https://doi.org/10.2352/ISSN.2470-1173.2017.14.HVEI-143

Determining the Influence of Image-based Cues on Human Skin Gloss Perception

Jing Wang, Thrasyvoulos N. Pappas; Northwestern University; Evanston, IL/US
Jim Mayne, Carla Kuesten, Gopa Majmudar; Amway; Ada, MI/US

Abstract

We explore the influence of surface and subsurface reflections on skin gloss perception. We rely on multimodal photography to separate the surface and subsurface reflection images. Since the original data consists of a limited number of images (25 subjects, front and side view, before and after skin cleansing), we apply different transformations to surface and subsurface reflection images, in order to generate a broad range of appearance of skin images. We conducted two empirical studies with the expanded set of data, at both the macro-scale level (whole face) and the meso-scale level (local skin patch). We found that increasing the contrast of surface reflection results in higher gloss perception, while a decrease in the amount of subsurface reflection (lower average lightness, darker complexion) results in higher gloss perception; however, the differential effect of subsurface reflection on gloss diminishes as the average lightness becomes very low. We also computed the statistics of the two reflection images and found their effects (sometimes opposite for the corresponding statistic) on gloss perception. We then learned a regression model based on the concatenation of statistics from the surface and subsurface reflection images to predict relative gloss differences. Our results indicate that using the statistics from both modalities provides more consistent correlation with human judgments than using only the statistics from a single modality.

Introduction

Gloss is an important attribute of visual texture perception, and the visual appearance of human skin, in particular. The objective and subjective evaluation of human skin gloss is important for cosmetology and dermatology. Too much facial skin gloss indicates oily and sweaty skin and too little gloss results in a dull and unhealthy appearance, while the right amount of gloss generates a youthful and vibrant impression. Multiple cosmetics and health care products have been developed to improve the skin condition for a just "perfect" gloss appearance. Thus, techniques for effective measurement of skin gloss are valuable for the evaluation of skin consultation, marketing, product claims, etc.

Existing techniques for quantitative measurement of skin gloss are mostly at the physical level using optical instruments like the SkinGlossMeter.[1] Such instruments collect data by pressing a probe at different points of the skin and recording the reflected light. However, the collected data are not representative of the overall skin condition, and more importantly, do not correlate well with visual gloss.

In contrast, the goal of this paper is to develop image-based metrics of perceived gloss that are consistent with human judgments. The appearance of surface gloss depends on multiple extrinsic and intrinsic factors including the illumination source, the viewing conditions, the surface geometry, and the material optical reflectance and transmittance [1]. For human skin, the intrinsic properties are further complicated by the multi-layer structure of the skin. The air-oil layer, the outer layer of the epidermis, is translucent and only partially reflects incident light while the rest of the light is transmitted through, scattered, or absorbed in the inner skin layers. The scattered component exits the skin as reflected light in random directions. Thus, the reflection from human skin consists of a surface reflection and a subsurface reflection component. Understanding the distinct effects of these two components is important in the study of gloss perception.

To separate the surface reflection and subsurface reflection, we utilize multimodal polarized photography to capture high resolution images of the face. With appropriate placement of polarizers in front of the light source and the camera (parallel or perpendicular to each other), we can separate the surface reflection from the subsurface reflection by subtracting cross-polarized image (subsurface only) from the parallel-polarized image (surface and subsurface).

To enrich the (typically limited) amount of data obtained by direct multimodal photography, we apply an S-transformation [2] to the extracted surface reflection image and a λ-transformation [3, 2] to the subsurface reflection image. By manipulating the two reflections, we are able to generate multiple facial skin images with gradually varying appearance, which can be used as stimuli in our empirical studies. A comparison of the image statistics and the subjective evaluations can then be used to derive image-based metrics of perceived gloss.

We designed two empirical studies to uncover and separate the influence of the surface and subsurface reflections on skin gloss perception. We obtained images of facial skin at two levels: macro-scale (whole face) and meso-scale (skin patch). Our results show that the contrast of the surface reflection has a strong positive influence on skin gloss perception. Keeping the surface reflection constant, a darker skin tone in the subsurface reflection tends to result in glossier appearance. However, the differential effect of skin tone on perceived gloss diminishes when the average lightness is low.

