Hairstyle Suggestion Using Statistical Learning

Wei Yang, Masahiro Toyoura, and Xiaoyang Mao

University of Yamanashi
4-4-11 Takeda, Kofu, Yamanashi, Japan 400-8511
{g10mk039,mtoyoura,mao}@yamanashi.ac.jp

Abstract. Hairstyle is one of the most important features people use to characterize one's appearance. Whether a hairstyle is suitable or not is said to be closely related to one's facial shape. This paper proposes a new technique for automatically retrieving a suitable hairstyle from a collection of hairstyle examples through learning the relationship between facial shapes and suitable hairstyles. A method of hair-face image composition utilizing modern matting technique was also developed to synthesize realistic hairstyle images. The effectiveness of the proposed technique was validated through evaluation experiments.

Keywords: Hairstyle retrieval, example-based, statistical learning, non-parametric sampling, hairstyle image synthesis.

1 Introduction

With the advance of media acquisition and processing technologies, multimedia is now playing a very important role in enriching our daily life. This paper presents a new technology for suggesting a user with his/her suitable hairstyles by combing statistical learning and image processing techniques. Hairstyle is one of the most important features people use to determine their appearance and mood. People can look completely different simply by changing their hairstyles. Everyone would like to have a suitable hairstyle to make them look attractive, but it is usually difficult to find one as we cannot easily try out various styles with our real hair. Several commercials or free software have been developed allowing users to simulate how they look with different hairstyles by manually selecting hairstyle samples and superimposing them over their facial images. Although these systems do provide some general guidelines on choosing hairstyles, they do not provide any hints on what a suitable hairstyle is for a particular face. Therefore, users usually need to go through a very tedious process of trying out many different hairstyles before the one they like can be found. On the other hand, several papers have been published related to hairstyles [1,2,3,4] in the field of computer graphics. However, to the best of our knowledge, all these have focused on how to model and render hairstyles with computer graphics and have mainly been applied to create virtual characters and animations. In this paper, we propose a new technique for automatically retrieving a suitable hairstyle for a given face from a collection of successful hairstyle examples. What is a suitable hairstyle

K. Schoeffmann et al. (Eds.): MMM 2012, LNCS 7131, pp. 277–287, 2012.

for someone? Suitability is a perceptual attribute and it is difficult to model it computationally. A hairstyle that appears to be attractive to one person may not look that acceptable to another. Personal aesthetics may be affected by many factors, such as one's cultural background and living environment. There are also many other factors that may affect how one looks in a particular hairstyle. Despite this, there is still some common aesthetics sense on hairstyles, and a skilled hair stylist can usually successfully create styles that conform to such common aesthetics sense. We started our project by interviewing several hair stylists. An important fact we observed is that although there are no stylists who can tell any explicit rules about their designs, they all viewed the shape of the face as the most important attribute in designing a hairstyle. This inspired us to adopt an example-based framework, which is an approach that has been successfully used for texture synthesis [5, 6] and style transferring [7,8,9,10] in recent years. Our new technique finds suitable hairstyles for a given face by learning the relationship between facial shapes and successful hairstyles.

The four major contributions of this paper can be summarized as follows:

1. A new framework for retrieving suitable hairstyles through learning the relationship between facial shapes and hairstyles from successful hairstyle examples.
2. The design of a compact feature vector space enabling fast non-parametric sampling in statistical learning.
3. A method of hair-face image composition utilizing modern matting techniques for synthesizing realistic hairstyle images automatically.
4. An evaluation experiment demonstrating the validity of the feature vector and the effectiveness of the example-based approach.

2 Example-Based Framework

Given a face image I_{input}, we want to create another image I_{output} with hairstyle S matching the face best. We achieve this in two steps:

Statistical Learning: Find the most suitable hairstyle S, through learning the relationship between facial shapes and suitable hairstyles.

Composition: Superimpose hairstyle S over face image I_{input} to obtain realistic image I_{output} of the face in a suitable hairstyle.

2.1 Statistical Learning

The statistical learning step can be described within the Bayesian inference framework. Based on Bayes theorem, posterior probability $P(S|I_{input})$, i.e., the probability for face image I_{input} to have its best hairstyle S, can be represented as:

$$P(S|I_{input}) = \frac{P(I_{input}|S)P(S)}{P(I_{input})}. \quad (1)$$

$P(S)$ is the prior probability of S, and $\mathrm{P(I_{input}|S)}$ is the probability of observed face image I_{input} given hairstyle S. Consequently, the aim of finding S can be turned into an optimization problem maximizing $P(S|I_{input})$. Since evidence $P(I_{input})$ can be treated as a constant, S is actually the one maximizing the product of likelihood $P\left(I_{input}|S\right)$ and prior $P(S)$,

$$S = arg\, max_S\, P(I_{input}|S)P(S). \tag{2}$$

We obtain S by using non-parametric sampling, which has been proved to be an easy yet very efficient method in style transferring and texture synthesis applications [5,6,8,9]. Professionally designed hairstyle examples are used as the training data to learn prior $P(S)$ and likelihood $P(I_{input}|S)$.

To compute S, we first construct an approximation to conditional probability distribution $P(S|I_{input})$ and sample from this. Assuming $d(I^T, I_{input})$ is the distance between two facial images under some metric, if we define set

$$\Omega\left(I_{input}\right) = \left\{ I^T \left| I^T \subset T, d\left(I^T, I_{input}\right) = 0\right.\right\}. \tag{3}$$

containing all occurrences of I in training data set T, then the conditional probability distribution of $P(H)$ can be estimated with a histogram of all the hairstyles for faces I^T in $\Omega\ (I_{input})$. In other words, we can obtain a distribution of possible hairstyles for I_{input}. However, since we only have a finite number of examples from the training data set, we may fail to find any matching facial image I^T with $d\left(I^T, I_{input}\right) = 0$. To obtain an approximation to $\Omega\ (I_{input})$, let us specify small distance allowance e, and obtain $\Omega'\ (I_{input}) = \{ I^T\ |I^T \subset T, d\left(I^T, I_{input}\right) < e\}$. The distance function d will be discussed in Section 3. Allowance e is determined by multiplying the minimum of $d\left(I^T, I_{input}\right)$ with a given constant, in our current implementation.

