value of a visual field, therefore simulating the HVS at a closer viewing distance when halftone noise becomes perceptible. Assigning a higher pixels-per-degree value of a visual field for the CSF filtering might improve iCID's performance [53]. Based on this, we found out that iCID is fairly robust to halftone noise at higher pixels-per-degree value (Figure 7).

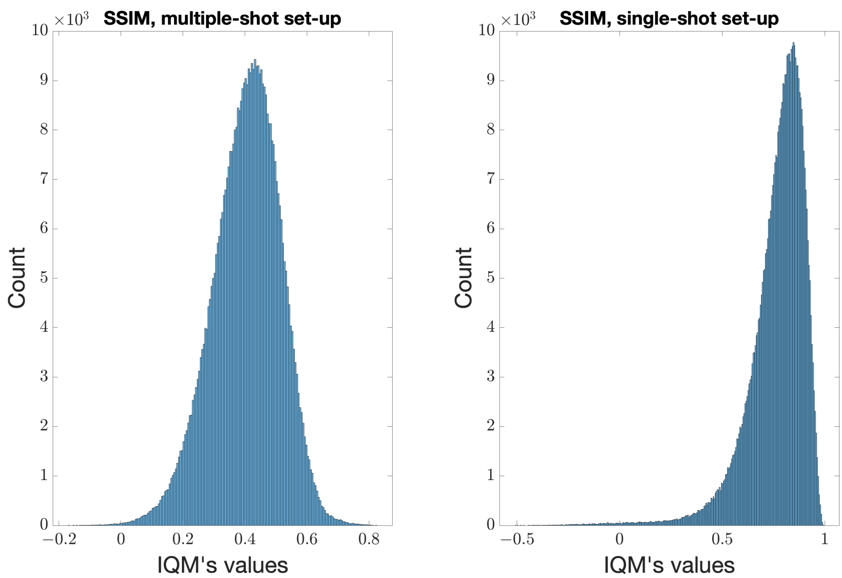

Figure 6. SSIM's (mean) quality map histograms for the scissor image with 1 mm maximum height print used as test image and 0.25 mm maximum height print used as reference image from the MS (**left**) and the SS (**right**) captures. In the MS captures, the prints are noisier than in the SS ones.

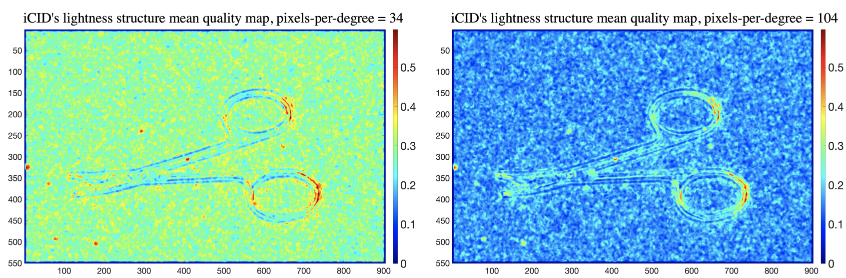

Figure 7. iCID's lightness structure mean quality maps for the scissor image with 1 mm maximum height print used as test image and 0.25 mm maximum height print used as reference image from the MS captures with lower (**left**) and higher (**right**) pixels-per-degree values. We see that halftone noise was reduced with higher pixels-per-degree value, where the quality values are closer to zero (meaning higher quality).

4.1.3. Printer Mode Set

The differences are expected to be in elevation and in shadows as well as in color. There is a height difference between Alto and Brila modes, where the former has a maximum height of 0.5 mm and the latter 0.25 mm. In addition, the Brila mode introduces a color shift. All selected IQMs have the highest response on the edges in both set-up captures. It is expected that the appearance is different between Alto and Brila modes because the elevation in the Alto mode is opaque while it is relatively transparent in the Brila mode. This difference is captured by the IQMs by being responsive on the edges. The IQMs also showed responsiveness to the differences on the surface of the elevated parts and on the background in both set-up captures. For instance, CIEDE2000 and SCIELAB have the highest response on the edges than on the surface of the elevated parts and on the background in flower and snowflake images. In the packaging image, their response is similar on the background and on the surface of the elevated part. This is because the Brila mode added a yellowish color on the elevated parts of the prints and the content of the packaging image has a somewhat similar color in the background and elevated part. Hence, CIEDE2000 and SCIELAB found similar color differences between Alto and Brila mode prints on both the background and the surface of the elevated part. SSIM, MS-SSIM (both terms, scale 2), and iCID's lightness structure map detected halftone noise in the three images in the MS captures and in the flower image in the SS captures along with being responsive on the edges, on the background, and on the elevated parts' surfaces. The detection of halftone noise by iCID's lightness structure map, SSIM, and MS-SSIM (both terms, scale 2) can be explained with the same reasoning given in the previous set for these metrics. The IQMs better captured differences in the MS captures than in the SS ones in this set.

