https://doi.org/10.2352/ISSN.2470-1173.2017.8.MAAP-285

Towards a Consistent, Tool Independent Virtual Material Appearance

Darýa Guarnera [a]***, Giuseppe Claudio Guarnera*** [a,b]***, Cornelia Denk*** [c]***, and Mashhuda Glencross*** [d]
[a] ***Loughborough University, Loughborough, United Kingdom***
[b] ***NTNU - Norwegian University of Science and Technology, Gjøvik, Norway***
[c] ***BMW Research, Munich, Germany***
[d] ***Pismo Software, Manchester, United Kingdom***

Abstract

Current materials appearance is mainly tool dependent and requires time, labour and computational cost to deliver consistent visual result. Within the industry, the development of a project is often based on a virtual model, which is usually developed by means of a collaboration among several departments, which exchange data. Unfortunately, a virtual material in most cases does not appear the same as the original once imported in a different renderer due to different algorithms and settings. The aim of this research is to provide artists with a general solution, applicable regardless the file format and the software used, thus allowing them to uniform the output of the renderer they use with a reference application, arbitrarily selected within an industry, to which all the renderings obtained with other software will be made visually uniform. We propose to characterize the appearance of several classes of materials rendered using the arbitrary reference software by extracting relevant visual characteristics. By repeating the same process for any other renderer we are able to derive ad-hoc mapping functions between the two renderers. Our approach allows us to hallucinate the appearance of a scene, depicting mainly the selected classes of materials, under the reference software.

Introduction

Current materials appearance is mainly tool dependent and requires time, labour and computational cost to deliver consistent visual result. Visual appearance of a material is important in a range of different areas, including Visual Special effects, Interior/Exterior Modelling, Architectural Modelling, Cultural Heritage, Computer Games and Automotive Design. In many industries, final decisions about products are often based on a virtual prototype, which drives the demand for more realistic and tool independent virtual material.

Current material representations are dependent on the software where virtual material is designed, optimised and rendered. A virtual material in most cases does not appear the same as the original once imported in a different renderer due to different algorithms and settings. In fact, each renderer has a specific set of settings and limitations to interpret the properties of a given material data. Hence, a standard material model should be consistent and tool independent, to allow coherent material representation.

Although a broad range of solutions for material representation exists, there is still no straightforward and clear path to follow [1]. The lack of a standardized material representation method has prevented the use of measured data, often stored in some format not directly usable in 3D packages (e.g. raw uncompressed data organized in 4-D tables [2, 3]) and needs to be encoded and compressed in such way that it is readable from a specific renderer, leaving the choice of the file format to the end user of the data.


(a) Rendered Image (3ds Max)


(b) Hallucinated Image (Maya)

Figure 1. *An input-output pair from our tool. From the input image (on the top, rendered using 3D Max) we estimated the appearance of the same scene rendered using Autodesk Maya (on the bottom). The output image does not display any visual artifact and looks plausible.*

A recent development in this area provides a new file formats for BRDF representation, proposed by NVIDIA, is the Material Definition Language (MDL) [4]. MDL is a procedural programming language that allows to define properties of physically plausible material models and integrate them into any supported application. However, it requires some programming skills which may not be part of a typical 3D artist curriculum and it is not yet widespread among them. Furthermore, previously rendered scenes cannot benefit from it in order to achieve consistency.

Based on these observations, the aim of this research is to provide artists with a general solution, applicable regardless the file format and the software used, thus allowing them to uniform the output of the renderer they use with a "golden standard" application, arbitrarily selected. In this way, within an industry it would be possible to select a reference "golden standard", to which all the renderings obtained with other software will be made visually consistent though our method (see Figure 1 for an example).

