Construction of Manga Materials Database for Analyzing Perception of Materials in Line Drawings

Yuma Saito, Keita Hirai, Takahiko Horiuchi; Graduate School of Advanced Integration Science, Chiba University, Japan

Abstract

Interest in the visual perception of the materials that objects are made of has been growing. Most past studies on visual material-category perception have used stimuli with rich information such as color, shape, and texture. However, we can distinguish material categories from even simple black-and-white line drawings. This paper presents a new attempt to analyze material perception from Japanese "manga" comics which are composed of line drawings typically printed in black and white. In this study, we first collected 400 material objects captured from manga comics, and created 400 corresponding patches that were close-up images of the objects and that excluded shape information. Through psychophysical experiments, 274 pairs of images giving consistent material impressions to observers were chosen. According to our experiments, observers could distinguish material categories from patches with on accuracy of 88.4%. Then for each material, we investigated the low-dimensional image features that contributed to the material perception of black-and-white drawings. In particular, we found image features that represented metals very well.

Introduction

In everyday life, we can distinguish object categories without difficulty by recognizing different shapes and the functions of objects based on visual information. Building on this and on rich scientific information, researchers have recently undertaken studies of the perception of materials contributing to the understanding of object perception.

In the field of computer vision, most object recognition systems have relied on low-level material invariant features such as color; for example, the scale invariant feature transform (SIFT) [1], [2] and the histogram of oriented gradients (HoG) [3] have tended to ignore material information altogether. It may seem trivially easy to represent surface properties using the bidirectional reflectance distribution function (BRDF) [4], [5], the bidirectional texture function (BTF) [6] and the bidirectional surface scattering reflectance distribution function [7] depending on the materials; however, it is nearly impossible to estimate such features from a single image without employing simplifying assumptions [4], [8]. Recently, a few approaches have been proposed in order to directly study the relationships between image features and several perceptual attributes and to estimate the attribute values for a given image [8]-[11]; this research was performed using large image datasets such as the Flickr Materials Database [12].

Several studies have investigated the mechanism of the processing of material sensations in the human brain using various approaches such as functional magnetic resonance imaging (fMRI) [13] and psychophysical studies [14]. Most of these studies have focused on the visual estimation of specific properties of materials [15]-[17], such as glossiness, translucency, and roughness. Taken together, these findings support the general idea that the human visual system can estimate the properties of materials from relatively low-level visual features.

There is experimental evidence to support the hypothesis that human observers excel at recognizing and categorizing materials. For example, Sharan et al. [12] have shown that participants can identify a wide range of materials from photographs after even a very brief exposure. Recently, Fleming et al. [18] showed participants photographs of materials from different categories and asked them to rate various subjective qualities, such as hardness, glossiness, or prettiness. Although the participants were not explicitly informed that the samples belonged to different classes, the subjective ratings of the samples clustered systematically into categories, suggesting that the participants could theoretically classify materials by making visual judgments concerning their properties.

As has been shown by many previous studies, both surface properties (i.e., color, texture, surface reflectance, etc.) and shape are influential in distinguishing materials, as they provide relevant visual information. However, we can also perceive material properties from more primitive information. Tanaka and Horiuchi [19] investigated the perceptual qualities of static surfaces using real materials and four types of image reproductions by varying chromaticity (color vs. gray) and resolution (high vs. low) of the material images. In order to study material perception that is not influenced by shape or saturated color, they used square images of materials in gray tones.