To conclude, CIEDE2000, SCIELAB, iCID's lightness difference, lightness contrast, and light structure maps are found to be responsive to the differences between the prints in both set-up captures. The differences detected by these IQMs are more visible in the MS captures than in the SS ones. Based on this, the above-mentioned IQMs are considered in the next subsection. We keep iCID because it shows a capability to be used for perceptual QA of 2.5D prints as it is fairly robust to halftone noise arising from the printing process. We use iCID with an initial lower pixels-per-degree value for further analysis because the location of detected differences is the same regardless of the value of pixels-per-degree (Figure 7). There can be a slight change in the scale which is adequate because at a closer viewing distance we expect to see more differences than at longer distances.

4.2. Can We Obtain Insights on 2.5D Prints' Quality from 2D IQMs?

We present cases where the selected 2D IQMs and their quality maps are informative regarding relevant attributes between the three sets' images and their 2D reference images.

4.2.1. Naturalness Set

In both set-up captures, CIEDE2000, SCIELAB, and iCID's lightness difference map detected more differences on the elevated parts' surfaces in the prints with surface roughness while on the background in the prints with unnatural elevation in the three images. The same pattern was observed by iCID's lightness contrast and light structure maps in brick and wood images.

From the line profile taken from the (mean) quality map (the quality map was rotated to -65° to make the grout lines to be relatively vertical), CIEDE2000, SCIELAB, and iCID's lightness difference map show a similar pattern that there is less difference on the grout line than on the tiles and edges in the three prints from both set-up captures in the tiles image. To demonstrate this, we give an example of a line profile from the MS captures from SCIELAB in Figure 8.

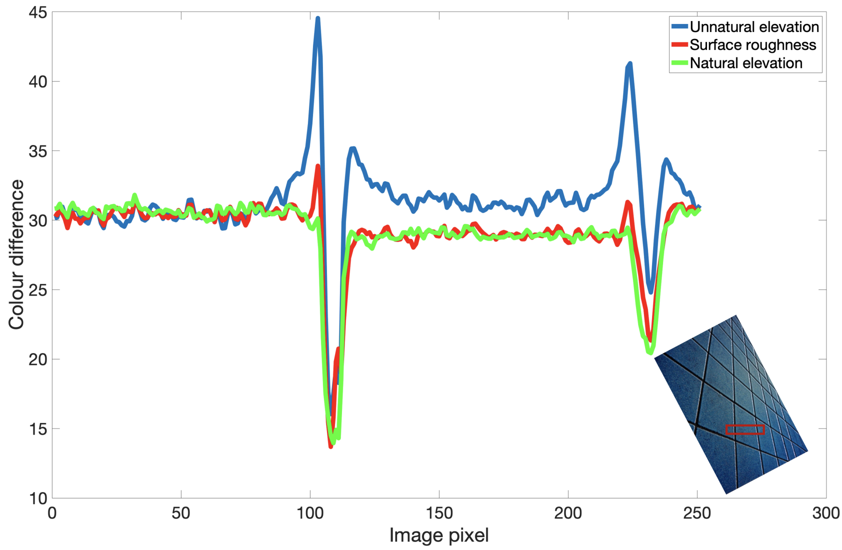

Figure 8. The line profile of the grout line (around 105 to 115 pixels) from the MS captures from SCIELAB. The image (rotated to −65°) on bottom right side shows approximate location, in red, of the extracted grout line in the tiles image. Zero means there is no color difference.

Based on Figure 8, we found the following observations: the shadow makes larger color differences on the left side of the grout line and this is because of the elevation made on the grout line in the print with unnatural elevation. There is also a larger color difference in the print with unnatural elevation on the right side of the grout line (around 115 to 130 pixels). This can be explained due to the elevation of the grout line in the print with unnatural elevation where light might hit and reflect back making the right side of the print slightly brighter (i.e., inter-reflection). In addition, there is some gloss on the elevated grout line and this might cause inter-reflections. As a result, we assume that CIEDE2000, SCIELAB, and iCID's lightness difference map are not able to detect whether the grout line is elevated or not, but they are able to detect shadows from the grout line.

iCID's lightness structure map found more differences in the surface roughness print than in the other two prints in the three images in the MS captures (Figure 9) while it found more differences in the unnatural elevation print than in the other two prints in the wood and tiles images in the SS captures. Moreover, SCIELAB and CIEDE2000 seem to be able to detect fine details from the surface roughness print of the wood image in both set-up captures (Figure 10).

In addition, more differences were detected by CIEDE2000 (tiles image), SCIELAB (tiles, wood images), iCID's lightness difference (tiles, wood images), and lightness structure (tiles image) maps in the three prints in the MS captures than in the SS ones.