Our contributions include the following:

- A dataset of reference scenes rendered with different software, carefully selecting the most closely matching settings. Each scene depicts spheres with known materials applied on them, used to characterize the output of each selected renderer with meaningful statistics, per color channel.
- An algorithm aimed to standardize the material appearance which makes use the aforementioned statistics to hallucinate the look of the relevant materials in the scene, providing the user with a post-processed rendering which matches the characteristics of the golden standard renderer, on a per material basis, in a transparent way. Our approach takes in input a rendered image, a label indicating the renderer used to generate the input image and a label specifying the renderer to match in output, and a normal map of the scene.

In this paper we sketch our system and we provide some preliminary results, showing the potential of the proposed approach.

Related Work

The Bidirectional Reflectance Distribution Function [5] (BRDF) is a radiometric function which describes how incident energy is redirected in all directions across a hemisphere above the surface. The BRDF is a simplified reflectance representation for opaque surfaces: it assumes that light entering a material leaves the material at the same position, whereas more complex reflectance functions can describe light transport between any two incident rays on a surface, *e.g.* the Bidirectional Scattering-Surface Reflectance-Distribution Function (BSSRDF). Given the incident direction $\mathbf{v_i} = (\theta_i, \phi_i)$ and the outgoing direction $\mathbf{v_o} = (\theta_o, \phi_o)$ in spherical coordinates, the BRDF f_r, measured in inverse steradian $[1/sr]$, is defined as the ratio of outgoing radiance to incoming irradiance:

$$f_r(\mathbf{v_i}, \mathbf{v_r}) = \frac{dL_r(\mathbf{v_r})}{L_i(\mathbf{v_i}) \cos\theta_i d\omega_i} \tag{1}$$

where L_i is incident radiance and L_r is the reflected radiance. A BRDF should ensure some basic physical properties, namely non-negativity, reciprocity and energy conservation.

Zubiaga *et al.* [6] focused on how the properties of BRDFs influence the rendered picture, by working locally in Fourier space and analysing how BRDF moments up to order 2 induce colouring, warping and blurring of reflected radiance. In their work a subset of unimodal materials in the MERL database [2] has been used, limiting the analysis to 2D slices of the selected BRDFs, assumed to act as filters on the incident lighting, whose parameters need to be estimated. Although a BRDF is a 4D function, the choice of 2D slices of the BRDF is justified by the consideration that when the radiance reaching the eye from a surface point is computed, the view direction is kept fixed.

The idea that a material acts as a filter in the image has been successfully exploited by Zubiaga *et al.*[7] in the context of MatCaps, images of spheres in orthographic projection in which lighting and material properties are baked in; MatCaps have proved to be a tool to design plausible material appearance included in many physically-based renderers, although they do not allow easy manipulation of the material. In [7] a MatCap is decomposed into high and low frequency components, unwarped into a spherical representation thanks to the filter parameters interactively estimated. The described representation allows dynamic appearance manipulation of lighting and material, thus overcoming the typical limitation of MatCaps, able to describe only static appearance. Different MatCaps can be used on different object parts by giving in input a map of the materials IDs.

Khan *et. al.* [8] leveraged some limitations of the human vision, which proves to be tolerant to some physical inaccuracies, in order to develop an image-based material editing tool. It requires in input a high-dynamic range photograph of an object, together with an alpha matte to separate it from the background, to produce in output a new photograph of the same object, in which its material is replaced with an entirely new one. A depth map is estimated from the pixels belonging to the object itself though an approximated shape-from-shading approach; the gradient of the depth map are then used to compute a surface normal **n** for each pixel. The object is removed from the image thanks to the alpha matte, and the missing pixels are inpainted by preserving the statistical properties of the remainder of the environment. A HDR environment map is then created by cutting a circle from the middle of the image, then placed in the image plane and extruded to become a hemispherical environment map. The estimated information are then used for a range of transformations ranging from applying a texture to the object to the application of an arbitrary BRDF, or even the simulation of transparency and translucency.

