This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2880983, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Estimation of Perceptual Surface Property Using Deep Networks with Attention Models

HYUNJOONG CHO[1], YE SEUL BAEK[2], YOUNGSHIN KWAK[2], AND SEUNGJOON YANG[1], (MEMBER, IEEE)

[1]School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Korea (e-mail: hyeunjoong@unist.ac.kr, syang@unist.ac.kr)

[2]School of Design and Human Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Korea (e-mail: iole1004@naver.com, yskwak@unist.ac.kr)

Corresponding author: Seungjoon Yang (e-mail: syang@unist.ac.kr).

This work was supported by National Research Foundation of Korea under Grant NRF-2016R1D1A1B01016041 and Ulsan National Institute of Science and Technology Outstanding Basic Research Fund under Grant 1.180051.01.

ABSTRACT How we perceive property of surfaces with distinct geometry and reflectance under various illumination conditions is not fully understood. One widely studied approach to understanding perceptual surface property is to derive statistics from images of surfaces with the goal of constructing models that can estimate surface property attributes. This work presents machine learning-based methods to estimate the lightness and glossiness of surfaces. Instead of deriving image statistics and building estimation models on top of them, we use deep networks to estimate the perceptual surface property directly from surface images. We adopt the attention models in our networks, to allow the networks to estimate the surface property based on features in certain parts of images. This approach can rule out image variations due to geometry, reflectance, and illumination when making the estimations. The networks are trained with perceptual lightness and glossiness data obtained from psychophysical experiments. The trained deep networks provide accurate estimations of surface property that correlate well with human perception. The network performances are compared with various image statistics derived for estimation of perceptual surface property.

INDEX TERMS Appearance model, neural network, perceptual surface property.

I. INTRODUCTION

In many industries, measuring and controlling the visual appearance of product surfaces is important to their success. For example, the color and gloss of product surfaces are carefully designed and manufactured to provide a certain appearance. Moreover, product surfaces are often designed to provide different appearances depending on illumination conditions. For these businesses, it is critical to understand how humans perceive surface appearance and how various conditions, such as surface geometry, the reflectance of surface materials, or the direction of illumination, affect human perception. However, the mechanism by which humans perceive surface property is not yet fully understood; how the physical and optical property of surface correspond to human visual perception is not clearly known. Nevertheless, even from a single surface image, human observers can easily judge surface qualities such as color and the degree of lightness or glossiness. One widely used approach to explain visual perceptions of surface property is to compute image statistics from a surface [1]–[9]. If we can find image statistics that predict the surface appearance appropriately, those results can be used to develop visual appearance measurement devices for the materials industry or to develop manufacturing methodologies to assess visual appearance.

The appearance of a surface depends on geometry, reflectance, and illumination. When surfaces with identical geometrical shapes are illuminated by a fixed lighting condition, their surface appearances depend solely on reflectance. However, when surfaces with different shapes are illuminated under different lighting conditions, their surface appearances are affected by a complicated interplay among geometry, reflectance, and illumination. Surface property can be mea-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2880983, IEEE Access

sured by various methods. For example, glossiness can be measured using a method suggested by the American society of testing and materials (ASTM1999523) [10] or by measuring the bi-directional reflectance distribution (BRDF). However, how these measurement data are related to the characterization of human perception is an open subject [3]. Image statistics computed from a surface image that describe the relationship of geometry, reflectance, and illumination to human perceptions of surface property have been studied. In [11], human perceptions of stucco-like materials under different illumination was studied. The skewness of luminance histograms was proposed as an image statistic that influences surface gloss and albedo. In [12], the perception of the glossiness or roughness of a sphere under different illuminations was studied. Inspired by studies of texture analysis and synthesis, this study used statistics in the wavelet domain and adopted the 10th percentile of luminance pixel values and the variance of a high wavelet subband as statistics to describe surface appearances. In [13], the perception of glossiness of concave and convex surfaces was studied. The human perception of glossiness is affected by the surface geometry and the locations of specular reflection. In [14], the perception of qualitative shapes of matte and glossy surfaces under different lighting directions was studied. Glossy surface perception improves when the surface is viewed at high slant angles. The study of image statistics is intended to improve scientific understanding of human perceptions of appearance attributes and to derive soft metrology that correlates with human responses. Statistics derived from images have often been correlated to specific appearance attributes, but identifying image statistics suitable for building a model that accurately describes the complex perceptions of human observers is difficult.