4.2.2. Height Set

In both set-up captures, CIEDE2000, SCIELAB, and iCID's lightness structure map found more differences on the edges of the three prints in the three images. iCID's lightness contrast map also found more differences on the edges of the three prints in the scissor and speed sign images. More differences were found in the 1 mm maximum height print than in the other two prints in speed sign and running track images by CIEDE2000, SCIELAB, iCID's lightness difference, and lightness structure (Figure 11) maps in both set-up captures. This can be explained by the level of elevation.

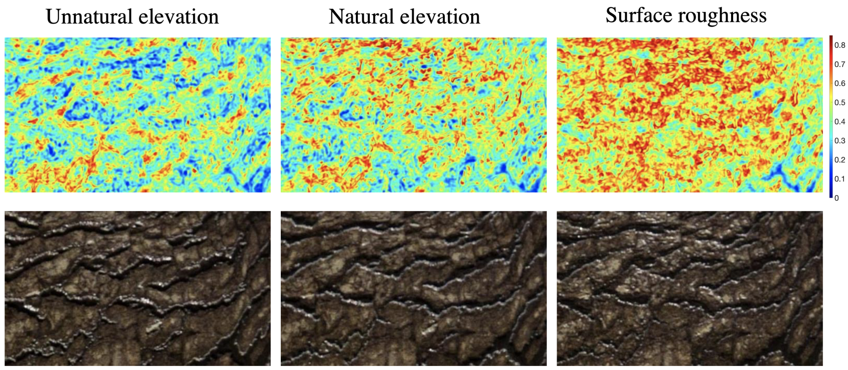

Figure 9. Example of more difference detection in print with surface roughness by iCID's lightness structure map in the MS captures in the wood image. Zero means there is no difference while one means that there is more difference. The images were cropped for illustration purpose. The mean quality maps are given on top while registered 2.5D print captures of the MS at elevation angle = 30° and azimuth angle = −20° are given on bottom.

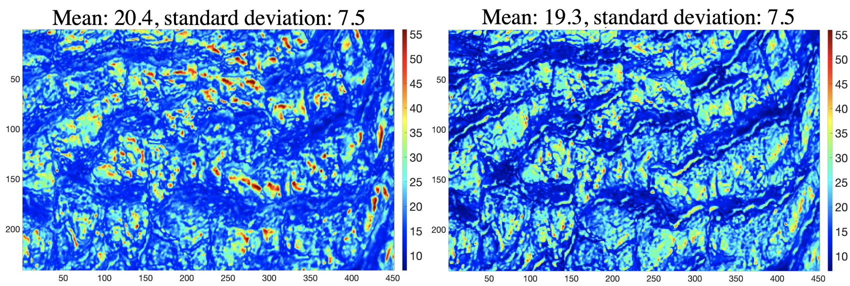

Figure 10. Example of fine details detection from surface roughness print of the wood image by SCIELAB in captures from the MS (**left**) and the SS (**right**). The images were cropped for illustration purpose. Zero means that there is no difference.

Additionally, more differences found by CIEDE2000, SCIELAB, iCID's lightness contrast, and lightness structure (running track, scissor images) maps in the three prints of the three images in the MS captures than in the SS ones. The differences were relatively similar in both set-up captures by iCID's lightness difference and lightness structure maps in the speed sign image.

4.2.3. Printer Mode Set

In both set-up captures, iCID's lightness difference map detected more differences on the surface of the elevated parts and on the background in both Alto and Brila mode prints in the three images. iCID's lightness contrast and lightness structure maps found more differences on the surface of the elevated parts in both prints in the flower and snowflake images. CIEDE2000 and SCIELAB detected more color differences on the edges in the Alto mode print and on the surface of the elevated part in the Brila mode print in the flower image (Figure 12). This is expected because the Brila mode introduces color shift. In addition, the Alto mode has higher elevation compared to the Brila mode. Therefore, more color differences are located on the edges in the Alto mode print. CIEDE2000, SCIELAB, and iCID's lightness structure map found more differences in the three images' both prints in the MS captures than in the SS ones.