Both [7, 8] require either a map with material IDs or a foreground/background map, which could be challenging to create manually for a complex scene. Luckily, computer vision techniques have recently become available to help in this task. In fact, in the last few years, several dataset of texture images, including man-made materials, have been released. The Material in Context Database (MINC) is large dataset containg about 3 milions material samples, used to train a convolutional neural network (CNN) and a conditional random field (CRF), combined together to recognize and segment materials in the wild [9]. In particular, a CNN is trained in order to produce a single prediction for a given input patch. The trained CNN is used as a sliding window to predict material across the image, at different scales. The prediction across the scales are then averaged and given in input to the CRF, in which all pairs of variables are directly connected by pairwise potentials; the fully connected pairwise reasoning outputs material predictions for each pixel.

Problem Definition

Material modelling generally involves a great deal of manual effort, ranging from a completely manual creation of a material to a fully automated acquired material which often cannot be used directly in rendering. The broad range of material models and complexity of the parameters requires from an artist an understanding of the underlying representation and material's micro/macrostructure; moreover current photo-realistic rendering systems use BRDFs to varying levels of accuracy, leading to a very different appearance in the final rendering.

In this paper we adopt a conceptually similar approach as in [6, 7], however we make a different hypothesis: we consider a rendering system to act as a set of filters, on a per material basis, on the rendered image. Hence, if the same scene is rendered using two different tools using the same settings, the visual differences can be attributed to the aforementioned filters, which we aim to estimate. Given any two renderers R_a and R_b, set of materials M_i, $i = 1,\ldots,k$, the knowledge of such filters f_{a,M_i} and f_{b,M_i} allows to remove the effect of R_a from an image I_a, by computing $\hat{I}_a = {f_{a,M_i}}^{-1}(I_a)$, $i = 1,\ldots,k$; the subsequent application of f_{b,M_i} on $\hat{I}_a$ provides an image which visually mimics the typical output of R_b.

In the following Sections we describe the main components and steps of our approach, summarized in the following bullet points:

- Characterization of the renderer R in terms of the parameters of the associated filter f_R. The lighting and the geometry used for the characterization are known. This process is performed only once for each renderer.
- Lighting estimation on a given rendered image which needs to be made visually consistent with a different renderer;
- Material segmentation on the input image, in order to apply a per material correction;
- Rendering of a set of spheres, one per each material in the scene, using the statistics of the selected input and output renderers and the estimated lighting;
- Appearance transfer at a pixel level. The surface normal $\mathbf{n}_p$ at pixel p and its material ID m_p are used as entries for lookup table approach on the rendered sphere with the same ID.

Statistical Analysis of Renderers

In previous work [7] it has been noted that the image I of a sphere of a material M, in orthographic projection, rendered under some environment lighting L, can be written as a 2D spherical convolution under the radial symmetry hypothesis: $I = M * L$. Thanks to the radial symmetry it is possible to restrict the analysis to a $1D$ slice of M; in particular, if an angular parametrization (θ,ϕ) based on screen-space normals is used, the analysis can be limited to the θ direction. In this paper we want to understand how a specific rendering software influences the final image and we make the hypothesis that the image I_R, obtained with the rendering tool R, can be written as:

$$I_R = M * L * f_R. \tag{2}$$

where f_R is the set of filters (one per material) introduced by R.

By keeping the material properties as coherent as possible across renderers, and thanks to the known lighting, the statistic proposed in [6] to characterize M can be readily adapted to define the properties of R instead, on which we focus. In particular, we derive the energy Γ, mean μ and variance ρ for each point of the sphere, as described in [6, 7]:

$$\Gamma(I_R) = \Gamma(L_\phi)\alpha(\theta), \tag{3}$$

$$\mu(\bar{I}_R) = \mu(\bar{L}_\phi) - \lambda(\theta), \tag{4}$$

$$\sigma(\bar{I}_R) = \sigma(\bar{L}_\phi) + \psi(\theta), \tag{5}$$

where L_ϕ is the lighting L integrated over the ϕ direction, $\bar{L}_\phi$ and $\bar{I}_R$ are normalized by the energy $\Gamma(I_R)$; the parameters of the filter f_R are given by (α,λ,ψ), which are related to important features like coloring, warping and blurring of the reflected radiance. The statistical analysis is performed on both the diffuse and specular component separately, which on our dataset (described in Section are both known.