Machine learning techniques have been used to estimate image qualities that reflect human perceptions [15]–[17]. The goal of these approaches is to develop measures that quantify the degradation of perceptual image property under various adverse conditions such as additive noise, compression, loss of data, etc. In [18], a convolutional neural network (CNN) was constructed to determine the relative lightness between patches of images. The network simply determines whether one given patch is brighter or darker than another. The trained network was used to evaluate the relative lightness of image regions and to assign lightness values to segment an image.

In this study, we used a machine learning technique to estimate how humans perceive surface property under various conditions. We built deep networks to estimate a human perception of surface property, which is affected by surface geometry, reflectance, and illumination. Samples of surface images were prepared by rendering 3D surface of various geometry and reflectance under various illumination conditions. Human perceptions of surface lightness and glossiness were obtained through psychophysical experiments. The deep networks consist of multiple blocks of convolutional layers, an attention model, and fully connected layers. The purpose of the attention model is to direct the attention of the deep networks to certain parts of the images. By directing their attention, the networks can estimate the surface property while ignoring variations in pixel values that occur due to the geometry, reflectance, and illumination. We compare the behaviors of these trained deep networks with estimation models that use specific image statistics derived for estimation of perceptual surface property. The experimental results show that the deep networks can predict the perceptual lightness and glossiness more accurately than do the specific image statistics.

II. PERCEPTUAL SURFACE PROPERTY

A. SAMPLE PREPARATION

We prepared images with varying statistics by rendering images of 3D surfaces with various geometry and reflectance property under various illumination locations. Fig. 1 shows some examples of the 3D surfaces used for rendering the test images. Eight 256 × 256-pixel patches were selected from the 3D surface: the bunny's thigh, which is highly convex, the elephant's ear, which exhibits well-defined structures, the horse's thigh, with a smooth contour, and the complex lines on the top of the skull are examples of the patches. The surface geometry for each patch is modified to a reverse convex surface with various levels of relief. The images are rendered using the Phong model [19] while varying the matte and specular reflection parameters and illumination locations. Table 1 summarizes the parameters used to prepare test images. Other parameters, such as the azimuth angle of a light source or the ambient reflectance parameter of the Phong model, were excluded through a pilot test. For each selected patch, we prepared 81 test images under different reflectance, geometry, and light source location parameters. By using 3D surface geometry data and the reflectance model, we were able to prepare physically meaningful test images with varying image statistics for our experiments. Our test images were not preprocessed by any image processing techniques because those may cause the rendered images to be physically impossible.

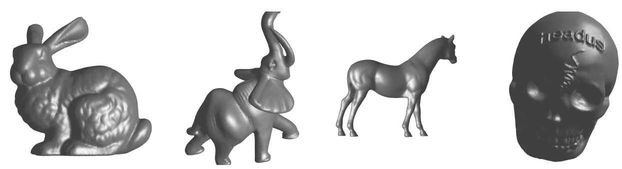


FIGURE 1: 3D surface used to render test image patches.

TABLE 1: Parameters used to prepare test images

parameter		range		
surface	concavity	as is	flipped	
	relief	×0.75	×1.00	×1.25
light source location	altitude	30°	60°	90°
Phong model	diffuse	0.25	0.50	0.75
	specular	0.25	0.50	0.75

Fig. 2 shows some examples of surface images prepared with different geometry and reflectance under different illumination conditions. For a given patch, there are 162 variations of the surface, reflection, and illumination parameters in Table 1. Eight of these 162 variations are shown in Fig. 2, which demonstrates that the perception of surface property is dramatically affected by the parameters used to prepare the test images.