We then compared the statistics of the surface and subsurface reflection images to the results of the subjective evaluations. We found that the statistics of both the surface and subsurface reflection have a strong effect on gloss perception, even though, in some cases, the corresponding statistics (average lightness, skewness, and cluster shade) of the two reflection images have opposite

[1] http://www.delfintech.com/en/product_information/skinglossmeter/

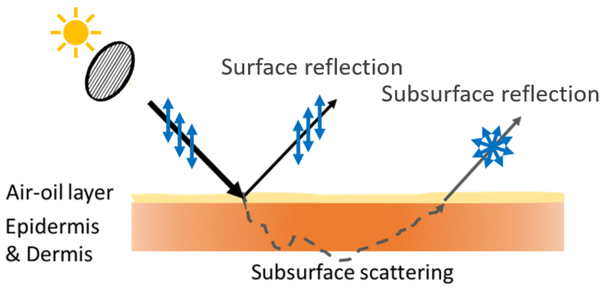


Figure 1: Propagation of polarized light via the surface and subsurface of the human skin.

(a) PP image (b) XP image (c) Surface reflection (d) Subsurface reflection

Figure 2: An example of polarized multimodal imaging and the surface and subsurface reflection images. The surface reflection (c) is extracted as PP lightness - XP lightness. The subsurface reflection (d) is obtained as the XP lightness.

effects on gloss perception. We then learned a regression model based on the concatenation of statistics from the two reflection images to predict relative gloss differences. The model performance shows that it correlates better with human judgments than the model that uses the same statistics from a single modality (overall reflection).

In summary, the objectives of this paper are:

- To investigate the influence (separate and combined) of surface reflection and subsurface reflection on gloss perception of human skin images in both macro-scale level and meso-scale level.
- To correlate statistical features of both surface and subsurface reflection images with gloss perception.
- To quantitatively predict the relative gloss difference of human skin before and after cleansing.

Previous work on gloss perception

In the past decades, studies on visual gloss of 2D images have been focusing on the relationship between image-based cues and gloss perception. Motoyoshi *et al.* [4] observed that darker and glossier appearance of a surface tends to have a positively skewed luminance histogram. Their skewness hypothesis was later challenged by several studies [5, 6], which found that the correlation between skewness and gloss perception only exists with certain lightness and texture surface restrictions.

To find additional image-based cues, many studies utilized computer generated surfaces under controlled settings for psychophysical tests. Marlow *et al.* [7] found that the specular sharpness, highlight coverage, and specular contrast correlate closely with gloss perception. Through the analysis of surface geometry, Ho *et al.* [8] found that increasing the stretch of surface relief height can change the surface gloss appearance. The relations among geometrical height relief, skewness, and gloss perception were further studied for Lambertian surfaces [9]. They found that, in near-frontal illumination, skewness positively correlates with the surface relief stretch of Lambertian surfaces, which mediates visual gloss; however, this does not extend to oblique illumination. Thus, the skewness hypothesis does not lead to a complete explanation of visual gloss ratings. Lambertian surfaces were also used for the study of the relation between roughness and gloss [10] at different spatial scales. They found complex non-linear interactions in the effects of two roughness parameters on visual gloss. Most of the work we discussed in this paragraph, relies on sets of simple synthetic objects, or surfaces rendered with reflection models in computer graphics. However, the image-based cues they considered and their effects on gloss perception do not necessarily extend to real material images [2] as the appearance of real world materials involves more complicated optical properties and physical processes.

Compared to computer-rendered images, the study of images of real materials is hindered by the difficulty of incrementally varying image appearance. Although Motoyoshi *et al.* [4] used real world images in their study, the stimuli and viewing conditions were quite constrained. A wider range of natural surface images were investigated by Wang *et al.* [2] and the recent work of Wiebel *et al.* [11]. To obtain a variety of controlled variations in image appearance, Wang *et al.* [2] proposed the use of S-curve and λ-curve [9] transformations to manipulate texture lightness. Similarly, Wiebel *et al.* [11] used histogram manipulations to study the relation between contrast and gloss. Both studies found that contrast manipulation has a stronger effect on perceived gloss than skewness. However, they also found that a single statistic like skewness or contrast is not sufficient to predict visual gloss.

For translucent materials, like human skin, we found that the statistical features of a single modality (overall reflection) cannot adequately account for gloss perception. Thus, in our study of human skin, we investigated the influence of the statistics of images that are generated from different skin layers on the overall perception of gloss.

Method

Skin optics and multimodal imaging

The motivation for utilizing statistics of multiple images to predict perceived skin gloss originates from the special geometric and optical properties of the layers of the human skin.

