Image Features and Perception of Object Surface Qualities

Takehiro Nagai[1]

nagai.t.aa@m.titech.ac.jp
[1]Tokyo Institute of Technology
4259-G2-1 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8502, Japan

Keywords: surface quality perception, image features, shitsukan, psychophysics

ABSTRACT

Different types of image cues for surface quality perception, such as glossiness, have been discovered by psychophysical studies. In this talk, I will discuss basic strategies of the visual system to grasp surface qualities and recent progress in research on surface quality perception by showing some psychophysical results.

1 Introduction

Past studies on human visual psychophysics and physiological studies have classically elucidated the mechanisms for relatively simple perceptual features such as color, motion, and shape of visual stimuli. In addition, the perception of object surface qualities or material (often called “shitsukan”), such as glossiness and transparency, is also an essential type of visual perception. Many psychophysical and neuroimaging studies have recently tried to uncover the mechanisms of shitsukan perception. The retinal images are created as a result of complex interactions between different physical factors: illumination environments and the object’s optical properties and shapes. Among these factors, the object’s optical properties are considered to associate with perceptual shitsukan. Although estimating the object’s optical properties only from retinal images is an ill-posed problem, the human visual system seems to grasp them as a shitsukan perception pretty well. Therefore, the mechanisms underlying this shitsukan perception are of great interest to researchers.

However, it is not easy to elucidate the shitsukan perception mechanisms. First, the visual stimuli must be naturalistic. Therefore, it is difficult parametrically manipulate the physical parameters of the stimuli. In addition, the shitsukan perception is composed of many perceptual features as compared with, for instance, color perception. In other words, the experimental methods used for simple perceptual features cannot be applied to the studies on shitsukan perception. Of course, recent advances in image technologies like computer graphics help us to create naturalistic visual stimuli, and understanding of the shitsukan perception have been progressed little by little.

In this paper, I introduce some basic findings reported by recent psychophysical studies.

2 Correlation of simple image features and surface quality perception

2.1 Glossiness perception

Glossiness perception has been investigated most actively among various kinds of surface qualities. For example, there have been several reports that the glossiness perception consists of at least two psychological dimensions [1,2]. In addition, a model associating these perceived dimensions with the object's optical properties has also been proposed [1]. This model helps create object images exhibiting desired glossiness in computer graphics rendering.

One of the interesting properties of glossiness perception is the correlation with simple image features in visual stimuli. In the 2000s, studies focusing on the relationship between image features and shitsukan perception have gradually increased. For instance, Motoyoshi et al.[3] found that the shape of the luminance histogram captured by the luminance skewness of an object image strongly correlates with perceived glossiness by analyzing many object images with different degrees of glossiness. Furthermore, they also reported that the glossiness of an object image could be manipulated by changing its luminance skewness. Their results motivated many researchers to tackle finding the relations between some image features and shitsukan perception based on the assumptions that the shitsukan perception relies on heuristics (empirical rules about image features) and that the mechanism of shitsukan perception may be understood in a simpler way as expected. These studies reported that, for instance, luminance rms (root-mean-square) contrast [4] and luminance gradient [5] are correlated with perceived glossiness, especially for more natural stimuli than the stimulus set in Motoyoshi et al. [3]. All these image features seem to be related to specular highlights that clearly contribute to glossiness perception. For instance, object images with specular highlights tend to have high luminance contrast.

In addition, chromaticity, not only luminance, contributes to glossiness perception. Adding chromatic colors on diffuse reflection components or both diffuse and specular reflection components of object images enhances perceived glossiness (Figure 1)[6]. Also, the magnitude of the glossiness enhancement depends on the stimulus hue. This hue dependence can be mainly explained by the perceptual brightness increase in highlight regions due to the Helmholz-Kohlaush effect (H-K effect). However, there also found some cases where glossiness was improved even when the H-K effect was not strong enough. Thus, there may also be a contribution of colors specific to glossiness other than the H-K effect, such as color contrast between highlights and shadows and/or higher-order color distributions relevant to glossy object surfaces.

