FACE IMAGE ENHANCEMENT USING 3D AND SPECTRAL INFORMATION

Charles Dubout[*†], *Masato Tsukada*[*], *Rui Ishiyama*[*], *Chisato Funayama*[*], *and Sabine Süsstrunk*[†]

[*]Common Platform Software Research Laboratories, NEC Corporation, Japan
[†]School of Computer and Communication Sciences, EPFL, Switzerland

ABSTRACT

This paper presents a novel method of enhancing image quality of face pictures using 3D and spectral information. Most conventional techniques directly work on the image data, shifting the skin color to a predefined skin tone, and so do not take into account the effects of shape and lighting. The proposed method first recovers the 3D shape of a face in an input image using a 3D morphable model. Then, using color constancy and inverse rendering techniques, specularities and the true skin color, i.e., its spectral reflectance, are recovered. The quality of the input image is improved by matching the skin reflectance to a predefined reference and reducing the amount of specularities. The method realizes the enhancement in a more physically accurate manner compared to previous ones. Subjective experiments on image quality demonstrates the validity of the proposed method.

Index Terms— face image enhancement, color constancy, spectral analysis, 3D morphable model

1. INTRODUCTION

With the recent, rapid popularization of imaging appliances digital pictures are now everywhere. All those pictures have ultimately been taken by a digital camera, and if those cameras are getting better and take photographs closer to reality every year, in the end what end users really want is in fact not only realistic pictures but nice looking ones. Humans are particularly sensible to the appearance of skin, making its enhancement both desirable and difficult.

The retouching of a digital photograph is still a tedious process, often requiring the intervention of an expert. Furthermore, since it is impossible to give a formal definition of perceived image quality, as it depends both on the photographed object and the viewer, assessing the quality of the result is complicated, and needs to be tested by a great number of users to be considered valid. It is tedious because the operations performed by the expert are dependent of the image specificities (luminance, scene type, etc.), and so cannot be simply repeated on another image and expected to yield a good result. The reason is that some of the information needed to enhance a photograph (for example, the white point of the illuminant or the surface reflectance of the objects in the scene) are not directly available but must be estimated, even indirectly, by the operator.

A way to solve this problem is to augment the image with additional information about the scene, so that the estimation of those parameters can be done directly, automatically, and so that only the intended parameters are modified. Conventional techniques [1][2] assume that there is a preferred skin tone, to which they can shift the image skin colors. But this assumption cannot be true, as those values depends heavily on the particular illumination of the scene and characteristics of the imaging system. In this paper, we instead make the assumption that there exists a preferred face skin reflectance, independent of lighting and imaging conditions.

Our proposed method works first by canceling the effects of lighting on the face using a 3D model and inverse rendering techniques, before applying a color constancy technique to recover the skin spectral reflectance. The reflectance is then matched to a predefined reference, taken as the recovered mean skin reflectance of a face in a target image. Some other lighting parameters, such as the amount of specularities can also be modified before reconstructing the output image. Techniques based on physical models already exist for skin synthesis [3][4], but we are not aware of any targeting face image enhancement.

Experimental results reveal the benefits of using a physical model to perform image enhancement, as the improved images look more physically correct and are clearly preferred in our subjective experiments.

2. PHYSICAL PARAMETERS RECOVERY

2.1. 3D shape and lighting condition determination

The appearance of a face under different lighting conditions can vary significantly, even though the spectral reflectance of the skin stays constant. However, as shown recently both by Basri and Jacobs [5] and Ramamoorthi and Hanrahan [6], if one neglects the effects of cast shadows and near-field illumination, the irradiance is then a function of the surface normal **n** only and can be well approximated analytically in terms of spherical harmonic coefficients. Those assumptions are reasonable since human heads are mostly convex and the distance to the light is usually much greater than the size of

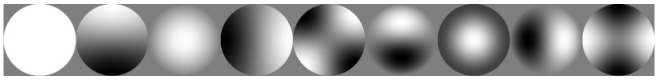

Fig. 1. First 9 spherical harmonic basis functions shown on a sphere. Positive values are in light gray, negative values in dark and zero is set to the gray of the background.