Skin is composed of two main components, the epidermis and dermis. The epidermis generates an air-oil layer that consists of a mixture of sebum, lipids, and sweat, and covers the skin. When incident light reaches the skin, it is either reflected by the air-oil layer or propagates into the subsurface of the skin tissue. We refer to the light that is reflected by the air-oil layer as *surface reflection.* As shown in Figure 1, when the incident light is polarized, the surface reflection is polarized in the same direction as the incident light. Due to the scattering, the light that emerges from the skin tissue is undirectional; we refer to it as *subsurface reflection.*

To separate surface reflection and subsurface reflection, we utilize multimodal photography to capture high resolution facial

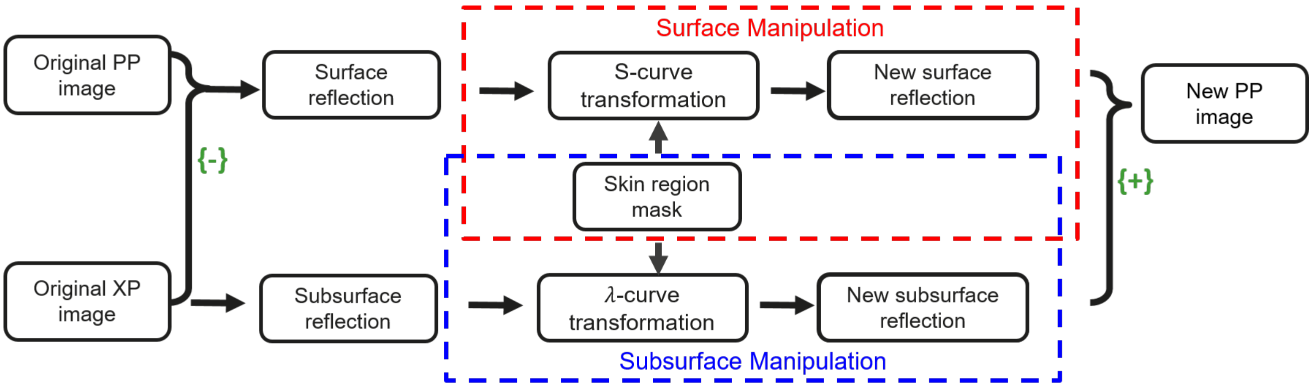


Figure 3: Surface reflection (SurfRefl) manipulation and subsurface reflection (SubSurfRefl) manipulation


(a) (b) (c-1) (c-2) (c-3) (c) Surface reflection (SurfRefl) manipulation (d-1) (d-2) (d-3) (d) Subsurface reflection (SubSurfRefl) manipulation

Figure 4: An example of surface reflection and subsurface reflection manipulation: (a): Original image; (b) Skin region mask; (c-1): Original surface reflection; (c-2): S-curve transformed surface reflection; (c-3): Image with new surface reflection. (d-1): Original subsurface reflection; (d-2): λ-curve transformed subsurface reflection; (d-3): Image with new subsurface reflection.

images using a facial imaging system called VISIA-CR,[2] manufactured by Canfield Scientific Inc. (Parsippany, NJ, US). The VISIA-CR system is equipped with multiple filters to simulate different lighting modalities. In our study we used two lighting modes: parallel-polarized (PP) and cross-polarized (XP). The system places a polarizing filter in front of the light source, and another polarizing filter in front of the digital camera. The angle of the camera filter varies from 0° (PP) to 90° (XP) relative to the polarization angle of the light filter. An example of a PP image and an XP image of one subject is illustrated in Figures 2(a) and 2(b).

As mentioned above, the subsurface reflection is not polarized regardless of the polarization of the incident light. Therefore it exists in both the PP and XP modes. On the other hand, the surface reflection is polarized in the same direction as the incident light. When the polarizer is placed perpendicular to the polarization plane of the light source, all the surface reflection is blocked out. Since the polarization of the subsurface reflection is equally distributed in all directions, the component that corresponds to the subsurface component remains the same in both modes. Thus, the surface reflection can be obtained as the difference between the PP and XP modes (PP − XP), while the subsurface reflection is obtained directly from XP. Figures 2(c) and 2(d) show the separated surface reflection and subsurface reflection components.