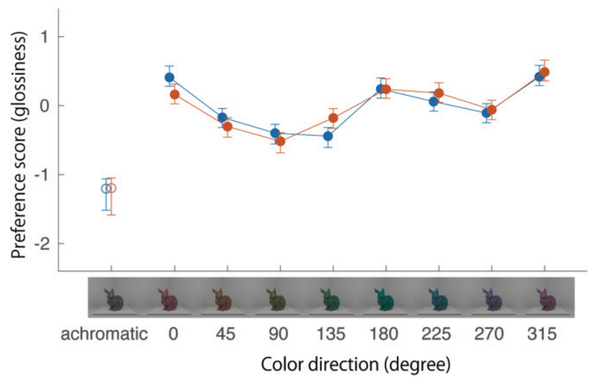

Fig. 1. Glossiness on colored object images.[6]

Glossiness can be perceived even on object surfaces on which specular highlights are not very clear. For instance, perceived glossiness can be maintained in some cases even when the luminance of the specular highlight region of a glossy object image is artificially lowered [7]. This phenomenon indicates that the glossiness cues exist in low-level image regions. On such object surfaces, rms contrasts of the high spatial frequency components correlate with perceived glossiness [8, 9]. In addition, manipulating the perceived contrast in the high spatial frequency bands by adaptation or image manipulation affects perceived glossiness [8, 10]. These high spatial frequency contrast corresponds to luminance edges made by the mirror-like reflection of the surrounding environments included in the specular reflection component and thus are thought to act as a cue for perceived glossiness. The spatial frequencies of the specular reflection components, of course, largely depend on the micro-scale surface roughness of the object. Thus, the spatial frequency band effective for manipulating perceived glossiness varies from object to object (Figure 2A). On the other hand, in the situation where the glossiness perception must depend on the low-luminance components, such as on stimuli with low highlight luminance, manipulation of a sub-band contrast in specific spatial frequency bands has a substantial impact on glossiness regardless of object surface roughness (Figure 2B) [10]. This dependency on specific spatial frequency components may imply the strategy of glossiness perception of the human visual system.

However, as easily expected, simply having luminance

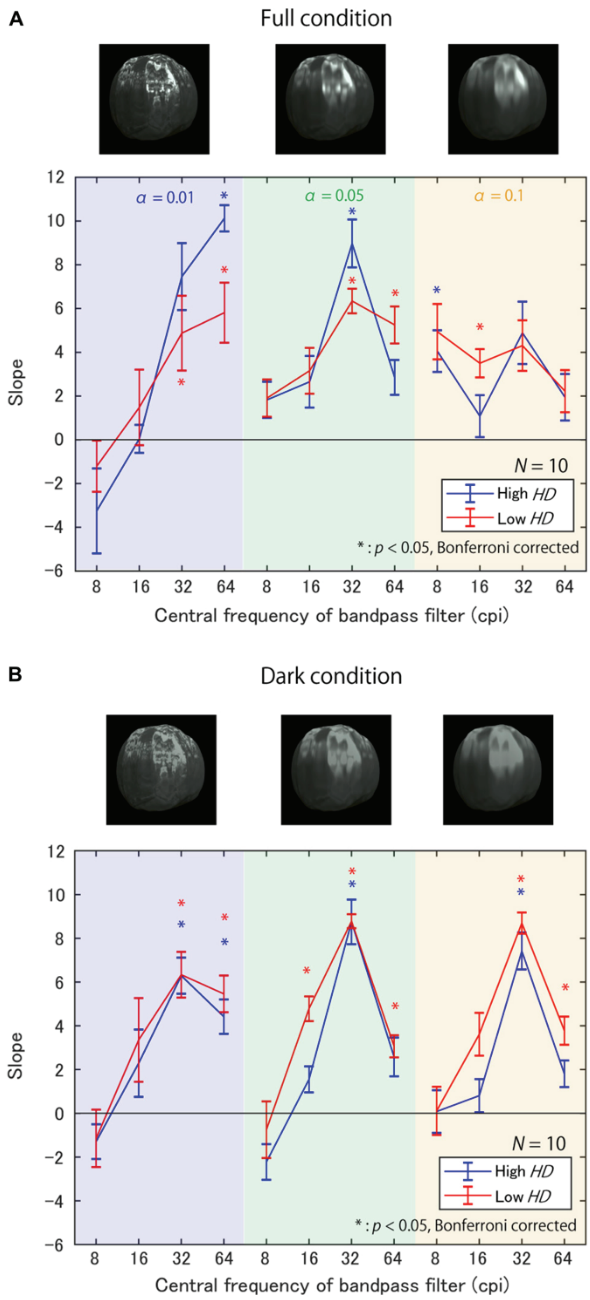

Fig. 2. Effects of increasing spatial-frequency sub-band contrast on perceived glossiness for stimuli with (A) strong and (B) weak specular highlights [9]

edges (or high spatial frequency components) does not necessarily lead to high glossiness. First, spatial alignment between luminance shading (or object shapes) and luminance edges is crucial for the luminance edges to contribute to glossiness [11]. This indicates that the relationship between simple image features and perceived glossiness, as described above, can be found only under complex conditions. However, among the image features above, the luminance gradient is considered to reflect a more sophisticated perceptual strategy than others because the luminance gradient may be relevant to the dissociation of specular and diffuse reflections [5].

