Cosmetic Features Extraction by a Single Image Makeup Decomposition

Kanami Yamagishi[1], Shintaro Yamamoto[1],Takuya Kato[1], and Shigeo Morishima[2]
[1]Waseda University [2]Waseda Research Institute for Science and Engineering, Japan
{k_yamagishi, s.yamamoto, takuya_lbj}@{fuji, fuji, ruri}.waseda.jp, shigeo@waseda.jp

Abstract

In recent years, a large number of makeup images have been shared on social media. Most of these images lack information about the cosmetics used, such as color, glitter or etc., while they are difficult to infer due to the diversity of skin color or lighting conditions. In this paper, our goal is to estimate cosmetic features only from a single makeup image. Previous work has measured the material parameters of cosmetic products from pairs of images showing the face with and without makeup, but such comparison images are not always available. Furthermore, this method cannot represent local effects such as pearl or glitter since they adapted physically-based reflectance models. We propose a novel image-based method to extract cosmetic features considering both color and local effects by decomposing the target image into makeup and skin color using Difference of Gaussian (DoG). Our method can be applied for single, standalone makeup images, and considers both local effects and color. In addition, our method is robust to the skin color difference due to the decomposition separating makeup from skin. The experimental results demonstrate that our method is more robust to skin color difference and captures characteristics of each cosmetic product.

1. Introduction

Many people wishing to showcase their skills have taken to sharing makeup images on social media. While many people wish to recreate the makeup through the images, information about the cosmetics is not always labeled. Skin color and lighting conditions make it difficult to specify the features of the cosmetics, as shown in Figure 1.

In the fields of computer vision, makeup simulation has been studied. Li *et al*. [2] proposed makeup simulation by building physics-based reflectance model. Although their results have some validity, local effects such as pearl or glitter are not considered. Crucially, their cosmetics parameter measurement requires a pair of makeup and non-makeup images while these pair images are often not available for practical use.

Figure 1. Skin color effect. (left) Cosmetic product image. (middle) Cosmetic product applied for Subject A. (right) Cosmetic product applied for Subject B.

Figure 2. A makeup image decomposition example. (left) Makeup image. (middle) DoG. (right) Residual.

Our goal is to estimate cosmetic features from a single makeup image. In this paper, we propose an image-based method that extracts cosmetic features which represent both color and local effects. Inspired by the work by Shih *et al*. [4], we decompose the makeup image into two components; makeup and skin color represented by Difference of Gaussian (DoG) and residual, respectively. Here, we assume only one cosmetic product is used for the makeup image. Our method is robust to skin color difference due to the separation of makeup and skin effects. To validate the effectiveness of DoG, we conducted comparison with color histogram-based method using makeup images. The experimental results demonstrate that feature extraction using DoG performs better and more robustly to skin color difference than that from the original makeup image.

2. Method

Given a makeup image I, we extract cosmetic features to estimate cosmetics used for the makeup. Our method removes the skin color effect by taking advantage of Difference of Gaussian (DoG). We employ the CIE-Lab as the color space in order to approximate human perception.

Inspired by the results of facial image style transfer method using DoG[4], the assumption can be made that the makeup effects remain on DoG of the makeup face image. Therefore, our feature extraction utilizes the ability of DoG $D(I)$ to represent makeup and removes skin color effect. As in [4], we decompose the makeup image I using convolution operator $\otimes$:

$$D(I) = I - I \otimes G(\sigma) \tag{1}$$

$$R(I) = I \otimes G(\sigma) \tag{2}$$

where $G(\sigma)$ is a 2D Gaussian kernel with standard deviation σ. Figure 2 shows the makeup image decomposition as mentioned above. As shown in Figure 2, DoG $D(I)$ contains makeup features while skin color is excluded to the residual $R(I)$. After obtaining DoG, we calculate color histogram $H(D)$ from $D(I)$ with $M \times N \times N$ bins of the makeup region S and adopt it as cosmetic features. $H(D)$ is normalized by the sum of pixels in S.

3. Experiment

We conducted an experiment by comparing cosmetic features extracted from images of different subjects using the same cosmetic products. Eight kinds of cosmetic products were applied to faces of five Asian subjects with various skin color, and color histogram was extracted from each DoG. We photographed their headshots under identical illumination conditions to obtain the images where the only variance was skin color. Here, we denote normalized color histogram from i-th subject using j-th cosmetics as $H(D_{ij})(1 \leq i \leq 5, 1 \leq j \leq 8)$. We compared with normalized color histogram $H(I_{ij})$ obtained from I_{ij} in CIE-Lab color space with $40 \times 100 \times 100$ bins.

In this experiment, we set standard deviation $\sigma = 2^8$, kernel size $K = \sigma + 1$ for both x and y axis and makeup region of the target image is manually specified. We calculated L2 norm of features extracted from every pair of subjects and cosmetic products. Table 1 and Table 2 shows the L2 norm between Subject 1 and Subject 2 using DoG and original makeup images, respectively. The L2 norm should be minimal for the same cosmetic product pair, regardless of the subject. For each feature from i-th subject $H(I_{i,j}), H(D_{i,j})$, features from i'-th subject $H(I_{i',j'}), H(D_{i',j'})(1 \leq i' \leq 5, 1 \leq j' \leq 8, i \neq i')$ with minimal L2 norm was searched. Table 3 shows how many times j equals j' *i.e.* L2 norm of features for the same cosmetic product is minimal: we define this as correct. With eight different cosmetic products, maximal correction number (C) of our experiment is eight. We compared correction number for every pair of subjects shown in Table 3.