We collected PP and XP full-face images from 25 subjects with self-perceived oily skin. Images were obtained from two viewing angles (front view and side view) and two skin conditions (before and after skin cleansing) for each subject, for a total of 100 full-face images. The image resolution is 4270×3612 pixels. We also selected multiple patches from different regions of full-face (macro-scale) images to obtain sets of meso-scale images.

[2] http://www.canfieldsci.com/imaging-systems/visia-cr/

Surface and subsurface manipulation

The 100 natural skin images we collected are insufficient to study gloss related image-based cues since they have very limited variations (two skin conditions, two viewing angles). To increase size of the database and to enrich the appearance range, we used two transformations to independently manipulate the extracted surface and subsurface reflection images. The processing steps are shown in Figure 3.

The two transformations we used are the S-curve transformation [2] and the λ-curve transformation [3]. The two transformations were originally proposed to control the histogram of an image, which has an effect on both the appearance and the statistics of the resulting image. As shown in Equation (1) and (2), I_{in} and I_{out} are the input and output luminance intensity values. Both curves have a controlling parameter (S or λ) that monotonically changes the curve shape and its effect.

$$\lambda\text{-curve: } I_{out} = \sqrt{\frac{I_{in}^2}{I_{in}^2 + \lambda^2(1 - I_{in}^2)}} \qquad (1)$$

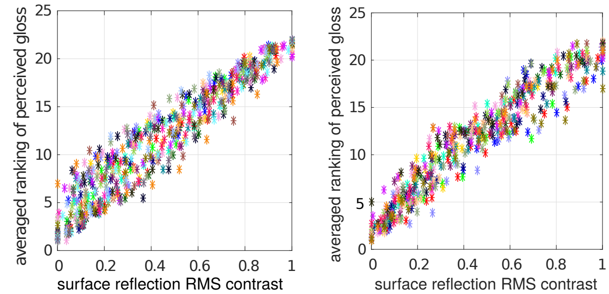
$$\text{S-curve: } I_{out} = \mu - \frac{\mu - I_{in}}{\sqrt{\alpha^2(\mu - I_{in})^2(1 - 1/S^2) + 1/S^2}}, \qquad (2)$$

where $\alpha = \begin{cases} \frac{1}{1-\mu}, & \text{if } I_{in} > \mu \\ \frac{1}{\mu}, & \text{if } I_{in} \le \mu \end{cases}$ and μ denotes the original mean.

In our framework, we used the transformations to independently modify only the luminance component of the surface and

Table 1: Visual Stimuli for Empirical Study 1

		Macro-scale	Meso-scale
Facial skin subjects		25	12
Conditions	Lighting	PP	PP
	Viewing	Front / Side	Front / Side
	Cleansing	Before / After	Before / After
Each condition: 11 images (1 orig. + 10 modified SurfRefl)			
Each trial: 22 images (Before + After cleansing images of ONE skin subject in ONE view)			



(a) macro-scale (b) meso-scale

Figure 5: RMS contrast of surface reflection vs. gloss perception

Table 2: Visual Stimuli for Empirical Study 2

		Macro-scale	Meso-scale
Facial skin subjects		6	48
Conditions	Lighting	PP	PP
	Viewing	Side	Side
	Cleansing	Before	Before
Each condition: 9 images (1 orig. + 8 modified SubSurfRefl)			
Each trial: 2 images of ONE skin subject in ONE view			

subsurface images. Figures 4(c-2) and 4(d-2) show the results of the transformations on the surface and subsurface luminance, respectively. To avoid awkward looking artifacts, we applied the transformations only to skin regions. For this we used a binary mask (Figure 4(b)) that was automatically generated based on an off-the-shelf face landmark detection [12] and color distribution algorithm [13]. After the manipulations, the two reflection images were added up to obtain a new lightness, which was combined with the unmodified color component, as illustrated in Figures 4(c-3) and 4(d-3). We applied the same technique to meso-scale facial skin images to generate multiple variations in appearance.

Empirical Studies

Using different subsets of the expanded set of data (original and modified, at meso-scale and macro-scale), we conducted two empirical studies to determine the influence of the surface and subsurface reflection images on skin gloss perception.

Empirical Study 1: Gloss perception versus surface reflection

Experiment Setup

Visual Stimuli The visual stimuli for Study 1 consisted of original and modified facial skin images with varying surface reflections but fixed subsurface reflection. As shown in Table 1, the stimuli include both macro-scale and meso-scale images. At the macro-scale level, we used original and modified PP images of the 25 subjects, in the two viewing conditions, before and after cleansing. At the meso-scale level, we selected images of 12 skin patches, in the two viewing conditions, before and after cleansing. Each original image was accompanied by 10 modified images with varying surface reflections. Thus, for each subject and each condition, there are 11 images. A complete set of images for one subject and one patch in the study are shown in the Appendix. In total, there were 1,100 macro-scale images and 264 meso-scale images.