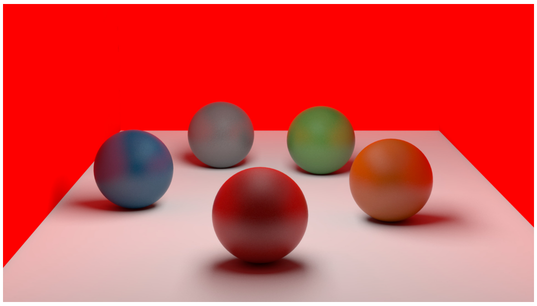
Training Set

In order to derive the filters characteristics, we built a dataset tailored for the automotive industry as a case study, which consists of a Cornell box containing spheres. Each spheres has a material assigned to it out of the following six material classes: car paint, plastic, wood, glass, fabric and leather, the most common ones in a car interior. All walls of the Cornell box are set, in turn, to the same primary RGB color; additionally, a setting with uniformly white walls is included. Inside each of these boxes is in turn placed a spheres with a different material sample. Hence, in total we have 6 classes, 5 samples per each class and 4 colored walls, with a total of 120 scenes, each of which rendered with all the selected renderers, chosen among the most used ones: Autodex 3Ds Max, Autodex Maya and Blender (see Figure 2, Figure 3 and Figure 4 for same examples and visual comparisons). Our dataset allows us to characterize, for each class of materials, the influence of a renderer on the visual appearance, together in addition to simple information such as global and local histograms per color channel.

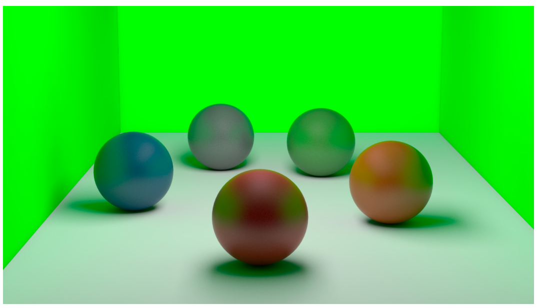
Textured materials with a rough surface, like the fabric sample in Figure 4, display mesoscopic effects like inter-reflections, self-occlusions and self-masking, hence a Bidirectional Texture Function (BTF) should be used to describe their properties, thus preventing a direct application of the framework described in [6]. However, as observed by Dana *et. al.* [10] the BRDF is able to describe material variation of a textured material at a coarse scale, hence averaging through the sample leads to its BRDF.

Image Segmentation and Materials Map

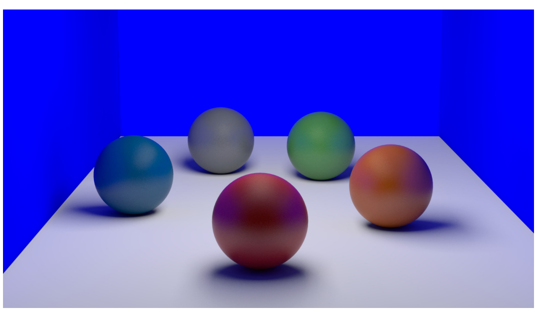
The material categories considered in this study are sufficient to faithfully describe most of the appearance of a car model, and at the same time they are included in the MINC dataset [8], thus enabling a straightforward use of the GoogLeNet [11] network for the image segmentation into material classes, useful to obtain a map with material IDs from an unlabelled input rendering. We augmented the MINC dataset with 100 images including car paint and trained the segmentation network. Once the material labels are know, we can remove the relevant material pixels from the image to estimate the lighting and select the correct statistics to use.