the face. They derived an analytic formula for the irradiance, showing that it can be treated as a convolution of the incident illumination with the Lambertian reflectance function (a clamped cosine). A key result of their work is that Lambertian reflection acts as a low-pass filter, so that the radiance lies very close to a nine-dimensional subspace. The eigenvectors of this subspace are simply quadratic polynomials of the Cartesian components of $\mathbf{n}$, and are illustrated in Fig. 1. It is thus possible to closely model the reflected radiance of a solid diffuse object under any distant illumination with just nine coefficients. In the case of a textured object, the irradiance E is simply scaled by the surface albedo $\rho(\mathbf{x})$ which depends on the position $\mathbf{x}$ and gives the reflected radiance B, directly related to image intensity.

$$B_k(\mathbf{x}, \mathbf{n}) = \rho_k(\mathbf{x})E(\mathbf{n}) \tag{1}$$

As our method takes only a single image as input, we fit a morphable face model [7] to recover the normal vector $\mathbf{n}$ at each pixel. We used an extended version of the original algorithm, based on [8], which can fit a 3D morphable model without any prior assumption on the illumination. Augmenting the image with 3D information enables us to decompose each pixel intensity into albedo, specularities and shading terms. This improves the effectiveness of the skin reflectance recovery, as it allows the estimation to be performed on the specularity and shading free skin albedo, which is the only thing we would like to modify. Under the assumption that skin albedo is constant at low frequency, one can solve for the nine spherical harmonics coefficients using a least square procedure [9]. The coefficients will be scaled by the constant skin albedo, which thus must be estimated to obtain the true irradiance. Once the irradiance has been recovered, one can simply invert Eq. (1), dividing the image intensities by the irradiance to get the albedo. An additional improvement comes from the fact that it is also easy to estimate specularities. Image pixel intensity of value greater than the recovered reflected radiance B are simply clamped, and the residual part is taken as specularities, i.e.:

$$\delta_k(\mathbf{x}) = max(\sigma_k(\mathbf{x}) - B_k(\mathbf{x}, \mathbf{n}), 0) \tag{2}$$

$$\rho_k(\mathbf{x}) = \frac{\sigma_k(\mathbf{x}) - \delta_k(\mathbf{x})}{E(\mathbf{n})} \tag{3}$$

where $\rho(\mathbf{x})$ is the albedo and $\delta(\mathbf{x})$ is the estimated specularity component. The whole process is described in Fig. 2.

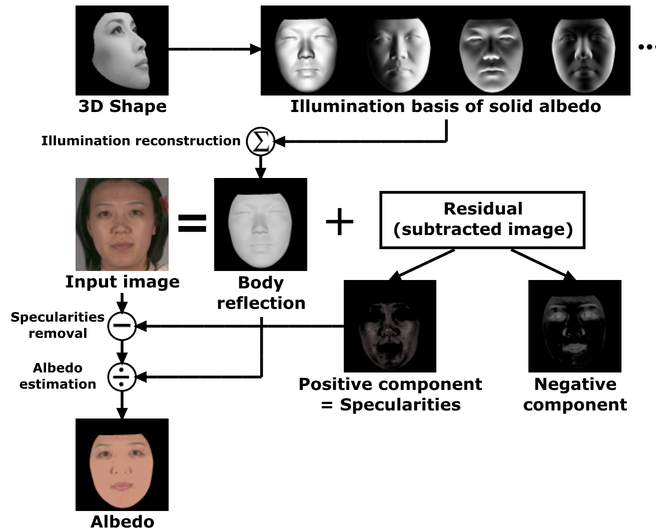

Fig. 2. Albedo and specularities recovery.

2.2. Skin spectral reflectance estimation

As detailed in [10], supposing that we can ignore the surface characteristics, lighting, and viewing geometry by using a relative SPD $E(\mathbf{x}, \lambda)$ instead of physical irradiance measures, the color response $\sigma_k(\mathbf{x})$ of a sensor k with sensitivity $R_k(\lambda)$ is:

$$\sigma_k(\mathbf{x}) = \int_{\upsilon s} S(\mathbf{x}, \lambda)E(\mathbf{x}, \lambda)R_k(\lambda)d\lambda \tag{4}$$