Apparatus As we describe below, the participants were asked to compare several images at both the macro-scale and the meso-scale. Due to the limited size of the LCD screen, it is impossible to display multiple macro-scale images on the screen at full resolution. Therefore, the macro-scale images were printed on $5 \times 7''$ photographic paper at 720 pixels per inch. For each trial, 22 photos were pinned on a $40 \times 30''$ canvas.

The tests with the meso-scale images were conducted using a calibrated LCD screen with linear gamma and 1920×1080 resolution. The viewing distance was approximately 600 mm such that a 256-px image subtended an angle of 9.39 degrees.

Procedure In each trial, a participant was shown 22 images at the same time (on the canvas or LCD screen) in random order, and was instructed to re-rank the images in order of increasing visual gloss. The images in each trial came from the same subject in one viewing condition (front or side), before and after cleansing. Skin gloss was explained as "a shiny or radiant appearance of human skin."

Participants There were 10 participants in this study, all female and all with normal vision. Before the test, all participants were asked to read and sign consent forms.

Result Analysis

The goal of this study was to investigate the influence of surface reflection on gloss perception. Figures 5(a) and 5(b) plot the averaged gloss ranking results of each image as a function of the RMS contrast of the surface reflection map. Different colors represent skin images of different subjects. For better visualization, the range of contrast values is normalized across subjects. As expected, visual gloss increases with increasing surface reflection contrast. The influence of surface reflection contrast on visual gloss is similar to the observations in [2], whereby the perceived perceived contrast of real world images had a strongly positive correlation with gloss. As we will see below, in addition to contrast, other statistics of the surface reflection image have an influence on gloss.

Empirical Study 2: Gloss perception versus subsurface reflection

Experiment Setup

Visual Stimuli The visual stimuli for Study 2 consisted of original and modified facial skin images with fixed surface reflection and varying subsurface reflections. As shown in Table 2, the stimuli include both macro-scale images and meso-scale images.

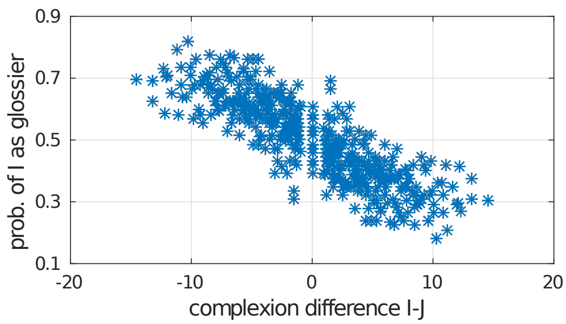
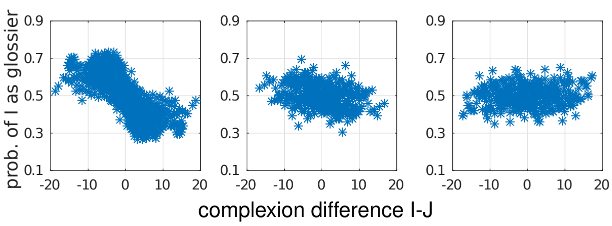
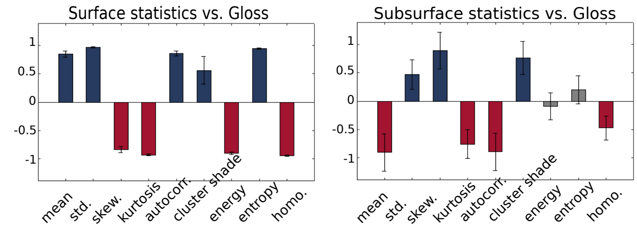


Figure 6: Selection results on macro-scale images.



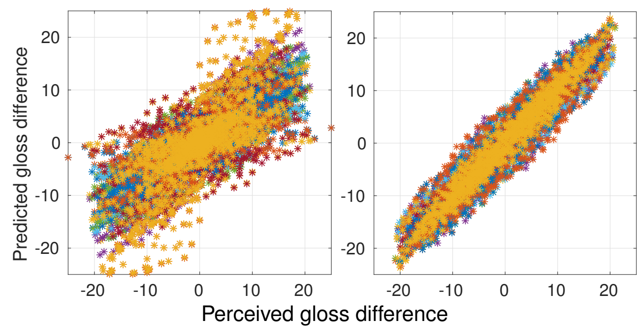
(a) (b) (c)

Figure 7: Selection results on meso-scale images.