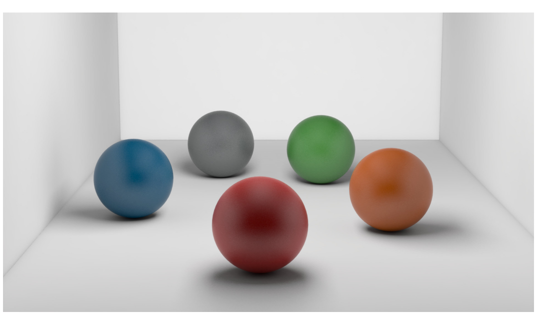
(a)



(b)

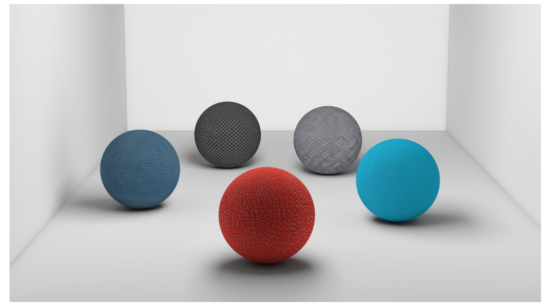


(c)

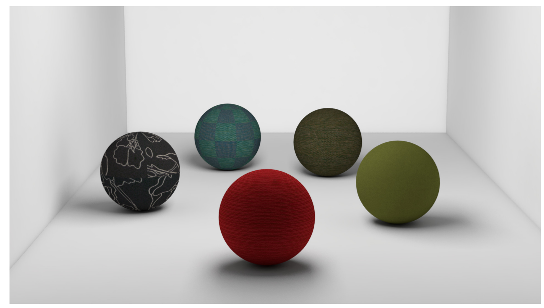


(d)

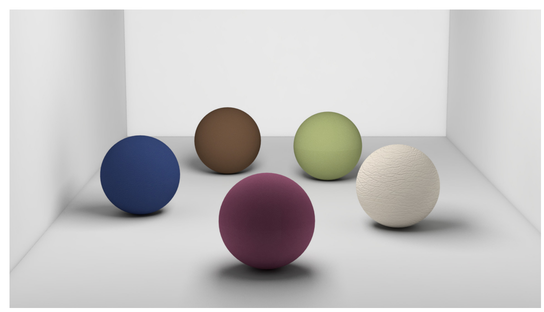
Figure 2. *Spherical samples for the category "car paint", inside the colored boxes, rendered with Autodesk Maya*



(a) Plastic

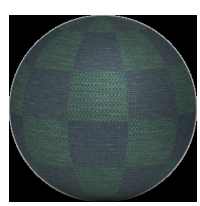


(b) Fabric

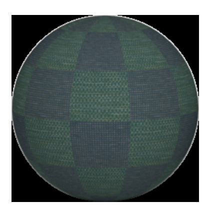


(c) Leather

Figure 3. *From top to bottom: 5 samples in the category "plastic"(a), "fabric"(b) and "leather"(c), placed inside the white box and rendered with Autodesk Maya*



(a) Blender



(b) Maya

(c) 3ds Max

Figure 4. *From left to right: fabric samples inside the white box rendered respectively with Blender, Autodesk Maya, 3ds Max. All the renderings have been linearly rescaled in order to have the same average value per each color channel.*

Incident Lighting Estimation

In order to estimated the incident lighting, we follow the approach described in [8]. All the background pixels together provide direct information for a subset of incident direction. The pixel belonging to the object and the area outside the image prevent a complete estimation of the incident lighting, which hence needs to be approximated. The object is removed from the image and the hole left by the removal process is filled by a simple inpainting technique, which copies pixels both from the left and from the right parts of the hole, using blending weights determined by the distance from the original location in the background to the new position in the hole [8]. From the inpainted image, placed on the xy plane, the biggest possible circle is cut from its central area (*i.e.* the center of the image and the center of the circle coincide and the diameter of the circle is equal to the smallest dimension of the image) and it is mapped onto a hemisphere, extruded along the z-direction; the conventional 3D sphere used for image based lighting is obtained by mirroring the hemisphere along the z-axis.