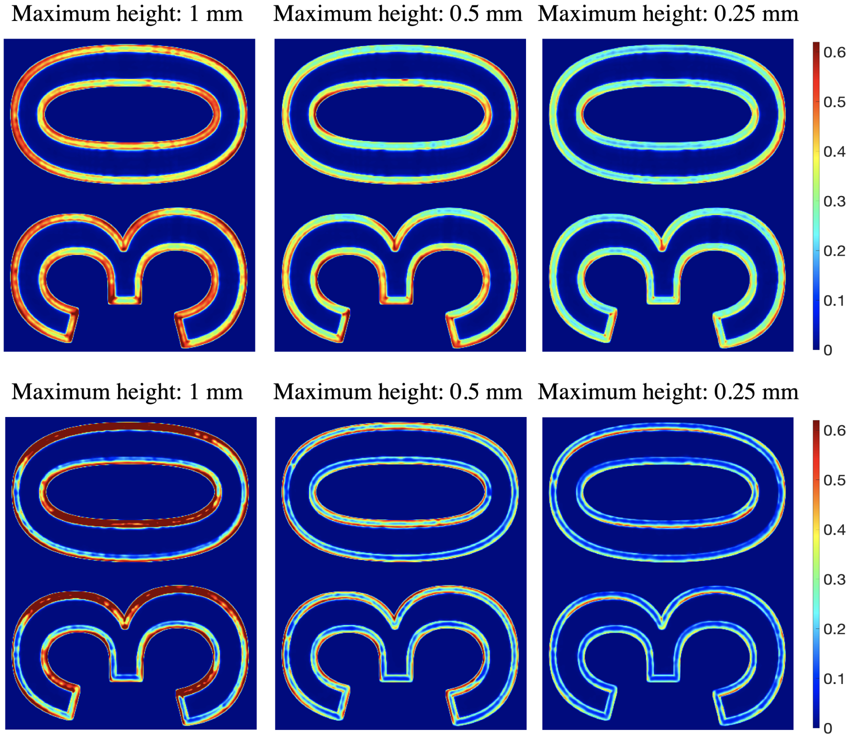

Figure 11. iCID's lightness structure (mean) quality maps from the MS (**top**) and the SS (**bottom**) captures. It detected more differences on the edges in the 1 mm maximum height print compared to other two prints in both set-up captures. The maps were cropped for illustration purpose. Zero means there is no difference.

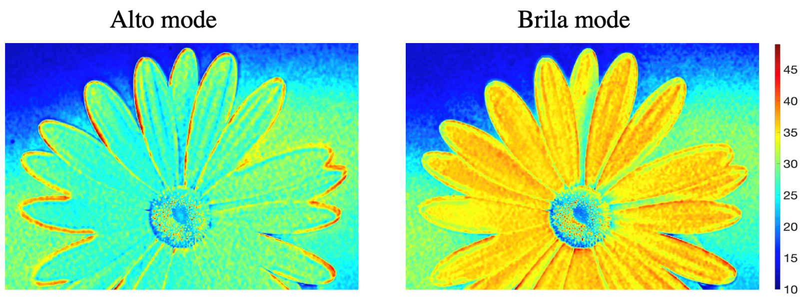

Figure 12. SCIELAB's mean quality maps of the flower image for the Alto mode (**left**) and the Brila mode (**right**) prints from the MS captures. The maps were cropped for illustration purpose. Zero means there is no difference.

To conclude, the quality maps of the tested IQMs are informative regarding 2.5D prints' QA. More specifically, they are informative in a similar way regarding difference detection on areas such as edges, elevated parts' surfaces, and background. For future work, it will be interesting to investigate quality maps along with height maps.

All differences detected by the tested IQMs (mainly on the edges, on the surface of the elevated parts, and on the background) for the examined set of images are consistent with the observers' feedback (mainly on elevation and aspects introduced by elevation (e.g., shadow)). This is better visible in the MS captures. As a result, this shows that the mentioned IQMs can be used for QA of 2.5D prints for specific attributes/features with a specific capture set-up.

4.3. Which Set-Up Is More Appropriate for 2.5D Prints Capture Based on IQMs' Performance?

We compare the correlation between (mean) quality maps of the tested IQMs from the MS and the SS captures to find which tested IQMs have more differences between the two set-ups and/or are more imperceptive to the set-up. For example, Figure 13 shows a boxplot of the correlation between the MS and the SS of height set images (i.e., for nine prints). Although the tested IQMs' median correlations between the two set-ups are relatively high (roughly around 0.7 for the given IQMs), the ranges of the correlation values are spread. In other words, the correlations of some IQMs are more variable than others. In particular, iCID's lightness contrast map has more prints which have lower correlations than the median value between the two set-ups. The tested IQMs seem to be more towards being imperceptive to the set-up in this set of images, meaning that they detect (more) differences (although the scale might vary between the set-ups) on the same areas irrespective of the set-up. This is in line with our observations in the previous subsections.