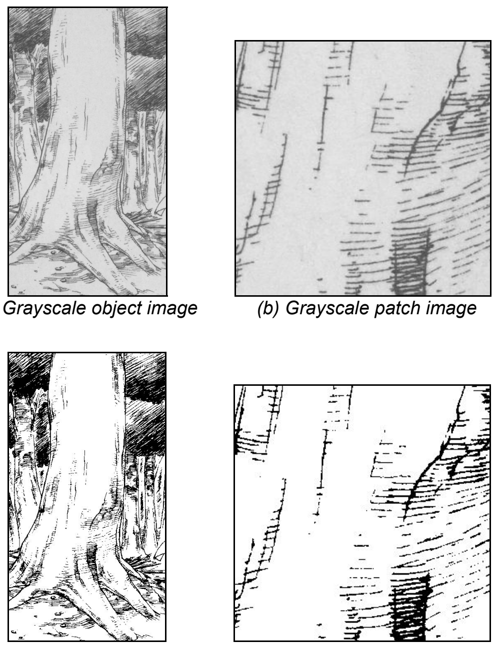
In this study, we focused on more primitive information such as that in black-and-white Japanese "manga" line drawings. Modern-day manga can be defined as comics corresponding to a Japanese style that originated during the mid-1900s. The Japanese manga (Japanese for "comic books") is distinct from traditional Western comic books that it presents fine details. Color manga can express even more semantics and artistic styles; however, mangas are seldom colored. In this study, we first construct a manga materials database. Then we analyze the low-dimensional image features within categorized images that contributed to the material perception of black-and-white images.

Construction of Manga Materials Database

No image dataset of manga materials had previously been created. Our first step was to build a manga materials database ourselves to analyze the line-drawing images.

Collection of Manga Materials

We collected various material images from commercially available Japanese comics drawn by six cartoonists. We digitized the material images with an EPSON GT-X820 scanner. The scanner used a resolution of 600 dpi and saved images in a grayscale 8-bit bitmap format. We extracted 400 rectangular images of objects from the material images captured. From these, we extracted 400 corresponding square patches that were close-up images of the objects and excluded shape information. Since manga consists of frames of various sizes, the shapes of the rectangular object images differed from each other. Locations and sizes of the material images were determined by the authors' subjective judgement in this stage. Finally, these captured grayscale images were converted to binary images by Otsu's algorithm [20]. Examples of a captured grayscale object image and its corresponding grayscale patch image are shown in Figs. 1(a) and (b), respectively, and their binary images are shown in Figs. 1(c) and (d), respectively.



(a) Grayscale object image *(b) Grayscale patch image*

(c) Binary object image *(d) Binary patch image*

Figure 1: A wood example of a manga material image

Evaluation of Manga Materials

As described in the previous subsection, the collected manga material images were selected based on the authors' subjective impression. We conducted an evaluation experiment using pairs of 400 material images that would give observers a common perception of a material category.

In the experiment, captured images were printed out side by side on a white sheet of paper. Subjects selected the most appropriate material from 20 material categories (Fabric, Paper, Metal, Plastic, Rock, Stone, Pottery, Soil, Sand, Mineral, Mirror, Glass, Vinyl, Ice, Water, Snow, Wood, Foliage, Leather, Fire) for each patch image. We instructed the participants as follows: "Each image patch shows a part of comic illustrations. Please select the most appropriate material category." If there was no appropriate category, subjects wrote down a suitable material category freely chosen. If they could not perceive any material from the image, they wrote the answer "Not Recognizable." The experiment was conducted under the standard illuminant D65, and the viewing distance was from 30cm to 40cm. After a few days, the same subjects also selected the most appropriate object images from the 20 material categories. Ten Chiba University students with normal vision participated in this experiment. Five participants evaluated 200 pairs of material images, and the other five participants evaluated the remaining 200 pairs of images.

We extracted material images that were assigned the same material category by more than half of the subjects. As a result, 376 and 276 images were extracted from the object and patch images, respectively. Furthermore, for 244 images, the material categories chosen for the extracted object image and for its patch images matched each other. This indicates that subjects were able to distinguish material categories from patch images alone with an accuracy of 88.4%. Table 1 shows the counts of the extracted object and patch images. Subjects were able to recognize material categories from rock and wood patch images with accuracy of 90%. These material patch images have peculiar material characteristic. On the other hand, the most difficult categories to distinguish material characteristics was soil. Most of the response to the patch images was "Not Recognizable."

Table1. Extracted patch and object images.

