A Similarity Measure for Material Appearance

Manuel Lagunas[1], Sandra Malpica[1], Ana Serrano[1], Elena Garces[2], Diego Gutierrez[1], Belen Masia[1]

[1] Graphics & Imaging Lab
Instituto de Investigación en Ingeniería de Aragón (I3A)
Universidad de Zaragoza, Mariano Esquillor s/n, 50018, Zaragoza, Spain.
Tel. +34-976762707, e-mail: *mlagunas@unizar.es*
[2] Universidad Rey Juan Carlos

Abstract

We present a model to measure the similarity in appearance between different materials, which correlates with human similarity judgments. To this end, we introduce a novel dataset, and gather data from crowdsourced experiments to train a deep learning architecture that correlates with perceived appearance similarity. Our evaluation shows that our model outperforms existing metrics.

Introduction

Humans are able to recognize materials, compare their appearance, or even infer many of their key properties effortlessly, just by briefly looking at them. Many works propose classification techniques, although it seems clear that labels do not suffice to capture the richness of our subjective experience with real-world materials [1]. Unfortunately, the underlying perceptual process of material recognition is complex, involving many distinct variables; such process is not yet completely understood [2].

Given the large number of parameters involved in our perception of materials, many works have focused on individual attributes (such as the perception of gloss [3], while others have focused on applications like material editing [4], or filtering [5]. However, the fundamentally difficult problem of establishing *a similarity measure for material appearance* remains an open problem.

We have developed a novel image-based material appearance similarity measure derived from a learned feature space. This is challenging, given the subjective nature of the task, and the interplay of confounding factors like geometry or illumination in the final perception of appearance. Very recent work suggests that perceptual similarity may be an emergent property, and that deep learning features can be trained to learn a representation of the world that correlates with perceptual judgements [6].

Inspired by this, we rely on a combination of large amounts of images, crowdsourced data, and deep learning. We create a diverse collection of stimuli covering a wide variety of isotropic appearances, shapes, and environment maps. Using triplets of images, we gather information where participants are asked which of two given examples has a more similar appearance to a reference. From this information, we learn a model that enforces learning of the perceptual information collected from humans, and the main features that describe a material in an image; this allows us to learn the notion of material appearance similarity, dependent on both the visual impression of the material, and the actual physical properties of it. Our work has been published in ACM Transactions on Graphics [7] and will be presented at SIGGRAPH 2019.

Collecting similarity information

By combining the 100 real-world measured materials in the MERL dataset [8], six unique illumination conditions, and 15 different scenes, we generate a total of 9,000 dataset samples. For each one we provide: the rendered HDR image, a corresponding LDR image, along with depth, surface normals, alpha channel, and ambient occlusion maps. Using the previously created dataset, we launch a study that deals with the perception of material appearance. We gather data in the form of relative comparisons, following a 2AFC scheme; the subject is presented with a triplet made up of one *reference* material, and two *candidate* materials, and their task is to answer the question *Which of these two candidates has a more similar appearance to the reference?* by choosing one among the two candidates. A total of 603 participants took part in the test. Users were not aware of the purpose of the experiment.

Implementation and results

We use the ResNet architecture [9], based on its generalization capabilities and its proven

performance on image-related tasks. For training, we use image data from our materials dataset together with human data on perceived similarity. We train our model using a loss function consisting of two terms, equally weighted. The first term introduces the collected MTurk information on appearance similarity, the second loss term maximizes the log-likelihood of the model choosing the same material as humans. We evaluate our model on the set of images of the material dataset not used during training. Table 1 compares the accuracy and perplexity on data collected from humans, against our model, and known metrics from the literature. The results show how our model outperforms existing metrics and agrees with human responses almost 81% of the time. The learned feature space allows for several applications, including perceptual clustering (see Figure 1), database summarization, visualizations, gamut mapping, or material suggestions. For a detailed explanation of the loss function, training details, and the proposed applications of our model, we refer the reader to [7].

Conclusions

We have presented and validated a model of material appearance similarity that correlates with the human perception of similarity. Our results suggest that a shared perception of material appearance does exist. Nevertheless, material perception poses many challenges; as such there are many exciting topics not fully investigated in this work. Several factors come into play that influence material appearance, i.e., the visual impression of a material, in a highly complex manner; fully identifying them and understanding their complex interactions is an open, fundamental problem.

REFERENCES

[1]. FLEMING, Roland W. Material Perception. Annual review of vision science, 2017, vol. 3, p. 365-388.

[2]. ANDERSON, Barton L. Visual Perception of Materials and Surfaces. Current Biology, 2011, vol. 21, no 24, p. R978-R983.

[3]. PELLACINI, Fabio; FERWERDA, James A.; GREENBERG, Donald P. Toward a Psychophysically-Based Light Reflection Model for Image Synthesis. Proceedings of the 27th annual conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., 2000. p. 55-64.

[4]. SERRANO, Ana, et al. An Intuitive Control space for Material Appearance. ACM Transactions on Graphics (TOG), 2016, vol. 35, no 6, p. 186.

[5]. JARABO, Adrian, et al. Effects of Approximate Filtering on the Appearance of Bidirectional Texture Functions. IEEE Transactions on Visualization and Computer Graphics, 2014, vol. 20, no 6, p. 880-892.

[6]. ZHANG, Richard, et al. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018. p. 586-595.

[7]. LAGUNAS, Manuel, et al. A Similarity Measure for Material Appearance. ACM Transactions on Graphics (TOG), 2019, vol. 38, no 4.

[8]. MATUSIK, Wojciech, et al. A Data-Driven Reflectance Model. ACM Transactions on Graphics (TOG), 2003

[9]. HE, Kaiming, et al. Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. p. 770-778.

Table 1. Accuracy and perplexity of our model compared to human performance, an oracle (which always returns the majority opinion), and six other metrics from the literature. For accuracy, higher values are better, while for perplexity lower are better.

EVALUATION OF OUR MODEL

Metric	Accuracy		Perplexity	
	Raw	Majority	Raw	Majority
Humans	73.10	77.53	-	-
Oracle	83.79	100.0	-	-
RMS	61.63	64.72	3.61	3.13
RMS-cos	61.60	64.67	3.86	3.33
Cube-root	63.71	67.40	1.96	1.86
L2-lab	63.76	67.21	2.16	2.07
L4-lab	60.60	62.93	15.36	11.66
SSIM	62.35	64.74	2.02	1.94
Our model	**73.97**	**80.69**	**1.74**	**1.55**

Figure 1. Material suggestions using our perceptual database clustering. The images show random materials assigned from three different clusters of varying appearance. The robot model (cKalten) was obtained from TurboSquid.