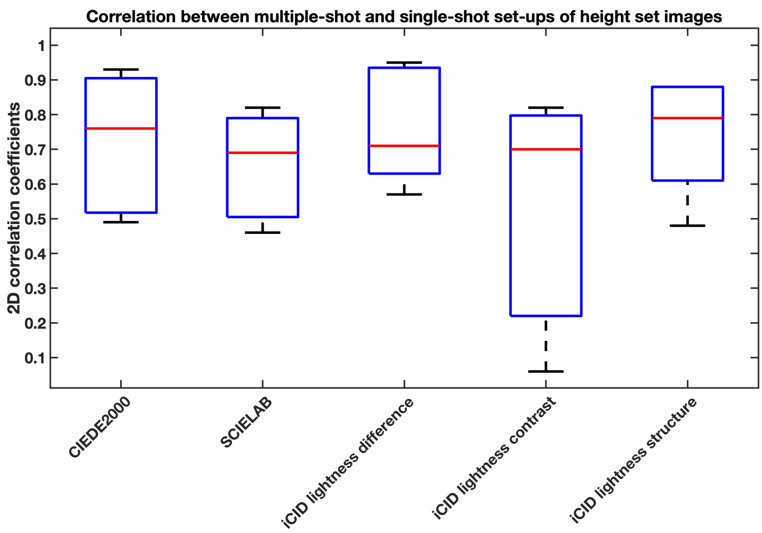

Figure 13. The boxplot of correlation between two set-ups of height set images by the tested IQMs. iCID's lightness contrast map has a very spread range among the IQMs.

The impact of elevation on the appearance can be better revealed in the MS captures. For example, Figures 4 and 11 show that the differences are better detected by the IQMs in the MS captures compared to the SS ones, where the latter depends on the illumination direction while the former can provide better error capture due to the mean of several illumination angles. In other words, the SS's light source is on top inside the light booth cabinet and shadowing will be more on one side than in the other side regardless of how diffuse the light is inside the light booth. This is a clear limitation of the SS with imperfect diffuse illumination inside the light booth. In contrast, the MS can provide diffuse illumination distributed evenly via capturing prints with a directional set of illumination angles (even generating mean quality maps from set of illumination angles can provide the similar effect). More information on the selected IQMs' ability to better detect differences in the MS captures than in the SS ones can be found in Section 4.1. Moreover, the tested IQMs can find more differences mostly in the MS captures in comparison to captures from the SS (refer to Section 4.2). SSIM and MS-SSIM (both terms, scale 2) detected differences that are perceptually difficult to see mostly in the MS captures as opposed to the SS ones.

Thus, they can in combination with the MS be useful in applications which need to detect such differences.

To summarize, both set-ups can be useful for 2.5D prints' QA. Nevertheless, there is a dominance by the MS based on our observations in terms of the tested IQMs' ability to capture more differences and that the captured differences are clearly visible by the examined IQMs' mean quality maps in the MS captures. Moreover, the MS (i.e., RTI) is relatively simple which needs a camera, a sample of interest, and a light source (which is movable to illuminate the sample from different illumination angles) [43].

4.4. Limitations

Because our primary goal is to focus on the IQMs and the QA side, there could be limitations on the data capture side in terms of camera calibration and accurate color acquisition. We have color checker captures with the 2.5D prints. Thus, there is an option to estimate the transformation from them to the 2D reference images and work with the transformed images, which can be considered for future work. Another limitation might be due to non-use of camera raw images. The raw images store unprocessed sensor data and are mostly used in science [57]. Thus, this can be considered for future work as well. Moreover, we worked with a small set of images, which limits quantitative analysis. It would be better to have other attributes included apart from the attributes that have been changed in the used dataset as well as other types of content that could be useful to generalize findings.

5. Conclusions and Future Works

The QA of 2.5D prints is currently attracting the attention of researchers, but objective QA of 2.5D prints is less studied. Therefore, an attempt to test existing 2D IQMs to predict the quality of 2.5D prints captured by the multiple-shot and single-shot set-ups was made. We acquired the following observations:

- iCID's lightness difference, lightness contrast, and lightness structure maps, CIEDE2000, and SCIELAB can find differences between 2.5D prints as well as between 2.5D prints and their 2D reference images in both set-up captures;
- More differences are detected mostly in the multiple-shot set-up captures than in the single-shot set-up ones and the captured differences are clearer visible in the multiple-shot set-up captures by iCID's lightness difference, lightness contrast, and lightness structure maps, CIEDE2000, and SCIELAB based on their (mean) quality maps;
- To create a metric for 2.5D prints' perceptual QA, it important to have a model of the HVS.

In conclusion, iCID's lightness difference, lightness contrast, and lightness structure maps, CIEDE2000, and SCIELAB were found to be relevant to use to detect differences on the edges, on the surface of the elevated parts, and on the background between 2.5D prints and their 2D reference images as well as between 2.5D prints, especially in the multiple-shot set-up captures. Our results on the responsiveness of the selected 2D IQMs on the quality variations of 2.5D prints can be useful in, for example a quality assurance application, where the industry is interested to detect defects or any other differences between prints (or with respect to the reference image/print) without necessarily defining which prints' quality is the best. We consider testing more IQMs and to apply the CSF in the preprocessing stage before applying the IQMs to account for the distance from the prints to the observer as a future work.