(a) surface reflection (b) subsurface reflection

Figure 8: Correlations between gloss perception and statistical features in (a) surface reflection and (b) subsurface reflection



(a) (b)

Figure 9: Prediction of relative gloss difference using statistical features extracted from (a) overall reflection image (b) combined surface and subsurface reflection images.

At the macro-scale, we used original and modified PP images of 6 subjects, in the side view, before cleansing. At the meso-scale, we selected images of 48 skin patches, in the side view, before cleansing. Each original image was accompanied by 8 modified images with varying subsurface reflections. A complete set of images for one subject and one patch in the study are shown in the Appendix. In total, there were 54 macro-scale images and 432 meso-scale images.

Apparatus The tests were all conducted using a calibrated LCD screen with the same settings as Study 1.

Procedure To reduce the risk that participants may confuse the task of gloss perception with that of brightness perception, we designed two forced alternative choice tests instead of the re-ranking tests. The participants were shown two images of the same subject with different subsurface reflections, and were asked to select the one that appears glossier.

Participants The participants were the same as those in Study 1.

Result Analysis

The goal of Study 2 was to investigate the influence of subsurface reflection on gloss perception. Figure 6 plots the probability of selecting one macro-scale image as a function of the complexion difference with another macro-scale image. The figure demonstrates that given two images I and J, there is a higher probability of selecting I as glossier when the complexion of I is darker than J. The complexion is computed as the average lightness of the subsurface image. The correlation coefficient between surface complexion and perceived gloss is -0.85.

However, the same negative correlation was not always observed with the meso-scale images. Figure 7 plots the probability of selecting one meso-scale image as a function of the complexion difference with another meso-scale image for all the skin patches in the study, organized in three clusters. The three clusters are based on the original average lightness values, from left to right: high (a), medium (b), low (c). Note that the negative correlation between complexion difference and gloss perception still holds for images with relatively high lightness (Figure 7(a)). However, as the average lightness decreases, the influence of complexion difference becomes weaker (Figure 7(b)) and gradually disappears (Figure 7(c)).

Statistical features for gloss difference evaluation

The above analysis of the two studies indicates that gloss perception depends on the joint effect of surface and subsurface reflection. Therefore, to quantitatively evaluate skin gloss, we need to combine statistical features of both reflection images. Figure 8 shows the correlations of nine statistics of the surface and subsurface reflection images on gloss perception. The gloss perception value of each image was derived from the subjective results of the two studies using the Bradley-Terry model. It is no surprise to find out that different statistics show different correlations with gloss perception. Note that the error bars for the subsurface reflection statistics are higher than those for the surface reflection statistics, which indicates that the influence of the surface reflection statistics is more stable than that of the subsurface reflection statistics.

Another interesting observation is that corresponding statistics (mean, skewness, cluster shade) of the two reflection images may have opposite effects on gloss perception. This is consistent with the analysis of the results of the empirical studies. For example, the averaged lightness of the surface reflection has positive correlation with gloss perception as it is related to the specular coverage on the skin surface, while the average lightness of the subsurface reflection correlates negatively with gloss perception since now the mean value indicates the skin complexion.

We then concatenated the statistics of both reflection images and learned a linear regression model to predict relative gloss difference. As a comparison, we also learned a linear regression model based on the same statistics extracted only from the overall reflection map (single modality) of each skin image. Figure 9 shows how the predicted results of the two models aligned with the results of human judgments. With 10-fold cross validation, the RMSE of single image statistics is 8.43, while the RMSE of the combined statistics of the surface and subsurface reflation images is 2.14, a considerable performance improvement.

Conclusions and future work

We investigated the statistical influence of surface and subsurface reflection on human skin gloss perception. We used multimodal photography and photometric transformations for independent manipulation of the surface and subsurface reflection images, in order to obtain a database of skin images with a broad range of appearance. Our empirical studies indicate that skin gloss perception is jointly affected by statistics of the surface reflection and the subsurface reflection images, each of which has a different effect on visual gloss. Compared with the image statistics of overall reflections, the breakdown into surface and subsurface reflection image statistics results in a model that is more consistent with human judgments. However, the models we developed were based on a fixed set of lighting conditions and comparisons within subjects. The development of models that are valid across subjects and a wider range of lighting conditions is a challenging problem that will be addressed in the future.