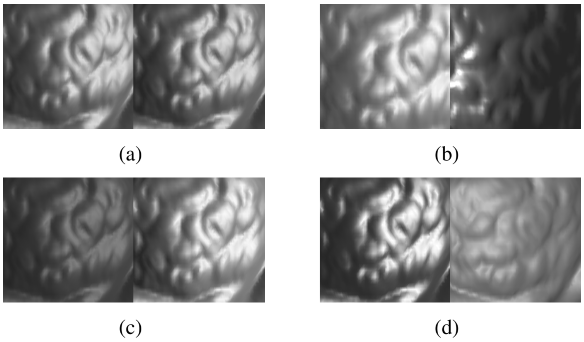


FIGURE 2: Examples of surface images with (a) different relief, (b) different concavity, (c) different reflectance, and (d) different illumination conditions.

There have been research on estimating parameters of a reflectance model from images. For example, in [20], the parameters of the Phong model were estimated from images of a cylinder and from images of general surface shapes. The estimation of model parameters provides information on the reflectance of surfaces. However, it does not provide information about how the human perception of surface changes through the intricate interplay of surface geometry, reflectance, and illuminations. Images of a surface with the same reflectance or the same Phong model parameters can look quite different depending on the surface geometry or illumination conditions. Our study is intended to estimate the human perception of surface property, not to estimate the parameters of a shading model from an image.

B. PSYCHOPHYSICAL EXPERIMENT

In this study, we obtained the lightness and glossiness of surface images as perceived by human observers through psychophysical experiments. We used a 24-inch EIZO-CG42W LCD monitor set up in a dark room in our experiments with a resolution of 1920 x 1080 pixels and whose full-white luminance is 250 cd/m^2. The display's color gamut is sRGB, and the white point was set to D65. The test images were displayed in the middle of the display along with the reference images. The background was set to gray at a luminance of 105 cd/m^2. For the visual lightness evaluation, the darkest and brightest images among our test images were displayed as references. For the visual glossiness evaluation, the images with the highest and lowest gloss were displayed as references.

Eighteen human observers, nine males and nine females, participated in the experiments. The observers were tested for normal color vision by the Ishihara test for color blindness. All the observers were university students. A 5-point Likert scale was adopted for the subjects to evaluate the perceived glossiness and lightness. The test images were shown in random order. The total number of test images shown to an observer in the experiments was 648; however, including the repetitions of the 27 reference images, the total number of observations was 1350.

From the psychophysical experiments, we collected 648 images labeled with the perceptual lightness and glossiness data. These labeled data were used to train deep networks for the estimation of perceptual surface property.

III. DEEP NETWORK BASED ESTIMATION OF PERCEPTUAL SURFACE PROPERTY

A. NETWORK ARCHITECTURE

Fig. 3 shows a schematic of a deep network used for the estimation of perceptual surface property. The network is built up with cascades of five convolutional blocks, each of which consists of two convolutional layers with 3×3 filters with rectified linear unit (ReLU) activation functions and a pooling layer using max pooling:

$$\begin{aligned} u_{l1} &= \mathrm{ReLU}(\mathrm{Conv}(x_l)) \\ u_{l2} &= \mathrm{ReLU}(\mathrm{Conv}(u_{l1})) \\ x_{l+1} &= \mathrm{MaxPool}(u_{l2}) \end{aligned} \tag{1}$$

for $l = 1, 2, \cdots, 5$, where x_l and x_{l+1} are the input and output of the lth layer, respectively. The blocks contain 64, 128, 512, 512, and 512 nodes. The first two blocks consist of the convolutional layers of VGG [21]. In addition, the attention model was prepared by

$$a = \mathrm{Sigmoid}(\mathrm{Conv}(x_2)) \tag{2}$$

using 1×1 filters. The input to the fourth block was modified by

$$x_3 \leftarrow x_3 \otimes a, \tag{3}$$

where $\otimes$ is a pixel-wise multiplication operation. The value of the attention a is between zero and one. The attention is responsible for activating the features in the part of an image to which the network should pay attention.