Appearance transfer at pixel level

At this stage, the only missing information is the diffuse and specular components of the materials in the input rendered image. In our work we do not focus on diffuse/specular separation but rather provide the user with a tool to select the diffuse color for each material, corresponding to $\alpha(0)$ in 3, subtracted from the corresponding pixels and hence providing an approximation of the specular component. The knowledge, for each pixel p, of the material IDs and surface normals $\mathbf{n}_p$, allows us to perform a mapping $(n_{p,x}, n_{p,x}, n_{p,x}) \to (\theta, \phi)$ on the sphere and generate a corrected value by applying Equations 3, 4 and 5 for the reference render, accounting for the estimated incident lighting as in Equation 2. The new pixel value is then applied to the input rendering image, thus emulating the appearance of a typical output of the reference renderer.

(a) Input Image (b) Histogram Matching

Figure 5. *Input image (a) and the corresponding output (b) applying a naive histogram matching to the reference image (3(c)).*

Experiments and Results

Please note that our method does not require to render the scene with the reference software, which is only used to derive the dataset for characterization and to obtain the ground truth in our experiments. Even with the additional key information given by the actual rendering with the reference software, a naive solution, such as a simple global (or even local) histogram matching, clearly cannot deliver accurate material appearance and introduces noticeable artifacts (see Figure 5).

Our method has been tested using typical scenes from the automotive industry (see Figure 1 for an example of an input-output pair, in which the histogram of the background has been matched to the reference data derived from the training set). preliminary results show the effectiveness of the proposed method. In Figure 6 a sketch car interior rendering 6(a) has been processed and the hallucinated output 6(d) can be visually compared to the ground truth obtained from the reference render 6(b); in the same Figure it is possible to see the output obtained by inverting the roles of the renderers 6(c). In Figure 7 a different car scene and model is reported.

The proposed solution can be readily extended to other renderers and the reference software can be seamlessly replaced by a different one.

Although our approach provides promising preliminary results, the employed approximations necessarily lead to some limitations, which we aim to address in future work. A first limitation is clearly due to the set of materials considered in this study, tailored for the automotive industry. A more general solution would require a broader set of materials, which could potentially lead to a vast amount of raw data to analyse and store.

The inclusion of other classes of materials could however pose additional issues related to the material segmentation, in particular in case of materials not included in the MINC dataset: in these cases, it is important to keep the dataset well balanced not to affect the segmentation performance, thus requiring thousands of samples to label; we partly faced this issue when including the "car paint" class, which can be misclassified (*e.g.* the assigned material label could be "metal", more represented in the dataset).

Another source of inaccuracy is due to the simple estimation of the environment lighting, which also poses the additional constraint of having a background area considerably bigger than the object to inpaint: such a limitation is particularly relevant in case of close up scenes.

Similarly to [7], our approach cannot mimic inter-reflections or shadowing effects, thus leading to noticeable errors in large shadow areas (see Figure 6). This could be mitigated by using an artist-created occlusion map. Currently our implementation requires a normal map in input, which can prevent the application of our technique in case of already existing rendered image for which it is only known the source renderer. In such situations a shape from shading approach could allow to derive a depth map and from this an approximated normal map [8], although it could be particularly challenging in case of metals and glass, a particularly common situation in the automotive industry.

Finally, each new release of a rendering tool might require a new characterization, although each new candidate release generally involve the use of benchmarks to guarantee visual consistency with previous versions. A different scenario would simply require a renderer/release characterization, which would simply increase the size of the dataset and of the database of parameters to store.