As in Table 3, features extracted from DoG showed more correction number comparing from the original makeup images. This suggests that our method successfully eliminates effect of skin color difference. Moreover, while the range of L2 norm from original image remains small regardless of whether the cosmetic products are the same or different, when using DoG, the range is greater, which suggests that our method successfully represents the characteristics of each cosmetic product. However, Figure 3 shows one example that j equals j' with features from original images, but not with features from DoG, where the target makeup image is close to the selected makeup image with similar cosmetic product with strong effect, rather than to the makeup image with the same cosmetic product applied lightly.

4. Conclusion

In this paper, we propose an image-based method that extracts cosmetic features representing both color and local effects. We decompose makeup image into Difference of Gaussian (DoG) and residual, representing cosmetic effects and skin color respectively. Experimental results indicate that our feature extraction using DoG performs more robustly with various skin colors and represents the difference of each cosmetic product. On the other hand, we have observed that our method is ineffective when there are changes to heaviness of the makeup or illumination due to the difficulty of removing lighting condition from DoG. As a future work, we will propose a cosmetic feature extraction that is robust to variations in heaviness of the makeup and illumination. Although we currently assume only one cosmetic product is used, we will expand our method to be able to work under multiple cosmetic products usage for more practical application. Furthermore, we will improve our results by deep learning which has showed successful achievement in the computer vision and computer graphics fields, some of which are about makeup [1, 3].

Acknowledgements

This work was supported by JST ACCEL Grant Number JPMJAC1602, Japan.

References

[1] H. Chang, J. Lu, F. Yu, and A. Finkelstein. Pairedcyclegan: Asymmetric style transfer for applying and removing makeup. In *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[2] C. Li, K. Zhou, and S. Lin. Simulating makeup through physics-based manipulation of intrinsic image layers. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4621–4629, June 2015.

[3] S. Liu, X. Ou, R. Qian, W. Wang, and X. Cao. Makeup like a superstar: Deep localized makeup transfer network. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, IJCAI'16, pages 2568–2575. AAAI Press, 2016.

[4] Y. Shih, S. Paris, C. Barnes, W. T. Freeman, and F. Durand. Style transfer for headshot portraits. *ACM Trans. Graph.*, 33(4):148:1–148:14, July 2014.

ID	0	1	2	3	4	5	6	7
0	**0.264113**	0.300809	0.468971	0.433649	0.448469	0.614736	0.440149	0.468077
1	0.365945	**0.326033**	0.488086	0.458896	0.46185	0.62915	0.462393	0.487231
2	0.477707	0.42462	**0.265402**	0.318102	0.394407	0.575243	0.386193	0.414803
3	0.513408	0.464395	0.446729	**0.305109**	0.436961	0.605222	0.429575	0.455468
4	0.430571	0.366692	0.416123	0.382134	**0.31704**	0.573998	0.38642	0.415123
5	0.573189	0.525193	0.524673	0.498147	0.503342	**0.209182**	0.500437	0.523232
6	0.503116	0.452334	0.44635	0.414845	0.426494	0.597506	**0.299082**	0.42929
7	0.46676	0.413017	0.407819	0.373055	0.385986	0.568515	0.347062	**0.203815**

Table 1. L2 norm of Subject 1 and Subject 2 using DoG; Horizonal and vertical ID represent each cosmetic product for Subject 1 and Subject 2, respectively. The L2 norm minimal of each row is in bold type.

ID	0	1	2	3	4	5	6	7
0	0.241638	0.204377	0.237473	**0.204105**	0.221685	0.27007	0.216159	0.213377
1	0.217855	0.178821	0.216072	**0.178753**	0.198588	0.251459	0.192402	0.189271
2	0.213681	0.170413	0.208959	**0.170087**	0.190827	0.245374	0.184378	0.181108
3	0.229098	0.189386	0.2247	**0.18907**	0.207945	0.258911	0.202043	0.199064
4	0.220517	0.178955	0.216001	**0.178667**	0.198144	0.251398	0.192322	0.18919
5	0.250165	0.214391	0.246144	**0.214132**	0.230949	0.277725	0.22565	0.222987
6	0.218463	0.176372	0.213848	**0.176058**	0.196168	0.249551	0.1872	0.184754
7	0.201094	0.154344	0.196075	**0.153984**	0.176625	0.234499	0.167049	0.165663

Table 2. L2 norm of Subject 1 and Subject 2 using original makeup image; Horizonal and vertical ID represent each cosmetic product for Subject 1 and Subject 2, respectively. The L2 norm minimal of each row is in bold type.

i	1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5
i'	2	3	4	5	1	3	4	5	1	2	4	5	1	2	3	5	1	2	3	4
C from Original	1	1	3	5	1	5	5	3	1	5	4	3	4	3	2	8	3	1	1	4
C from DoG	7	6	7	7	8	7	7	8	7	7	8	7	6	7	7	7	6	7	5	8

Table 3. Comparison of correction number (C) between all subject pairs.

Figure 3. An example that j equals j' with features from original images while not with features from DoG. (left) Target makeup image of Subject i, (middle) Makeup image with the same cosmetic product of Subject i', (right) Makeup image of Subject i' that L2 norm of features extracted from DoG is minimal with features from left image.