The five blocks of the convolutional layers with attention are followed by a last block consisting of three fully connected layers with two ReLU activation functions and the final softmax activation function:

$$\begin{aligned} v_1 &= \mathrm{ReLU}(\mathrm{FullConn}(x_5)) \\ v_2 &= \mathrm{ReLU}(\mathrm{FullConn}(v_2)) \\ y &= \mathrm{SoftMax}(v_2). \end{aligned} \tag{4}$$

The last fully connected layer classifies a given image into the 5-point Likert scale.

We prepared two networks: one to estimate perceptual lightness and the other to estimate perceptual glossiness. We also constructed the same networks but without the attention models for comparison purposes.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2880983, IEEE Access

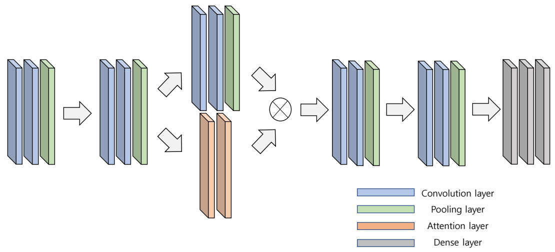
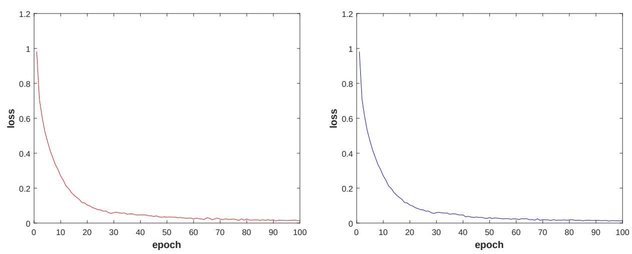


FIGURE 3: Schematics of network architecture

B. TRAINING

The networks were trained with two-thirds of the labeled data obtained through psychophysical experiments. The remaining one-third of the label data was reserved for testing. Because the psychophysical experiments involved human observers, it was difficult to collect vast amounts of labeled data. Consequently, we augmented the data as follows. Four 256×256 images from selected 3D surface with identical surface geometry, reflectance, and illumination parameters were prepared. One of the four images was used in the psychophysical experiments to evaluate the perceptual surface property. From each of the 256×256 images, we cropped 192×192 patches at random locations. Thus, all the patches of these four images share the same score from the psychophysical experiments. To augment the data, we also used rotated versions and mirror images of the patches and assigned the same score. The numbers of patches cropped from the images were controlled such that the numbers of samples with similar Likert scales were balanced in the training set. For the lightness and glossiness models, the total numbers of training samples were 21,656 and 20,410, respectively. The samples in the training and test sets are from separate sets of test images. We used ADAM [22] as the training method, initially setting the learning rate to 0.0001 and slowly reducing it. Fig. 4 shows the loss during training.



(a) (b)

FIGURE 4: Training the networks for the estimation of (a) lightness and (b) glossiness.

IV. PERFORMANCE ANALYSIS

The networks classify input surface images into 5-point Likert scales. We used ten patches cropped from a given test image and computed the average of the resulting 5-point Likert-scale scores as the estimation of the surface qualities. The scores estimated by the networks were compared to the average scores given by human observers. The networks were tested using the one-third of labeled data obtained from the psychophysical experiments that were not included in the training set. Table 2 shows the performances of the networks, whose accuracy, with or without the attention models, are 83.33% and 81.48% for lightness, respectively, and 80.09% and 76.39% for glossiness, respectively. The networks that include the attention models yield more accurate estimations of the lightness and glossiness. The mean square error (mse) between the estimations and the human scores were 0.081 and 0.097 for perceptual lightness and glossiness, respectively. The correlation coefficients between the estimation by the networks and the average scores by human observers were 0.954 and 0.962 for perceptual lightness and glossiness, respectively. These results suggest that the deep networks were properly trained and were able to estimate the perceptual lightness and glossiness accurately, relative to humans.