where $S(\mathbf{x}, \lambda)$ is the spectral reflectance of the object at position $\mathbf{x}$ and υs indicates the visible spectrum. As shown in [11], it is usually enough to represent the functions $R(\lambda)$, $S(\mathbf{x}, \lambda)$ and $E(\mathbf{x}, \lambda)$ by samples taken at 10 nm intervals over the spectral range of 400 to 700 nm. Using linear algebra notations, reflectance $S(\mathbf{x}, \lambda)$, illumination $E(\mathbf{x}, \lambda)$, and sensor sensitivity $R(\lambda)$ can thus respectively be expressed as the 31×1 vectors $\mathbf{s}$, $\mathbf{e}$, $\mathbf{r}$ and Eq. (4) can be simply written:

$$\sigma_k = \mathbf{s}^{\intercal} diag(\mathbf{e})\mathbf{r}_k \tag{5}$$

where $^{\intercal}$ indicates the transpose and *diag* is an operator that turns a vector into a diagonal matrix. Since our goal is to enhance images taken by a standard digital color camera, which process colors so as to be viewable by the human visual system, we used the CIE 1931 color matching functions, and appropriately converted input images to CIEXYZ.

Having first estimated the power spectral distribution of the illuminant, it is easy to recover the spectral surface reflectance vector $\mathbf{s}$ of each skin pixel in the input image using a color constancy technique. We first determined a skin reflectance basis by principal component analysis (*PCA*) over a database consisting of the skin spectral reflectance of 4407 Japanese men and women from the data of several cosmetic companies. We then solved the linear system obtained from Eq. (5) at each pixel location, setting $\mathbf{s} = \mathbf{c}_0 + \sum_{i=1}^{3} s_i\mathbf{c}_i =$

$\mathbf{c}_0 + \begin{bmatrix}\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3\end{bmatrix}\begin{bmatrix}s_1 s_2 s_3\end{bmatrix}^{\mathrm{T}}$. Since we have three color stimuli, only the first three coefficients s_i, $i = 1, 2, 3$, corresponding to the first three eigenvectors $\mathbf{c}_i$ of the basis can be recovered, where $\mathbf{c}_0$ is the mean skin reflectance of the da, which was subtracted before performing the *PCA* analysis. Three basis vector are enough to get a good approximation of the real skin reflectance, as human skin reflectance function is fairly smooth. Our *PCA* analysis reveals that for our database, the first three eigenvectors already account for 85% of the energy. The system that we have to solve is:

$$\begin{bmatrix}\sigma_1\\\sigma_2\\\sigma_3\end{bmatrix} = \begin{bmatrix}\mathbf{r}_1\mathbf{r}_2\mathbf{r}_3\end{bmatrix}^{\mathrm{T}} diag(\mathbf{e})(\mathbf{c}_0 + \begin{bmatrix}\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3\end{bmatrix}\begin{bmatrix}s_1\\s_2\\s_3\end{bmatrix}). \tag{6}$$

Defining the 3×31 matrix $\mathrm{M} = \begin{bmatrix}\mathbf{r}_1\mathbf{r}_2\mathbf{r}_3\end{bmatrix}^{\mathrm{T}} diag(\mathbf{e})$, converting skin reflectance to color responses, it is equivalent to:

$$\begin{bmatrix}s_1\\s_2\\s_3\end{bmatrix} = (\mathrm{M}\begin{bmatrix}\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3\end{bmatrix})^{-1}(\begin{bmatrix}\sigma_1\\\sigma_2\\\sigma_3\end{bmatrix} - \mathrm{M}\mathbf{c}_0). \tag{7}$$

The reflectance of every skin pixel can thus be estimated very efficiently by a matrix multiplication and vector subtraction.

3. ENHANCEMENT PROCESS

The spectral reflectance of every skin pixel can now be recovered by using Eq. (7) on the estimated albedo $\rho(\mathbf{x})$. We improve the perceived image quality of the face in the image by matching its mean reflectance $s_{avg}(\lambda)$ to a preferred reference $s_{ref}(\lambda)$ (taken as the estimated mean skin reflectance of a face in a target photograph). First we determine the function f, matching the mean reflectance to the reference:

$$s_{ref}(\lambda) = f(\lambda) \cdot s_{avg}(\lambda) \Leftrightarrow f(\lambda) = \frac{s_{ref}(\lambda)}{s_{avg}(\lambda)} \tag{8}$$

Using our algebraic notation, we can represent f by a linear transformation F, such that $\mathbf{s}_{ref} = \mathrm{F}\mathbf{s}_{avg}$ with:

$$\mathrm{F} = diag([\frac{\mathbf{s}_{ref,1}}{\mathbf{s}_{avg,1}} \dots \frac{\mathbf{s}_{ref,31}}{\mathbf{s}_{avg,31}}]) \tag{9}$$