	Patch Images	Object Images
Fabric	75	95
Metal	18	45
Stone	10	16
Rock	63	70
Soil	0	11
Sand	11	13
Mineral	1	1
Glass	0	1
Ice	0	4
Water	7	14
Snow	0	3
Wood	77	86
Foliage	13	16
Fire	1	1
Total	276	376

Constructed Materials Database

Finally, by removing those material categories that had only one extracted image, we built a manga materials database having 274 pairs of images, including 75 fabric, 18 metal, 63 rock, 10 stone, 11 sand, 7 water, 77 wood, and 13 foliage images. Examples of patches in the constructed manga materials database are shown in Fig. 2.

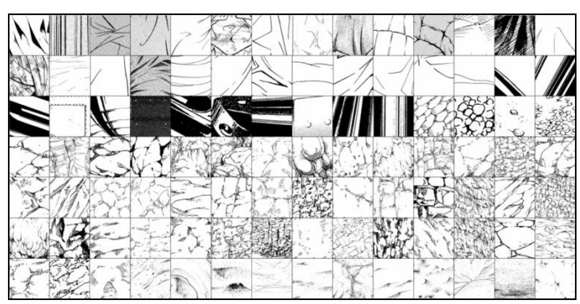


Figure 2: Samples portion of manga materials database

Feature Analysis of Line-Drawing Images

As described in the previous section, material categories can be discerned from the patch images in the constructed database. In this section, we investigate features of patch images that are representative of each material category.

Since manga readers do not pay careful attention to material textures, in this study we investigated the following simple low-dimensional features:

F1. The ratio of black pixels to total pixels.
F2. The average run-length of horizontal black pixels.
F3. The average run-length of vertical black pixels.
F4. The average run-length of horizontal white pixels.
F5. The average run-length of vertical white pixels.
F6. The standard deviation of horizontal black-pixel run-lengths.
F7. The standard deviation of vertical black-pixel run-lengths.
F8. The proportion of longer horizontal run-length for black pixels.
F9. The proportion of longer vertical run-length for black pixels.
F10. The proportion of longer horizontal run-length for white pixels.
F11. The proportion of longer vertical run-length for white pixels.
F12. The average horizontal run-length of low-passed black pixels.
F13. The average vertical run-length of low-passed black pixels.
F14. The number of neighboring black pixels.

Some features were also normalized by the total number of pixels.

By deriving the linear discriminant function from all patch images for each material category, the precision, recall, and F-measure were calculated for the 274 patch images. As features, we examined all combinations of up to four low-dimensional features. Therefore, the image features having the highest F-measures can be considered the low-dimensional features representative of images for the material category. Table 2 shows the F-measure of each material category.

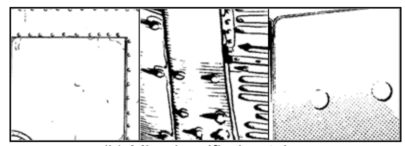
Table2. F-measure of each material category

	F-measure
Metal	0.81
Wood	0.68
Rock	0.62
Water	0.55
Fabric	0.51
Foliage	0.49
Sand	0.41
Stone	0.18

Metal Images

The F-measure value for the metal material category was 0.81 (recall: 0.83, precision: 0.79). This F-measure was the highest of the material categories in the database. The features selected were F1, F2 and F3. Figure 3(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, large black areas or thick straight lines appear, representing the metallic appearance. Since metal material category often have metallic luster expressed in the contrast of the thick black and white lines, the feature F1, F2 and F3 can be a strong clue for the classification. We found that in line-drawing images the ratio of black pixels to total pixels and the thicknesses of the black lines are relevant for representing metal surfaces. Figure 3(b) shows images mis-recognized by the linear discriminant classification but classified into the metal category by the psychophysical experiment. These three patches include boundary edges such as a screw in the images in addition to the surface gloss property.