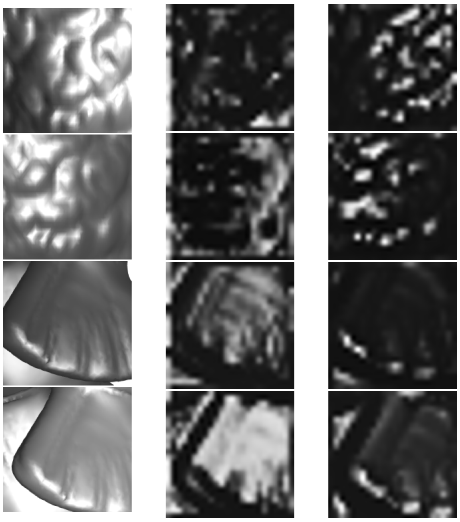
TABLE 2: Performances of the Deep Networks

	lightness		glossiness	
attention model	no	yes	no	yes
accuracy	81.48%	**83.33%**	76.39%	**80.09%**
mse	0.134	**0.081**	0.176	**0.097**
correlation	0.931	**0.954**	0.923	**0.962**

The results shown in Table 2 suggest that the inclusion of the attention models improves the performance of the networks when estimating perceptual surface property. The attention models allow the networks to make estimates of lightness or glossiness based on the features at particular regions of a given image. The networks learned to pay attention to specific parts of a given image. Fig. 5 shows examples of the attention capabilities of the networks. The attention a was used to estimate the lightness and glossiness shown in (b) and (c), respectively, for the input images in (a). Because the attention is multiplied in the features in the third block of the convolutional layers, the networks base their estimations only on the features where the attention is high (i.e., the attention value is close to one). The attention for lightness estimation is high when the surface geometry varies smoothly and low where the pixels are unusually bright, where specular reflection occurs, or unusually dim, where shadows are present due to the geometry. The attention for glossiness estimation is usually high where specular reflection occurs. The way the attention models work in our networks is quite similar to the attention of human observers when estimating surface property. For example, a human would certainly want to avoid being influenced by variations of pixel values due to the surface geometry, surface reflections or illumination conditions when estimating lightness and glossiness. Consequently, a human would pay attention to the parts of a given image conducive making robust and sound judgements.

Table 3 shows the correlation between the average scores from the psychophysical experiments and estimation by various methods. The image statistics proposed by [12] and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2880983, IEEE Access



(a) (b) (c)

FIGURE 5: Examples of the attention capabilities of the networks. (a) Source images, and network attention for (b) lightness; and (c) glossiness.

[11] are included for comparison. The estimations by the proposed deep networks result in the highest correlations to the average scores from the psychophysical experiments. The networks estimate surface property more accurately relative to humans than do the simple image statistics. The skewness proposed in [11] results in a low correlation to the human perceptual experiment scores. A possible explanation for this low correlation is that our images are rendered using a reflection model. Any skewness in the luminance histogram is possibly due to physical changes to the rendered image. In the experiments by [11], skewness was intentionally introduced by modifying images using an image-processing technique.

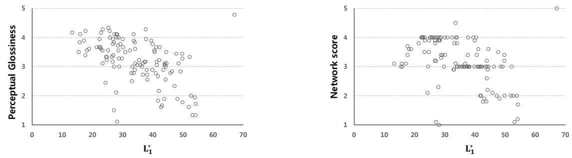
TABLE 3: Correlations Between Perceptual Surface Property and Estimation Methods

	estimation methods	R
perceptual	10th percentile of pixel value [12]	0.774
	proposed deep network	**0.954**
perceptual glossiness	variance of a high wavelet subband [12]	0.867
	skewness [11]	0.292
	proposed deep network	**0.962**

It was reported in [23] that perceptual glossiness is a function of surface lightness: darker surfaces appear glossier. For analysis, we use the bimodal Gaussian mixture model to estimate the surface lightness. The model is given by

$$p(x) = \pi_1 N(\mu_1, \sigma_1^2) + \pi_2 N(\mu_2, \sigma_2^2), \tag{5}$$

where π_1 and π_2 are mixture probabilities, and $N(\mu, \sigma^2)$ is the Gaussian distribution with a mean of μ and a variance of σ^2. Without loss of generality, we assume $\mu_1 < \mu_2$. Given an image of surface, the bimodal Gaussian mixture model is fitted to the luminance histogram using the expectation maximization (EM) algorithm [24]. The luminance of the lower mode, , L_1^*, is computed and used as the surface lightness. Fig. 6 (a) shows that this appearance attribute of darker surface appearing glossier was observed in our psychophysical experiment data. Fig. 6 (b) shows the estimation of the perceptual glossiness by the deep network vs. the luminance of the lower mode L_1^*. It can be seen that the deep network's estimate of perceptual glossiness decreases as the lightness increases. In other words, the deep network mimics the response of human observers by perceiving darker surfaces as glossier.