2.2 Translucency perception

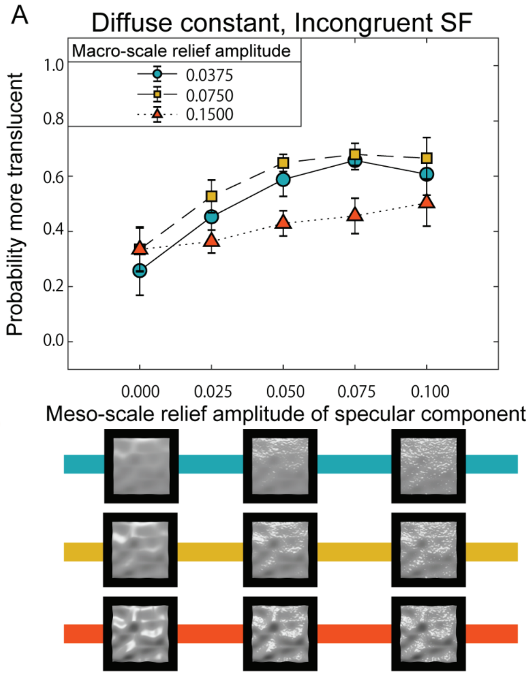

Fig. 3. Effects of manipulating relief height of specular reflection components on perceived translucency [13].

Transparency (and translucency) perception, which has been classically investigated with two-dimensional visual stimuli as represented by Metelli's law, has also been studied as one of the shitsukan features using three-dimensional object images. For instance, the contrast of the high-spatial frequency components in diffuse reflection has been reported to be relevant to translucency perception. The reduction or inversion of rms contrast in the high spatial frequency diffuse components induces translucency and/or transparency perception [12]. In addition, changing the relief height of specular reflections, not only the diffuse reflections, also alters the perceived translucency (Figure 3) [13]. The relations between specular and diffuse reflection components seem to underlie translucency perception.

Although the mechanisms underlying the relationship between these translucency and image features are still highly unclear, one possibility is that image features related to the differences in the shapes represented by the diffuse and specular reflection components may contribute to perceived translucency. For instance, the difference in the orientation distributions of the specular and diffuse reflection components explains the degree of translucency perception well [13]. The shape information of the object is reflected in the orientation distributions. Furthermore, a regression model based on the orientation distribution differences explained perceived translucency better than a model based on shape perception measured in a psychophysical experiment, suggesting the possibility that image features related to shape perception, not only perceived shape perception itself, contribute to translucency perception.

2.3 Other shitsukan features

Some other types of shitsukan perception seem relevant to simple image features, similarly to glossiness and translucency. For example, the perception of cloth qualities, such as thickness and softness, strongly correlates with spatial frequency subband contrasts [14]. Wet impressions also correlate with low-order moment image statistics of luminance and color [15]. These studies also showed that the manipulations of the image features alter perceived shitsukan. It should be noted, however, that in all cases, the correlations between image features and shitsukan perception are found in limited conditions, like comparing images of the same object. Thus, such correlations are insufficient to predict shitsukan perception across different objects.

3 Importance of more complex image features

There are severe limitations in the relations between shitsukan perception and simple image features. Even for glossiness, which seems to be one of the most basic surface qualities, simple image features can explain the perception only in limited situations, such as comparing glossiness between objects with the same shape and made of uniform materials. In addition, object shapes are also strongly associated with glossiness and translucency perception. Notably, just changing the perceived object shapes by motion and binocular disparity has been reported to modulate perceptual glossiness and translucency even if the luminance patterns on the stimuli were held constant [16]. This fact suggests the difficulty in predicting shitsukan perception based only on simple image features.

Instead, the visual system is considered to utilize more complex image information for shitsukan perception, including shape-related ones. For instance, as mentioned above, the relative correspondence between highlights and shadows seems essential for shitsukan perceptions. However, even the mechanisms for separating highlights from shadows, which is a prerequisite for this association, are still highly unclear.

Recently, some deep learning models have been found to behave similarly to human perception. There are systematic errors in shitsukan perception with physical properties. Some deep learning models have been shown to make errors in gloss judgments similar to humans [17]. These deep learning models may provide us clues to mechanisms of shitsukan perception, which cannot be obtained from simple image features.

4 Discussion

Simple features in object images are often correlated with shitsukan perception under certain conditions. This correlation with shitsukan perception can be applied to some image representation techniques. For instance, you may cut out an object region from a displayed image and manipulate the luminance contrast and skewness in that region to slightly manipulate the glossiness representation. Also, such knowledge about the image features for shitsukan perception may be informative in constructing a "shitsukan rendering index" of illumination environments.