(a) Patches correctly classified

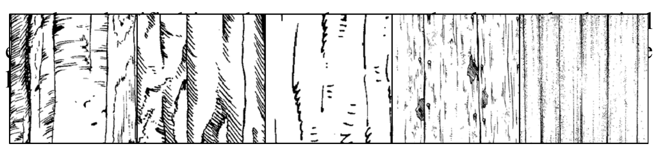


(b) Mis-classified patches

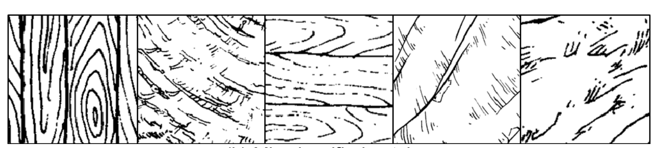
Figure 3: Results of linear discriminant classification for metal.

Wood images

The F-measure value for the wood material category was 0.68 (recall: 0.78, precision: 0.61). The features selected were F8, F9, F10 and F11. Figure 4(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, longer horizontal or vertical straight lines appear, representing the wood appearance. We found that in line-drawing images the frequency of horizontal and vertical black and white longer run-length were relevant for representing wood grain.. Figure 4(b) shows images mis-recognized by the linear discriminant classification but classified into the wood category by the psychophysical experiment. These images include a circular wood grain or lateral-direction wood grain.



(a) Patches correctly classified



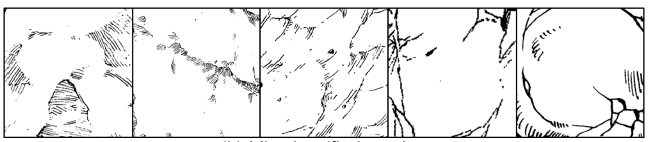
(b) Mis-classified patches

Figure 4: Results of linear discriminant classification for wood.

Rock images

The F-measure value for the rock material category was 0.62 (recall: 0.73, precision: 0.53). The features selected were F10 and F11. Figure 5(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, white areas of a characteristic size appear representing the rock appearance. We found that in line-drawing images the horizontal and vertical white longer run-length are relevant for representing a rock surface. Figure 5(b) shows images mis-recognized by the linear discriminant classification but classified into the rock category by the psychophysical experiment. These patches include larger white areas or areas not closed.

(a) Patches correctly classified

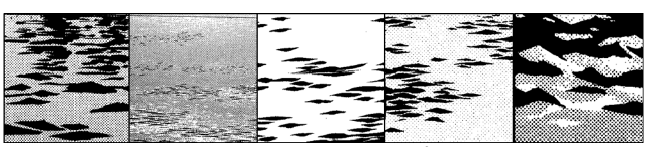


(b) Mis-classified patches

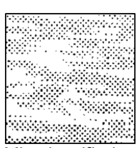
Figure 5: Results of linear discriminant classification for rock

Water images

The F-measure value for the water material category was 0.55 (recall: 0.86, precision: 0.40). The features selected were F8, F9, F10 and F11, the same ones as for the wood material. Figure 6(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, large horizontal black and white areas appear, representing the wave form of the water. We found that in line-drawing images a long run-length in the horizontal direction and a short run-length in the vertical direction are relevant. Figure 6(b) shows an image mis-recognized by the linear discriminant classification but classified into the water category by the psychophysical experiment. This patch includes a dazzling wave pattern by stippling.



(a) Patches correctly classified

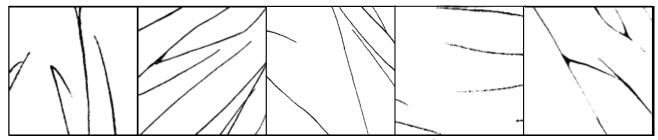


(b) Mis-classified patch

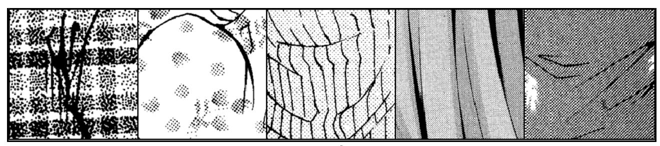
Figure 6: Results of linear discriminant classification for water.