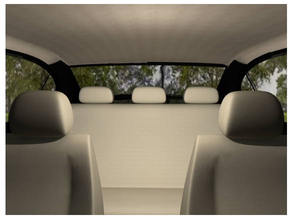
Conclusions

Up to now there are only few methods and tools that allows to edit BRDFs, typically in an interactive way on 3D scenes; a recent survey is reported in [1]. We do not editing BRDF materials but rather hallucinate their appearance in a post-rendering step, making them visually consistent with a reference software. Despite of the current limitation, we believe that the proposed work can be a significant step towards a tool independent, consistent material appearance across different renderers, in an intuitive and transparent way for artists.

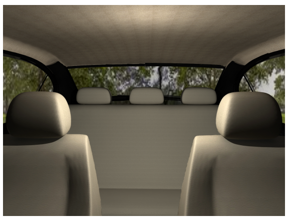
Acknowledgements

We would like to thank Marting Boeckenfeld who rendered the dataset used for characterizing the selected renderers. We thank Jessica Catania and Marting Boeckenfeld for their support with the renderings in Blender of the BMW M5 and I8 car models.

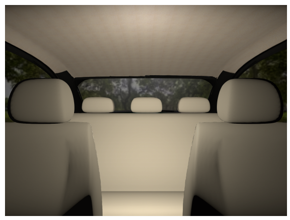
(a) Blender Output



(b) 3ds Max Output

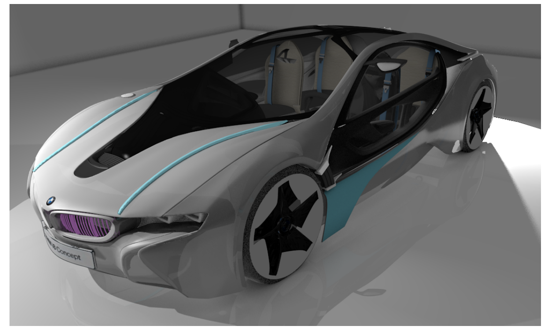


(c) 3ds Max to Blender hallucination

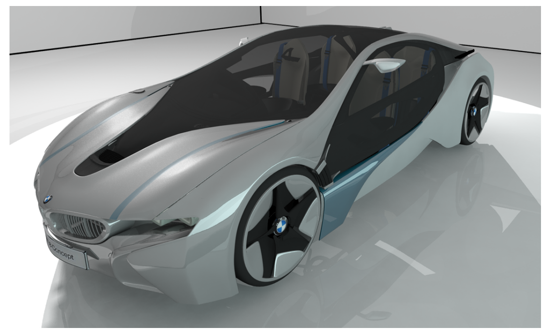


(d) Blender to 3ds Max hallucination

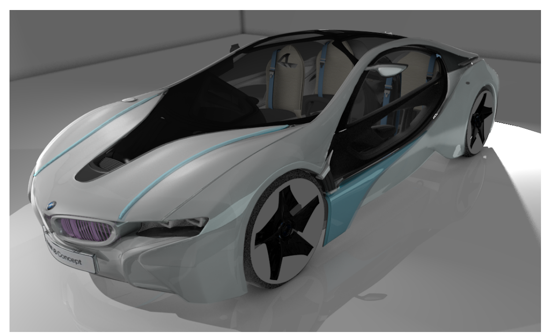
Figure 6. *Car interior scene from the design stage. In (a) the scene is rendered with Blender, in (b) with 3ds Max; In (c) the image (b) has been processed in order to match (a). In (d) the image (a) has been processed in order to match (b).*



(a) Blender Output



(b) 3ds Max Output



(c) Blender to Max hallucination

Figure 7. *Car exterior scene in an advanced design stage. In (a) the scene is rendered with Blender, in (b) with 3ds Max; In (c) the image (b) has been processed in order to match (a). Please note that the view point in (a) and (b) are slightly different, as well as the lighting. The image has been cropped for visualization purposes, removing part of the background.*

References

[1] D. Guarnera, G. Guarnera, A. Ghosh, C. Denk, and M. Glencross, "Brdf representation and acquisition," *Computer Graphics Forum*, vol. 35, no. 2, pp. 625–650, 2016.