References

[1] Adelson, E. H., "On seeing stuff: The perception of materials by humans and machines," in [*Human Vision and Electronic Imaging VI*], Rogowitz, B. E. and Pappas, T. N., eds., *Proc. SPIE* **4299**, 1–12 (Jan. 2001).

[2] Wang, J., Pappas, T. N., and de Ridder, H., "Effects of contrast adjustment on visual gloss of natural textures," in [*SPIE/IS&T Electronic Imaging*], 93940F–93940F, International Society for Optics and Photonics (2015).

[3] Wijntjes, M. W. and Pont, S. C., "Illusory gloss on lambertian surfaces," *Journal of Vision* **10**(9), 13–13 (2010).

[4] Motoyoshi, I., Nishida, S., Sharan, L., and Adelson, E. H., "Image statistics and the perception of surface qualities," *Nature* **447**, 206–209 (May 2007).

[5] Anderson, B. L. and Kim, J., "Image statistics do not explain the perception of gloss and lightness," *Journal of vision* **9**(11), 10–10 (2009).

[6] Kim, J. and Anderson, B. L., "Image statistics and the perception of surface gloss and lightness," *Journal of Vision* **10**(9), 3–3 (2010).

[7] Marlow, P. J., Kim, J., and Anderson, B. L., "The perception and misperception of specular surface reflectance," *Current Biology* **22**(20), 1909–1913 (2012).

[8] Ho, Y.-X., Landy, M. S., and Maloney, L. T., "Conjoint measurement of gloss and surface texture," *Psychological Science* **19**, 196–204 (Feb. 2008).

[9] Wijntjes, M. W. A. and Pont, S. C., "Illusory gloss on Lambertian surfaces," *Journal of Vision* **10**(9), 1–12 (2010).

[10] Qi, L., Chantler, M. J., Siebert, J. P., and Dong, J., "The joint effect of mesoscale and microscale roughness on perceived gloss," *Vision research* **115**, 209–217 (2015).

[11] Wiebel, C. B., Toscani, M., and Gegenfurtner, K. R., "Statistical correlates of perceived gloss in natural images," *Vision research* **115**, 175–187 (2015).

[12] Xiong, X. and De la Torre, F., "Supervised descent method and its applications to face alignment," in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], (2013).

[13] Tan, W. R., Chan, C. S., Yogarajah, P., and Condell, J., "A fusion approach for efficient human skin detection," *IEEE Transactions on Industrial Informatics* **8**(1), 138–147 (2012).

Appendix

Subjective Study 1 was a ranking test on images with varying surface reflection. Figure A.1 and Figure A.2 show one group of stimuli in one trial at macro-scale and meso-scale levels respectively. In each trial, 22 images from one skin subject with before and after skin cleansing in one view were displayed randomly. Each original skin subject was accompanied with 10 varying surface reflection.

Subjective Study 2 was a 2AFC test on images with varying subsurface reflection. Figure A.3 and Figure A.4 are one group of stimuli at macro-scale and meso-scale levels respectively. Each original skin subject was accompanied with 8 varying subsurface reflection. In each trial, two images were randomly selected from the 9 images (1 original + 8 SubSurfRefl) of an arbitrary skin subject.