Author Contributions: Conceptualization, A.K., V.K., M.P.; methodology, A.K., V.K., M.P.; data acquisition and processing, A.K., V.K.; analysis, A.K., V.K., M.P.; writing—original draft, A.K.; writing—review and editing, A.K., V.K., M.P.; visualization, A.K.; supervision, M.P., V.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the ApPEARS project from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 814158.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset used in this work can be found at https://zenodo.org/record/4609031#.YRTiMt8RVPY, accessed on 15 March 2021.

Acknowledgments: We would like to thank Stephen Westland and Clemens Weijkamp for fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyzes, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Pedersen, M.; Amirshahi, S.A. Framework for the evaluation of color prints using image quality metrics. In Proceedings of the Conference on Colour in Graphics, Imaging, and Vision. Society for Imaging Science and Technology, Joensuu, Finland, 14–17 June 2010; Volume 2010, pp. 75–82.
2. Pedersen, M.; Zheng, Y.; Hardeberg, J.Y. Evaluation of image quality metrics for color prints. In Proceedings of the Scandinavian Conference on Image Analysis, Ystad, Sweden, 1 May 2011; pp. 317–326.
3. Liu, X.; Chen, L.; Ortiz-Segovia, M.V.; Ferwerda, J.; Allebach, J. Characterization of relief printing. In Proceedings of the Measuring, Modeling, and Reproducing Material Appearance. International Society for Optics and Photonics, San Francisco, CA, USA, 10 March 2014; Volume 9018, p. 90180P.
4. Baar, T.; Samadzadegan, S.; Brettel, H.; Urban, P.; Segovia, M.V.O. Printing gloss effects in a 2.5 D system. In Proceedings of the Measuring, Modeling, and Reproducing Material Appearance. International Society for Optics and Photonics, San Francisco, CA, USA, 24 February 2014; Volume 9018, p. 90180M.
5. O'Dowd, P.; Parraman, C.; Harding, M. Vector driven 2.5 D printing with non-photorealistic rendering. In Proceedings of the Electronic Imaging, San Francisco, CA, USA, 14–18 February 2016; pp. 1–6.
6. Wang, J.; Allebach, J.P.; Ortiz-Segovia, M.V. Semantically-based 2.5 D texture printing. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 2634–2638.
7. Jacobson, R.E. An evaluation of image quality metrics. *J. Photogr. Sci.* **1995**, *43*, 7–16. [CrossRef]
8. Engeldrum, P.G. *Psychometric Scaling: A Toolkit for Imaging Systems Development*; Imcotek Press: Winchester, MA, USA, 2000.
9. Fairchild, M. Image Quality Measurement and Modeling for Digital Photography. In Proceedings of the International Congress on Imaging Science, Tokyo, Japan, 13–17 May 2002; pp. 318–319.
10. Janssen, T.J.W.M. Computational Image Quality. Ph.D. Thesis, Technische Universiteit Eindhoven, Eindhoven, The Netherlands, 1999.
11. Yendrikhovskij, S.N. Color Reproduction and the Naturalness Constraint. Ph.D. Thesis, Technische Universiteit Eindhoven, Eindhoven, The Netherlands, 1998.
12. ISO. *ISO 12647: Graphic Technology-Process Control for the Manufacture of Half-Tone Colour Separations, Proof and Production Prints*; ISO: Geneva, Switzerland, 2013.
13. Field, G.G. Test image design guidelines for color quality evaluation. In Proceedings of the Color and Imaging Conference, Society for Imaging Science and Technology, Scottsdale, AZ, USA, 16–19 November 1999; Volume 1999, pp. 194–196.
14. Southworth, M.; Southworth, D. *Quality and Productivity in the Graphic Arts: How to Improve Quality, Productivity, and Profit Utilizing a Continuous Quality Improvement Program and Statistical Process Control (SPC)*; Graphic Arts Publishing Company: Portland, OR, USA, 1989.
15. Keelan, B. *Handbook of Image Quality: Characterization and Prediction*; CRC Press: Boca Raton, FL, USA, 2002.
16. ISO. *ISO 20462-1 Photography-Psychophysical Experimental Methods to Estimate Image Quality-Part 1: Overview of Psychophysical Elements*; ISO: Geneva, Switzerland, 2004.
17. Liu, X. Feature Extraction and Image Analysis with the Applications to Print Quality Assessment, Streak Detection, and Pedestrian Detection. Ph.D. Thesis, Purdue University Graduate School, West Lafayette, IN, USA, 2019.
18. Pedersen, M.; Bonnier, N.; Hardeberg, J.Y.; Albregtsen, F. Attributes of image quality for color prints. *J. Electron. Imaging* **2010**, *19*, 011016.
19. Pedersen, M.; Bonnier, N.; Hardeberg, J.Y.; Albregtsen, F. Validation of quality attributes for evaluation of color prints. In Proceedings of the Color and Imaging Conference. Society for Imaging Science and Technology, San Antonio, TX, USA, 8–12 November 2010; Volume 2010, pp. 74–79.
20. Kadyrova, A.; Kitanovski, V.; Pedersen, M. A study on attributes for 2.5 D print quality assessment. In Proceedings of the Color and Imaging Conference. Society for Imaging Science and Technology, Online, 4–19 November 2020; Volume 2020, pp. 19–24.