The estimated skin reflectance of each pixel is then multiplied by the function f (or the matrix F in algebraic notation), giving the enhanced skin reflectance values, which we convert back to color stimuli to get the enhanced image. The whole process can be summarized as:

$$\begin{bmatrix}\rho'_1\\\rho'_2\\\rho'_3\end{bmatrix} = \mathrm{MF}[\mathbf{c}_0 + \begin{bmatrix}\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3\end{bmatrix}(\mathrm{M}\begin{bmatrix}\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3\end{bmatrix})^{-1}(\begin{bmatrix}\rho_1\\\rho_2\\\rho_3\end{bmatrix} - \mathrm{M}\mathbf{c}_0)] \tag{10}$$

Where $\rho'_1, \rho'_2, \rho'_3$ are the enhanced color stimuli and the other variables are defined as before. An overview of the whole enhancement process is drawn in Fig. 3.

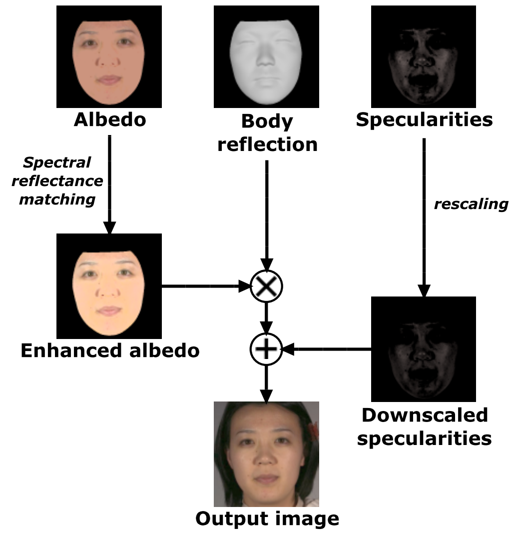

Fig. 3. Enhancement steps.

An additional improvement comes from smoothing the appearance of the skin by scaling down the specularities image. As each pixel is treated independently there is no blur effect as would be observed by trying to smooth the face image directly. Face specularities comes primarily from the skin surface lipid film (SSLF) [12]. Reducing its intensity corresponds thus roughly to reducing the amount of sebum and sweat on the skin surface. Such specularities reveal the skin's imperfections, and are thus undesirable to most people.

4. SUBJECTIVE EXPERIMENTS

To assess the performance of the proposed algorithm for face relighting, experiments were realized where volunteers had to compare pictures enhanced by a conventional method [1] and by the proposed method with and without 3D information.

We used the pairwise comparison method to determine the performance order of the different enhancement techniques. Two images randomly selected from four (the three enhanced images and the original one) were displayed on a monitor. They were not displayed simultaneously but alternatively in response to the subjects' mouse clicks. Subjects were instructed to select from the two images the one they preferred. The experiment was repeated twice to improve the accuracy of the data obtained. Ten images were compared by fourteen subjects (seven men and women, with normal color vision), all Japanese as they are the target users. There is $\binom{4}{2} = 6$ possible combinations of comparisons, making a total of $10 \times 14 \times 6 \times 2 = 1680$ tests. The results of all the experiments can be read in Table 1, and one of the ten images used is shown in Fig. 4.

It can be seen from Table 1 that the proposed method works well, as it outperforms the conventional method 256 times on 280. We confirmed by variance analysis that the re-

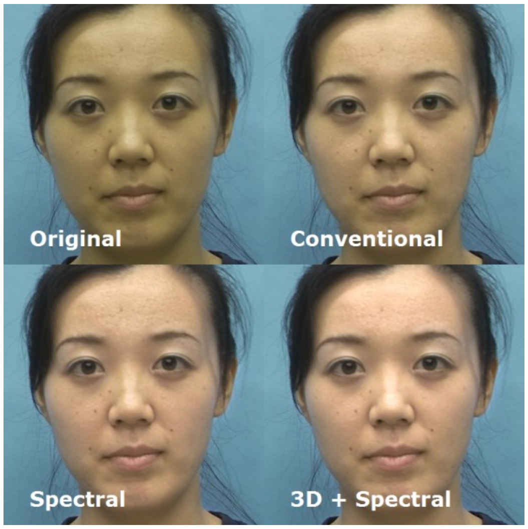

Fig. 4. One of the ten images used in the experiments, enhanced by each model.