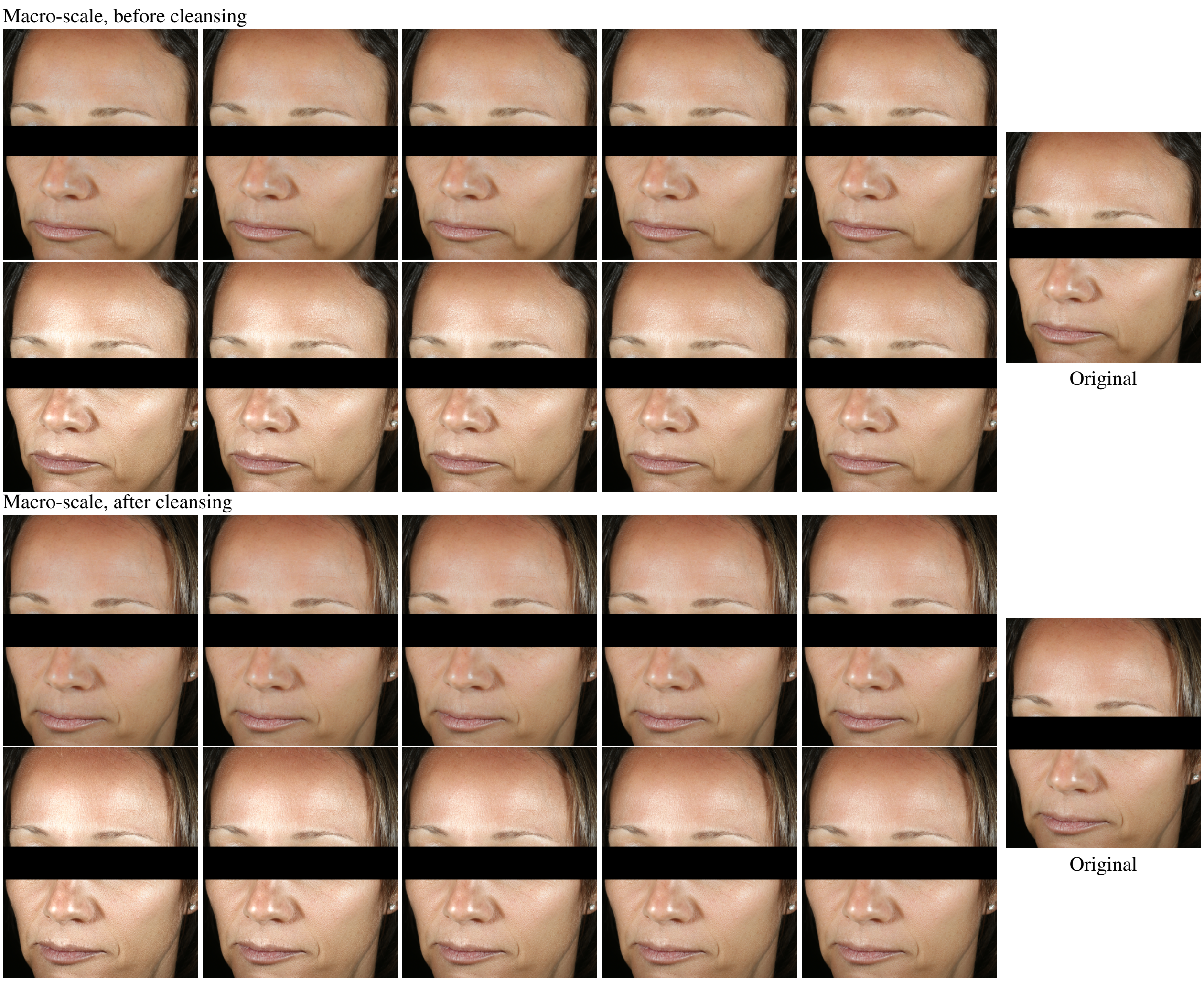


Figure A.1: Macro-scale stimuli examples for Study 1

Figure A.2: Meso-scale stimuli examples for Study 1

Figure A.3: Macro-scale example stimuli for Study 2.

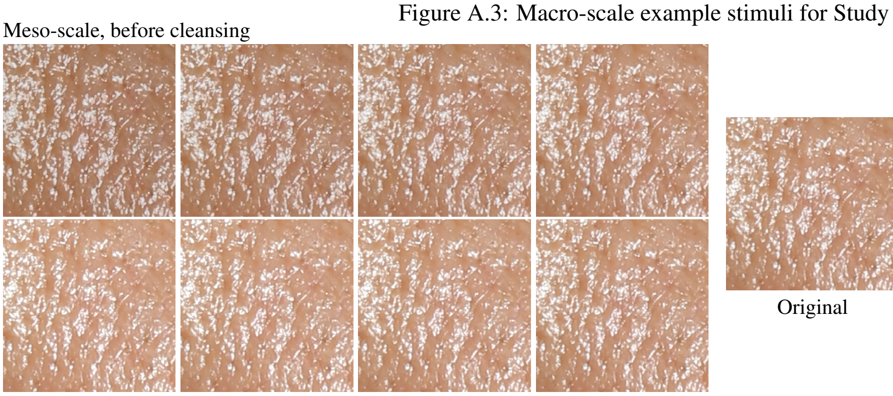


Figure A.4: Meso-scale example stimuli for Study 2.