21. Samadzadegan, S.; Blahová, J.; Urban, P. Color-Printed Gloss: Relating measurements to perception. In Proceedings of the Color and Imaging Conference. Society for Imaging Science and Technology, Boston, MA, USA, 3–7 November 2014; Volume 2014, pp. 207–211.
22. Fu, Y.; Wang, S. A no reference image quality assessment metric based on visual perception. *Algorithms Multidiscip. Digit. Publ. Inst.* **2016**, *9*, 87. [CrossRef]
23. Liu, H.; Heynderickx, I. Visual attention in objective image quality assessment: Based on eye-tracking data. In *IEEE Transactions on Circuits and Systems for Video Technology*; IEEE: Minneapolis, MN, USA, 2011; Volume 21, pp. 971–982.
24. Wang, B.; Wang, Z.; Liao, Y.; Lin, X. HVS-based structural similarity for image quality assessment. In Proceedings of the 2008 9th International Conference on Signal Processing, Beijing, China, 26–29 October 2008; pp. 1194–1197.
25. Babu, R.V.; Perkis, A. An HVS-based no-reference perceptual quality assessment of JPEG coded images using neural networks. In Proceedings of the IEEE International Conference on Image Processing 2005, Genoa, Italy, 11–14 September 2005; Volume 1, p. I-433.
26. Zhang, Y.; Wu, L.; Wang, S.; Wei, G. Color image enhancement based on HVS and PCNN. *Sci. China Inf. Sci.* **2010**, *53*, 1963–1976. [CrossRef]
27. Nadenau, M.J.; Winkler, S.; Alleysson, D.; Kunt, M. Human vision models for perceptually optimized image processing—A review. *Proc. IEEE* **2000**, 0243191.
28. Akyazi, P.; Ebrahimi, T. An improved objective metric to predict image quality using deep neural networks. In Proceedings of the Electronic Imaging, Society for Imaging Science and Technology, Burlingame, CA, USA, 13–17 January 2019; Volume 2019, p. 21-1.
29. Amirshahi, S.A.; Pedersen, M.; Yu, S.X. Image quality assessment by comparing CNN features between images. *J. Imaging Sci. Technol. Soc. Imaging Sci. Technol.* **2016**, *60*, 60410. [CrossRef]
30. Hou, W.; Gao, X.; Tao, D.; Li, X. Blind image quality assessment via deep learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *26*, 1275–1286. [PubMed]
31. Bianco, S.; Celona, L.; Napoletano, P.; Schettini, R. On the use of deep learning for blind image quality assessment. *Signal Image Video Process.* **2018**, *12*, 355–362. [CrossRef]
32. Chetouani, A.; Pedersen, M. Image Quality Assessment without Reference by Combining Deep Learning-Based Features and Viewing Distance. *Appl. Sci. Multidiscip. Digit. Publ. Inst.* **2021**, *11*, 4661.
33. Zhang, X.; Silverstein, D.A.; Farrell, J.E.; Wandell, B.A. Color image quality metric S-CIELAB and its application on halftone texture visibility. In Proceedings of the IEEE COMPCON 97. Digest of Papers, San Jose, CA, USA, 23–26 February 1997; pp. 44–48.
34. Yanfang, X.; Yu, W.; Ming, Z. Color reproduction quality metric on printing images based on the s-cielab model. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Wuhan, China, 12–14 December 2008; Volume 6, pp. 294–297.
35. Eerola, T.; Kämäräinen, J.K.; Lensu, L.; Kälviäinen, H. Framework for applying full reference digital image quality measures to printed images. In Proceedings of the Scandinavian Conference on Image Analysis, Oslo, Norway, 15–18 June 2009; pp. 99–108.
36. Baar, T.; Brettel, H.; Segovia, M.V.O. A survey of 3D image quality metrics for relief print evaluation. In Proceedings of the 2013 Colour and Visual Computing Symposium (CVCS), Gjøvik, Norway, 5–6 September 2013; pp. 1–6.
37. Grice, J.; Allebach, J.P. The print quality toolkit: An integrated print quality assessment tool. *J. Imaging Sci. Technol. Soc. Imaging Sci. Technol.* **1999**, *43*, 187–199.
38. Zhao, P.; Pedersen, M.; Thomas, J.B.; Hardeberg, J.Y. Perceptual spatial uniformity assessment of projection displays with a calibrated camera. In Proceedings of the Color and Imaging Conference. Society for Imaging Science and Technology, Boston, MA, USA, 3–7 November 2014; Volume 2014, pp. 159–164.
39. Zhao, P.; Pedersen, M.; Hardeberg, J.Y.; Thomas, J.B. Camera-based measurement of relative image contrast in projection displays. In Proceedings of the European Workshop on Visual Information Processing (EUVIP), Paris, France, 10–12 June 2013; pp. 112–117.
40. Zaman, T.; Jonker, P.; Lenseigne, B.; Dik, J. Simultaneous capture of the color and topography of paintings using fringe encoded stereo vision. *Herit. Sci.* **2014**, *2*, 1–10. [CrossRef]
41. Elkhuizen, W.S.; Lenseigne, B.A.; Baar, T.; Verhofstad, W.; Tempelman, E.; Geraedts, J.M.; Dik, J. Reproducing oil paint gloss in print for the purpose of creating reproductions of Old Masters. In Proceedings of the Measuring, Modeling, and Reproducing Material Appearance 2015. International Society for Optics and Photonics, San Francisco, CA, USA, 14 April 2015; Volume 9398, p. 93980W.
42. Elkhuizen, W.S.; Callewaert, T.W.; Leonhardt, E.; Vandivere, A.; Song, Y.; Pont, S.C.; Geraedts, J.M.; Dik, J. Comparison of three 3D scanning techniques for paintings, as applied to Vermeer's 'Girl with a Pearl Earring'. *Herit. Sci.* **2019**, *7*, 1–22. [CrossRef]
43. Reflectance Transformation Imaging (RTI). Available online: http://culturalheritageimaging.org/Technologies/RTI/ (accessed on 2 July 2021).
44. Parraman, C.; Segovia, M.V.O. *2.5 D Printing: Bridging the Gap between 2D and 3D Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2018.
45. Parraman, C.; Klein, S.; Aure, X. *RTI for RGB: Photographing Mica Based Inks on Black Paper Using Reflectance Transform Imaging*; Repository of University of the West of England: Bristol, UK, 2019.
46. Pintus, R.; Dulecha, T.; Jaspe Villanueva, A.; Giachetti, A.; Ciortan, I.; Gobbetti, E. Objective and Subjective Evaluation of Virtual Relighting from Reflectance Transformation Imaging Data. In Proceedings of the 16th EUROGRAPHICS Workshop on Graphics and Cultural Heritage (EG GCH), Vienna, Austria, 12–15 November 2018.