Fabric images

The F-measure value for the metal fabric category was 0.51 (recall: 0.56, precision: 0.46). The features selected were normalized versions of F4 and F5. As shown in Fig. 7(a), wide white areas or thick white lines appear, representing the fabric surface. We found that in line-drawing images the average run-length in the horizontal and vertical directions are relevant. Figure 7(b) shows images mis-recognized by the linear discriminant classification but classified into the fabric category by the psychophysical experiment. In most cases, these erroneously classified patches represented a screen tone or patterns instead of line drawings. A screen tone is the manga technique to represent specific patterns or colors, and it is represented by stippling, thereby decreasing the run-lengths.



(a) Patches correctly classified

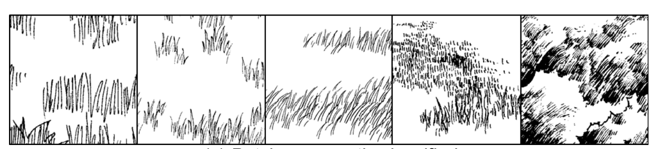


(b) Mis-classified patches

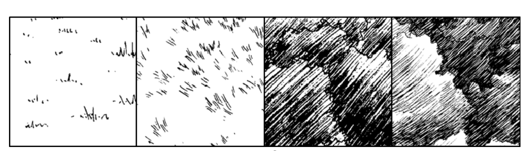
Figure 7: Results of linear discriminant classification for fabric

Foliage images

The F-measure value for the foliage material category was 0.49 (recall: 0.69, precision: 0.38). The features selected were (F2 – F3) and (F6 – F7). Figure 8(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, short black and white lines having almost the same length appear, representing the foliage appearance. We found that in line-drawing images the differences between horizontal and vertical run-lengths and between their respective standard deviations are relevant. Figure 8(b) shows images mis-recognized by the linear discriminant classification but classified into the foliage category by the psychophysical experiment. These patches include detached lines and shorter vertical lines.



(a) Patches correctly classified

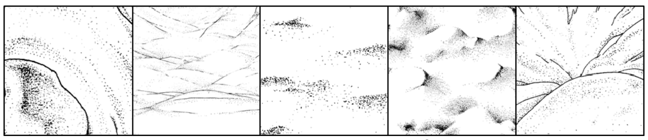


(b) Mis-classified patches

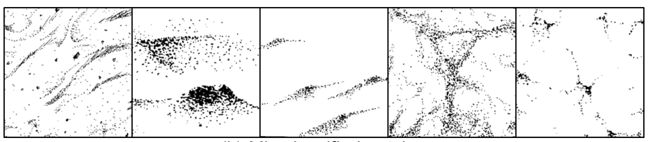
Figure 8: Results of linear discriminant classification for foliage.

Sand images

The F-measure value for the sand material category was 0.41 (recall: 0.55, precision: 0.33). The feature selected was the normalized versions of F14. Figure 9(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, many dots serve to represent the grains of sand. We found that in line-drawing images the number of neighboring black pixels is relevant. Figure 9(b) shows images mis-recognized by the linear discriminant classification but classified into the sand category by the psychophysical experiment. Dots in these patches cohere too much, so that they become large chunks of pixels.



(a) Patches correctly classified

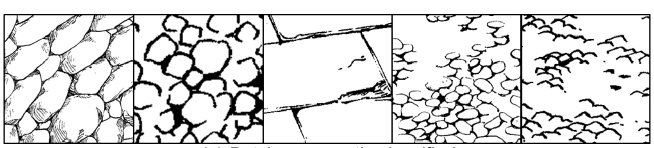


(b) Mis-classified patches

Figure 9: Results of linear discriminant classification for sand

Stone images

The F-measure value for the stone material category was 0.18 (recall: 0.80, precision: 0.10); this was the lowest score of our material categories. The features selected were F6, F7, F10 and F11. Figure 10(a) shows images recognized correctly by the linear discriminant classification. As shown in the figure, characteristic sizes of white areas appear, representing the stone appearance. We found that in line-drawing images the horizontal and vertical white longer run-length are relevant for representing a stone surface. These features represented not only the stone material patches but also many rock material patches; therefore, it was difficult to distinguish stone and rock using the low-dimensional features. Figure 10(b) shows an image mis-recognized by the linear discriminant classification but classified into the stone category by the psychophysical experiment. This patch includes only a few stones, whereas most of the drawings included sets of small stones.