(a) (b)

FIGURE 6: Relation between perceptual glossiness and lightness. (a) perceptual glossiness vs. L_1^*, and (b) estimation by the deep network vs L_1^*.

The image statistics derived from regression models are useful for interpreting surface appearance in terms of known variables; lightness, variance, 10th percentiles, and so on are statistics that we can intuitively understand and relate to surface property attributes. In contrast, the operations of deep networks and the aspects of the features that deep networks utilize for classification are difficult to interpret intuitively. However, in terms of performance in predicting the visual perception of surface property attributes, the deep networks demonstrate clear advantages over the use of specific image statistics. For applications in which accurate prediction of visual perception is a critical issue, the proposed deep network is suitable for use in assessing and monitoring the perceptual surface property.

Textures and patterns that humans perceive in a surface image can stem either from reflections from the illumination or from the true textures or patterns of the surface. In real-world illumination conditions, the surface image is affected by the distribution of light hitting the surface [25]. Because of these ambiguities, image statistics obtained from real-world surface images or—following the same line of thought—the use of a deep network to estimate surface property may not be suitable when the surfaces and illumination conditions are complicated. However, for manufacturing purposes or for quality control purposes, surfaces and illumination can be controlled, and such ambiguities can be removed. Thus, the proposed deep network is suitable for assessing perceptual surface property under controlled illumination conditions, and further, for replacing difficult and time-consuming psychophysical tests on human observers. We are currently investigating the use of deep networks to estimate surface property in more complicated illumination conditions.

V. CONCLUSION

This work presents two deep networks with attention models to estimate surface property of perceptual lightness and glossiness. The networks learned to pay attention to certain parts of surface images to rule out any pixel values variations in surface images resulting from variations in geometry, reflectance, and illumination. The trained networks can accurately predict the perceptual lightness and glossiness of surfaces from images of the surfaces. The networks show higher correlations to human perceptual data than do various image statistics. The networks can be used to measure perceptual surface property in controlled environments and to simulate psychophysical tests on human observers.