On the other hand, this finding does not mean that the visual system really grasps shitsukan based on these image features. Instead, more complicated image information is crucial for shitsukan perception; the correlation of simple image features and perception may arise from the correlations of the simple image features with more complex image features the visual system relies on for shitsukan perception. For instance, glossiness differences across objects with different materials and shapes should be hard to be predicted just from simple image features. Also, even the discrimination between specular highlights and high lightness (white surface) regions is difficult for models based on simple image features. Recently, deep learning models are expected to be a tool to understand human shitsukan perception. However, the information represented in the deep learning models is not always understandable for researchers. It may be crucial to consider what we can learn from deep learning in the near future. It will require considerable time before the mechanism of shitsukan perception is firmly understood.

5 Conclusions

This paper outlined the relationships between shitsukan perception and simple image features. Unfortunately, we are still far from a deep understanding of the mechanisms of shitsukan perception. Nevertheless, such image features may be utilized as subelements to manipulate shitsukan impressions in images. I hope we further advance our understanding of the shitsukan perception mechanisms from both the scientific and engineering perspectives.

References

[1] J. A. Ferwerda, F. Pellacini, and D. P. Greenberg, "Psychophysically based model of surface gloss perception," *Proc. SPIE 4299, Human Vision and Electronic Imaging VI*, pp. 291-301 (2001).

[2] M. Toscani, D. Guarnera, G. C. Guarnera, J. Y. Hardeberg, and K. R. Gegenfurtner, "Three Perceptual Dimensions for Specular and Diffuse Reflection," *ACM Trans. Appl. Percept.* vol. 17, no. 2:6 (2020).

[3] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson, "Image statistics and the perception of surface qualities," *Nature*, vol. 447, no. 7141, pp. 206-209 (2007).

[4] C. B. Wiebel, M. Toscani, and K. R. Gegenfurtner, "Statistical correlates of perceived gloss in natural images," *Vision Res*, vol. 115, no. part B, pp. 175-187 (2015).

[5] M. Sawayama and S. Nishida, "Material and shape perception based on two types of intensity gradient information," *PLoS Comput. Biol.*, vol. 14, no. 4: e1006061 (2018).

[6] K. Koizumi, and T. Nagai. Highlight brightness increased by H-K effect is a dominant factor of color-induced glossiness enhancement on chromatic surfaces (in revision).

[7] J. Kim, P. J. Marlow, and B. L. Anderson, "The dark side of gloss,", *Nat Neurosci*, vol. 15, no. 11, pp. 1590-1595 (2012).

[8] J. Kim, K. Tan, and N. S. Chowdhury, "Image statistics and the fine lines of material perception," *i-Perception*, vol. 7, no. 4 (2016).

[9] H. Kiyokawa, T. Tashiro, Y. Yamauchi, and T. Nagai, "Luminance edge is a cue for glossiness perception based on low-luminance specular components," *J Vision*, vol. 19, no. 12:5 (2019).

[10] H. Kiyokawa, T. Tashiro, Y. Yamauchi, and T. Nagai, "Spatial frequency effective for increasing perceived glossiness by contrast enhancement," *Front Psychol*, vol. 12:625135 (2021).

[11] J. Kim, P. Marlow, and B. L. Anderson, "The perception of gloss depends on highlight congruence with surface shading," *J Vision*, vol. 11, no. 9:4 (2011)

[12] I. Motoyoshi, "Highlight-shading relationship as a cue for the perception of translucent and transparent materials," *J Vision*, vol. 10, no. 9:6 (2010).

[13] H. Kiyokawa, T. Nagai, Y. Yamauchi, and J. Kim, "The perception of translucency from surface gloss," *Vision Research* (in press).

[14] M. Giesel and Q. Zaidi, "Frequency-based heuristics for material perception," *J Vision*, vol. 13, no. 14:7 (2013).

[15] M. Sawayama, E. H. Adelson, and S. Nishida, "Visual wetness perception based on image color statistics," *J Vision*, vol. 17, no. 5:7 (2017).

[16] P. J. Marlow, J. Kim, and B. L. Anderson, "Perception and misperception of surface opacity," *P Natl Acad Sci USA*, vol. 114, no. 52, pp. 13840-13845 (2017).

[17] K. R. Storrs, B. L. Anderson, and R. W. Fleming, "Unsupervised learning predicts human perception and misperception of gloss," *Nat Hum Behav*, vol. 5, no. 10, pp. 1402-1417 (2021).