(a) Patches correctly classified

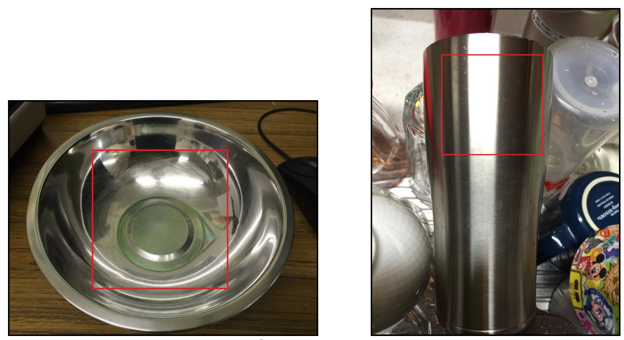
(b) Mis-classified patches

Figure 10: Results of linear discriminant classification for stone

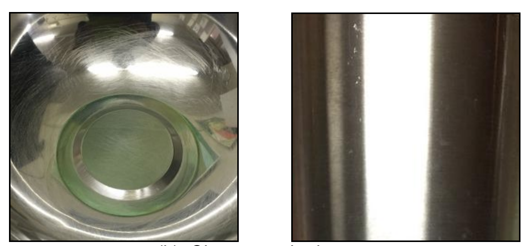
Transfer of Material Appearance from Color to Black-and-White Drawing

The metal category had the highest F-measure in our constructed materials database. As an application, we created black-and-white drawings of metal objects from color images. As described in the previous section, in line-drawing images the ratio of black pixels to total pixels and the thicknesses of the black lines were found to be relevant for representing metal surfaces. By converting a color image into a binary image satisfying the low-dimensional features of the metal category,

Figure 11(a) shows examples of metal object images captured by a digital color camera. Figure 11(b) shows the close-up images in part of red square in Fig.11(a). The resolution of the patch image is 300 × 300 pixels which is close to that for the images in the constructed database. We applied the median filter to the image for noise reduction and then converted it to a binary image by the discriminant percentile method, setting the percentile parameter to 70%. Figure 11(c) shows the converted images. We confirmed that the converted images were classified into the metal category by the linear discriminant function. Please note that this feature condition is a necessary condition but not a sufficient condition; higher-level features may be required to derive a condition that is both necessary and sufficient.



(a) Color input images



(b) Close-up color images

(c) Black-and-white binary images

Figure 11: Color-to-binary conversion of image of metal object

Conclusions

In this study, we constructed a manga materials database containing 274 images in eight material categories, consisting of object and patch images taken from black-and-white line drawings. These images were selected by a psychophysical experiment evoking a consistent material perception. Through the experiment, we confirmed that we could classify even patch images into the appropriate material category with an accuracy of 88.4%.

By an analysis of the database drawings, we found that the material appearance was well represented by certain low-dimensional image features such as the ratio of black pixels to total pixels and the average run-length of horizontal black pixels. As an application, we converted a color image of a metal object to a black-and-white line drawing with metal features.

In this study, we examined all combinations of up to four low-dimensional features, but this number of combinations is not optimum. We plan detailed analysis in these combinations of features. Our evaluation experiment used a large dataset, so there is a possibility that the subject may have a lot of pairwise comparisons [21]. Verification of this effect is required. Also, further investigation is required to represent line drawings of each material category. We also plan to increase the number of material categories and samples in the database.