sults obtained in the subjective experiments were significant at the 1% level ($F(3, 36) = 4.38$, $P < 0.01$). Interval scales were calculated from the evaluation results of the fourteen subjects using Thurstone's law. Fig. 5 shows the results. If the conventional method indeed leads to an improvement of the original image, the proposed method obtains significantly higher scores, both with and without 3D. Using 3D also leads to a notable improvement, although maybe not as much as expected. A likely explanation is that human faces are relatively flat, except for the curvature of the sides and the protrusion of the nose, making the lighting artifacts sometimes hard to spot. Comparing the two bottom images of Fig. 4, one can see that ignoring shape information makes the face look flat, as the lighting is wrongly estimated.

Table 1. Total number of times an image enhanced with model (i) was chosen over one corrected by model (j).

i | j	Original	Convent.	Spectral	3D + Spectral
Original	0	63	24	18
Convent.	217	0	48	24
Spectral	256	232	0	107
3D + Spectral	262	256	173	0

5. CONCLUSIONS AND FUTURE WORK

This paper presents a method to enhance the perceived quality of an image of a face, using physical parameters instead of directly modifying the image data. If additional work is required to estimate those parameters more reliably, we believe that our approach of using physical parameters for automatic image enhancement is promising, as proved by the results obtained in our subjective experiments.

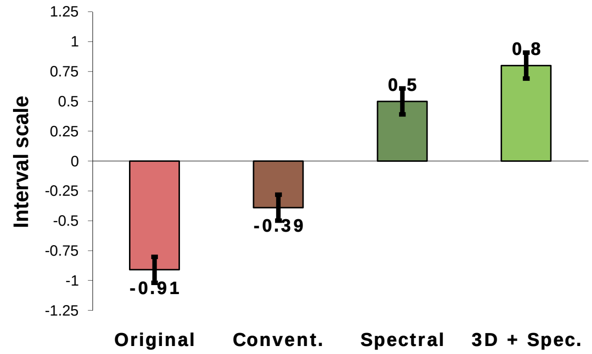

Fig. 5. Interval scales and confidence intervals calculated using Thurstone's law.

6. REFERENCES

[1] M. Tsukada, C. Funayama, and J. Tajima, "Color correction method for preferred color reproduction," in *Proc. IEICE*, 2000, vol. 2, p. 343.

[2] D.-H. Kim, H.-C. Do, and S.-I. Chien, "Preferred skin color reproduction based on adaptive affine transform," *IEEE Transactions on Consumer Electronics*, vol. 51, pp. 191–197, 2005.

[3] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, and Y. Miyake, "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin," *ACM Transactions on Graphics*, vol. 22, pp. 770–779, 2003.

[4] T. Weyrich, W. Matusik, H. Pfister, B. Bickel, C. Donner, C. Tu, J. McAndless, J. Lee, A. Ngan, H. Wann Jensen, and M. Gross, "Analysis of human faces using a measurement-based skin reflectance model," *ACM Transactions on Graphics*, vol. 25, pp. 1013–1024, 2006.

[5] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," in *Proc. ICCV*, 2001, vol. 25, pp. 383–390.

[6] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering," in *Proc. SIGGRAPH*, 2001, pp. 117–128.

[7] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in *Proc. SIGGRAPH*, 1999, pp. 187–194.

[8] R. Ishiyama and S. Sakamoto, "Fast and accurate facial pose estimation by aligning a 3d appearance model," *ICPR*, vol. 4, pp. 388–391, 2004.

[9] Z. Wen, Z. Liu, and T. Huang, "Face relighting with radiance environment maps," in *Proc. IEEE Conf. on Computer Vision & Pattern Recognition*, 2003, pp. 158–165.

[10] M. Peres (Editor) and S. Süsstrunk, *The Focal Encyclopedia of Photography, Fourth Edition: Digital Imaging, Theory and Applications, History, and Science*, Focal Press, 2007.

[11] B. Smith, C. Spiekermann, and R. Sember, "Numerical methods for colorimetric calculations: Sampling density requirements," *COLOR Research and Application*, vol. 17(6), pp. 394–401, 1992.

[12] T. Igarashi, K. Nishino, and S. Nayar, "The appearance of human skin," Tech. Rep., 2005.