[2] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Trans. Graph.*, vol. 22, pp. 759–769, July 2003.

[3] A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of brdf models," in *Proceedings of the Eurographics Symposium on Rendering*, pp. 117–226, Eurographics Association, 2005.

[4] L. Kettner, M. Raab, D. Seibert, J. Jordan, and A. Keller, "The material definition language," in *Proceedings of the Third Workshop on Material Appearance Modeling: Issues and Acquisition*, MAM '15, (Aire-la-Ville, Switzerland, Switzerland), pp. 1–4, Eurographics Association, 2015.

[5] F. Nicodemus, J. Richmond, J. Hsia, I. Ginsberg, and T. Limperis, "Geometrical considerations and nomenclature for reflectance, natl," *Bur. Stand. Rep., NBS MN-160*, 1977.

[6] C. J. Zubiaga, L. Belcour, C. Bosch, A. Muoz, and P. Barla, "Statistical analysis of bidirectional reflectance distribution functions," vol. 9398, pp. 939808–939808–14, 2015.

[7] C. J. Zubiaga, A. Muñoz, L. Belcour, C. Bosch, and P. Barla, "MatCap Decomposition for Dynamic Appearance Manipulation," in *Eurographics Symposium on Rendering 2015*, (Darmstadt, Germany), June 2015.

[8] E. A. Khan, E. Reinhard, R. W. Fleming, and H. H. Bülthoff, "Image-based material editing," *ACM Trans. Graph.*, vol. 25, pp. 654–663, July 2006.

[9] S. Bell, P. Upchurch, N. Snavely, and K. Bala, "Material recognition in the wild with the materials in context database," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3479–3487, June 2015.

[10] K. J. Dana, B. Van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Transactions on Graphics (TOG)*, vol. 18, no. 1, pp. 1–34, 1999.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, June 2015.

Author Biography

Dar$^{\prime}$ya Guarnera is a PhD student at Loughborough University. She obtained her Postgraduate Certificate in Education from the Liverpool Hope University and first-class degree in Computer Science from the Liverpool Hope University. In 2006 she received Student of the year award at graduation from Liverpool Community College in 3D modelling. She also obtained BSc in Architecture from the Odessa State Academy of Civil

Engineering and Architecture. Her interests includes 3D Architectural Visualisation, virtual materials, digital art and sculpture.

Giuseppe Claudio Guarnera is a Research Associate at NTNU - Norwegian University of Science and Technology (Norway). Previously he was a Research Associate at Loughborough Universiy (UK), working on material representation and acquisition. His research interests include Computer Vision, Pattern Recognition, Appearance Modelling and Computational Photography. He received his PhD in Computer Science from the University of Catania (Italy), with a doctoral dissertation in the area of Computer Vision and Pattern Recognition.

Cornelia Denk is currently working at BMW Group Research and Technology in Munich, Germany. A key part of her role is to establish a number of high profile international joint industry-university research initiatives for BMW Group. Her specialist research areas are in Computer Graphics, Global Illumination, Augmented & Virtual Reality, Process Validation, Sensors and I/O Devices. She is involved in the Computer Graphics community through a range of activities, including service on the board of the Munich ACM SIGGRAPH Chapter, ACM-W Europe subcommittee and mentoring initiatives to help female students in technology.

Mashhuda Glencross obtained her PhD in Computer Science from the University of Manchester. Prior to joining Loughborough as a Lecturer, she worked at ARM in Cambridge and as a postdoc at the University of Manchester. Her research interests include virtual reality, 3D reconstruction/relighting, material appearance acquisition, novel user interfaces and human visual perception. She is treasurer of the ACM Europe Council, chair of the ACM SIGGRAPH Professional and Student chapters committee and served as courses chair for SIGGRAPH 2014. She is currently also a member of the ACM Publication Board Conference Committee.