47. Kitanovski, V.; Hardeberg, J.Y. Objective evaluation of relighting models on translucent materials from multispectral RTI images. In Proceedings of the IS&T International Symposium on Electronic Imaging Science and Technology, Society for Imaging Science and Technology, Online, 11–25 January 2021.
48. Appearance Printing European Advanced Research School. Available online: www.appears-itn.eu (accessed on 12 November 2020).
49. Eerola, T.; Lensu, L.; Kalviainen, H.; Kamarainen, J.K.; Leisti, T.; Nyman, G.; Halonen, R.; Oittinen, P. Full reference printed image quality: Measurement framework and statistical evaluation. *J. Imaging Sci. Technol. Soc. Imaging Sci. Technol.* **2010**, *54*, 10201-1. [CrossRef]
50. Okarma, K.; Fastowicz, J.; Lech, P.; Lukin, V. Quality Assessment of 3D Printed Surfaces Using Combined Metrics Based on Mutual Structural Similarity Approach Correlated with Subjective Aesthetic Evaluation. *Appl. Sci. Multidiscip. Digit. Publ. Inst.* **2020**, *10*, 6248. [CrossRef]
51. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
52. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 9–12 November 2003; Volume 2, pp. 1398–1402.
53. Preiss, J.; Fernandes, F.; Urban, P. Color-image quality assessment: From prediction to optimization. *IEEE Trans. Image Process.* **2014**, *23*, 1366–1378. [CrossRef] [PubMed]
54. Sharma, G.; Wu, W.; Dalal, E.N. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. In *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*; Wiley Online Library: Hoboken, NJ, USA, 2005; Volume 30, pp. 21–30.
55. Zhang, X.; Wandell, B.A. A spatial extension of CIELAB for digital color image reproduction. In *SID International Symposium Digest of Technical Papers*; Citeseer: Princeton, NJ, USA, 1996; Volume 27, pp. 731–734.
56. Pedersen, M. Image Quality Metrics for the Evaluation of Printing Workflows. Ph.D. Thesis, University of Oslo, Oslo, Norway, 2011.
57. Burggraaff, O.; Schmidt, N.; Zamorano, J.; Pauly, K.; Pascual, S.; Tapia, C.; Spyrakos, E.; Snik, F. Standardized spectral and radiometric calibration of consumer cameras. *Opt. Express Opt. Soc. Am.* **2019**, *27*, 19075–19101. [CrossRef] [PubMed]