REFERENCES

[1] R. W. Fleming, R. O. Dror, and E. H. Adelson, "Real-world illumination and the perception of surface reflectance properties," Journal of vision, vol. 3, no. 5, pp. 3–3, 2003.
[2] CIE, "International commission on illumination technical report 175:2006, "a framework for the measurement of visual appearance"," in CIE, Vienna, 2006.
[3] C. Eugène, "Measurement of Òtotal visual appearanceÓ: a cie challenge of soft metrology," in 12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement, 2008, pp. 61–65.
[4] J. Kim and B. L. Anderson, "Image statistics and the perception of surface gloss and lightness," Journal of Vision, vol. 10, no. 9, pp. 3–3, 2010.
[5] M. Olkkonen and D. H. Brainard, "Perceived glossiness and lightness under real-world illumination," Journal of vision, vol. 10, no. 9, pp. 5–5, 2010.
[6] ——, "Joint effects of illumination geometry and object shape in the perception of surface reflectance," i-Perception, vol. 2, no. 9, pp. 1014–1034, 2011.
[7] P. J. Marlow, J. Kim, and B. L. Anderson, "The perception and misperception of specular surface reflectance," Current Biology, vol. 22, no. 20, pp. 1909–1913, 2012.
[8] F. Leloup and P. Hanselaer, "Progress in the soft metrology of appearance: the contribution of digital image representations," in SID Display Week'14 International Symposium. SID, 2014, pp. 603–606.
[9] R. W. Fleming, "Material perception," Annual review of vision science, vol. 3, pp. 365–388, 2017.
[10] D. ASTM, "523. standard test method for specular gloss," American Society for Testing and Materials, Philadelphia, 1999.
[11] 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.
[12] E. H. Adelson, "Image statistics and surface perception," in Electronic Imaging 2008. International Society for Optics and Photonics, 2008, pp. 680 602–680 602.
[13] I. S. Kerrigan and W. J. Adams, "Highlights, disparity, and perceived gloss with convex and concave surfaceshighlights, disparity, and perceived gloss," Journal of Vision, vol. 13, no. 1, pp. 9–9, 2013.
[14] A. Faisman and M. S. Langer, "How does lighting direction affect shape perception of glossy and matte surfaces?" in Proceedings of the ACM Symposium on Applied Perception. ACM, 2013, pp. 9–14.
[15] S. Bosse, D. Maniry, T. Wiegand, and W. Samek, "A deep neural network for image quality assessment," in Image Processing (ICIP), 2016 IEEE International Conference on. IEEE, 2016, pp. 3773–3777.
[16] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1733–1740.
[17] H. Tang, N. Joshi, and A. Kapoor, "Blind image quality assessment using semi-supervised rectifier networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2877–2884.
[18] T. Narihira, M. Maire, and S. X. Yu, "Learning lightness from human judgement on relative reflectance," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2965–2973.
[19] B. T. Phong, "Illumination for computer generated pictures," Communications of the ACM, vol. 18, no. 6, pp. 311–317, 1975.
[20] S. Tominaga and N. Tanaka, "Estimating reflection parameters from a single color image," IEEE Computer Graphics and Applications, vol. 20, no. 5, pp. 58–66, 2000.
[21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
[22] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.
[23] J. A. Ferwerda, F. Pellacini, and D. P. Greenberg, "Psychophysically based model of surface gloss perception," in Photonics west 2001-electronic imaging. International Society for Optics and Photonics, 2001, pp. 291–301.
[24] G. McLachlan and D. Peel, "Mixtures of factor analyzers," Finite Mixture Models, pp. 238–256, 2000.
[25] R. O. Dror, E. H. Adelson, and A. S. Willsky, "Surface reflectance estimation and natural illumination statistics," 2001.

HYUNJOONG CHO received a B. S. degree from the Ulsan National Institute of Science and Technology, Ulsan, South Korea in 2015, where he is currently pursuing a master's degree in electrical and computer engineering. His research interests include image processing, human perception, and human recognition.

YE SEOL BAEK received the Ph. D. degrees in color science from Daejin University, Pocheon, South Korea, in 2013. Since 2009, she is working as a postdoctoral researcher at School of Design and Human Engineering, Ulsan National Institute of Science and Technology (UNIST), South Korea. Her main research interests include human color perception, visual appearance, and display quality assessment.

YOUNGSHIN KWAK received the B.Sc. and M.Sc. degrees in physics from Ewha Womans University, Seoul, South Korea, in 1995 and 1997, respectively, and the Ph.D. degree from the Colour & Imaging Institute, University of Derby, U.K., in July 2003. Following her Ph.D. degree, she worked for Samsung Electronics, South Korea. Since 2009, she is working as the assistant professor at School of Design and Human Engineering, Ulsan National Institute of Science and Technology (UNIST), South Korea. Her main research interests include human color perception, color emotion, visual appearance, and image quality of 2D and 3D images.

SEUNGJOON YANG (S'99-M'00) received a B.S. degree from Seoul National University, Seoul, Korea in 1990 and M.S. and Ph.D. degrees from the University of Wisconsin-Madison in 1993 and 2000, respectively, all in electrical engineering. He worked at the Digital Media R&D Center at Samsung Electronics Co., Ltd. from September 2000 to August 2008 and is currently with the School of Electrical and Computer Engineering at the Ulsan National Institute of Science and Technology in Ulsan, Korea. His research interests include image processing, estimation theory, and multi-rate systems. Professor Yang received the Samsung Award for the Best Technology Achievement of the Year in 2008 for his work on the premium digital television platform project.