Acknowledgments

This work was supported by Grant-in-Aid for Scientific Research on Innovative Areas "Shitsukan" (No. 25135706) from MEXT, Japan.

References

[1] Koenderink, J. & van Doorn, A. (1987). Representation of local geometry in the visual system. Biological Cybernetics, 545, 367–375.

[2] Lowe, D. G. (2004). Distinctive image-features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91–110.

[3] Dalal, N. & Triggs, B. (2005). Histograms of oriented gradients for human detection. Proc. Computer Vision and Pattern Recognition, 2, 886–893.

[4] Debevec, P., Hawkins, T., Tchou, C., Duiker, H. P., Sarokin, W. & Sagar, M. (2000). Acquiring the reflectance field of a human face, Proc. ACM SIGGRAPH, 145–156.

[5] Marschner, S., Westin, S. H., Arbree, A. & Moon, J. T. (2005). Measuring and modeling the appearance of finished wood. Proc. ACM SIGGRAPH, 727–734.

[6] Dana, K. J., Van-Ginneken, B., Nayar, S. K. & Koenderink, J. J. (1999). Reflectance and texture of real world surfaces, ACM Trans. on Graphics, 18(1), 1–34. .

[7] Jensen, H. W., Marschner, S., Levoy, M. & Hanrahan, P. (2001). A practical model for subsurface light transport, Proc. ACM SIGGRAPH, 511–518.

[8] Dror, R. O., Adelson, E. H. & Willsky, A. S. (2001). Recognition of surface reflectance properties from a single image under unknown real-world illumination. Proc. the Workshop on Identifying Objects Across Variations in Lighting at Computer Vision and Pattern Recognition.

[9] Abe, T., Okatani, T. & Deguchi, K. (2012). Recognizing surface qualities from natural images based on learning to rank, Proc. International Conference on Pattern Recognition, 3712-3715.

[10] Varma, M. & Zisserman, A. (2009). A statistical approach to material classification using image patch exemplars. IEEE Trans. Pattern Analysis and Machine Intelligence, 31(11), 2032–2047.

[11] Liu, C., Sharan, L., Adelson, E. H., & Rosenholtz, R. (2010). Exploring features in a Bayesian framework for material recognition. Proc. Computer Vision and Pattern Recognition, 239–246.

[12] Sharan, L., Rosenholtz, R. & Adelson, E. (2009). Material perception: What can you see in a brief glance?, Journal of Vision, 9(8), 784.

[13] Hiramatsu, C., Goda, N. & Komatsu, H. (2011). Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. NeuroImage, 57(2), 482–494

[14] Motoyoshi, I., Nishida, S., Sharan, L. & Adelson, E.H. (2007). Image statis-tics and the perception of surface qualities. Nature, 447(7141), 206–209.

[15] Anderson, B. L. (2011). Visual perception of materials and surfaces. Current Biology, 21(24), R978–983.

[16] Thompson, W. B., Fleming, R. W., Creem-Regehr, S., & Stefanucci, J. (2011). Visual perception from a computer graphics perspective. Wellesley, MA, USA: CRC Press.

[17] Zaidi, Q. (2011). Visual inferences of material changes: Color as clue and distraction. Wiley Interdisciplinary Reviews: Cognitive Science, 2(6), 686–700.

[18] Fleming, R. W., Wiebel, C., & Gegenfurtner, K. (2013). Perceptual qualities and material classes. Journal of Vision, 13(8), 9. http://dx.doi.org/10.1167/13.8.9.

[19] Tanaka, M. & Horiuchi, T. (2014). Investigating perceptual qualities of static surface appearance using real materials and displayed images. Vision Research, doi:10.1016/j.visres.2014.11.016.

[20] Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9, 62–66.

[21] Clarke, A. D. F., Xinghui, D. and Chantler, M. J. (2012). Does free-sorting provide a good estimate of visual similarity? Predicting Perceptions. Proceedings of the 3rd International Conference on Appearance. Lulu Press, Edinburgh UK, 17-20.