There are several possible ways of computing S from its conditional probability distribution:

1. Integrate the distribution and obtain an expected hairstyle image by compositing all the possible hairstyles.
2. Take the hairstyle with the highest probability. If there is no maximum value in the distribution, randomly choose a hairstyle.
3. Take the hairstyles of the *k*-nearest-neighbor of I_{input}.

Method 1 may produce some unrealistic hairstyle images due to the discontinuity in the hairstyles from the distribution. Since the training dataset usually consists of no more than one hairstyle for one single facial image, the result from Method 2 is usually a randomly chosen hairstyle from the set of hairstyles $I^T \subset \Omega'(I_{input})$. Method 3 is used in our current implementation and the hairstyles of the *k*-nearest-neighbor of I_{input} are recommended as candidates for the best suitable hairstyles. The *k* can be adjusted interactively.

2.2 System Overview

As shown in Fig 1, our system assumes n successful hairstyle images $I_i^T(i = 1,2,\ldots,n)$ are available. The three operations below are executed in the training phase to build a training data set $\mathrm{T}(V_i^T, \alpha_i^T)(\mathrm{i} = 1,\ldots \mathrm{n})$ where V_i^T is the feature vector characterizing the shape of each face I_i^T and α_i^T is the α-matte indicating the probability of a hair area in the image:

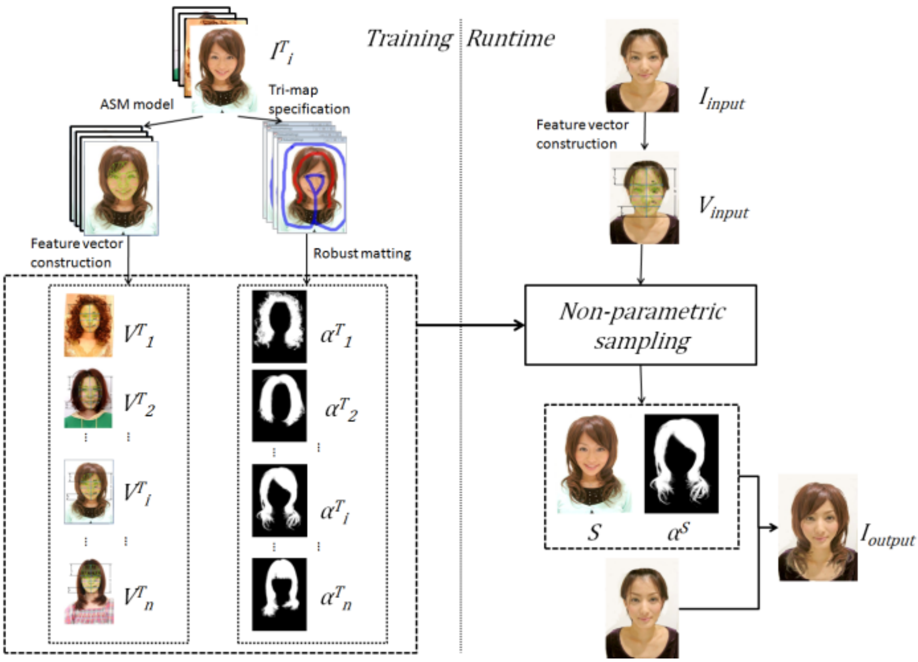


Fig. 1. System framework

1. Apply robust matting technique [11] to $I_i^T(i = 1,2,\ldots,n)$ to create $\alpha_i^T(i = 1,2,\ldots,n)$.
2. A trained Active Shape Models (ASM) [12] model is used to detect facial feature points on $I_i^T(i = 1,2,\ldots,n)$.
3. Construct feature vector $V_i^T(i = 1,2,\ldots,n)$ from the ASM model feature points.

While the first operation requires example strokes to be manually specified to create the tri-map for estimating the α-matte, the other two operations are performed fully automatically.

Given face image I_{input} in the runtime phase, the system performs six operations to compute a set of suitable hairstyles for I_{input}.

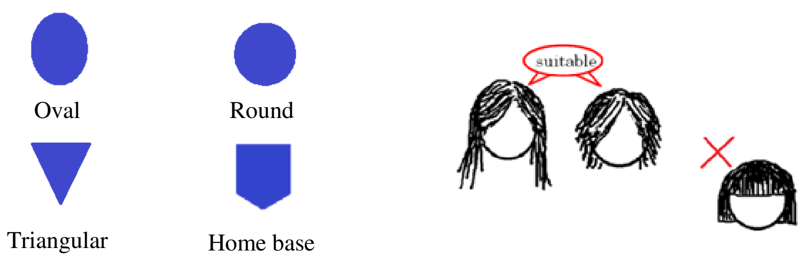
1. Apply a trained ASM to detect the facial feature points of I_{input}.
2. Construct feature vector V_{input} characterizing the shape of the face in I_{input}.
3. Search through all images in T in the feature vector space.

4. If $d(I^T, V_{input}) < e$ add I_i^T to $\Omega'(I_{input})$.
5. Sort images in $\Omega'(I_{input})$ by $d(I^T, V_{input})$.
6. Take top k images $I_j^T (j = 1,2, \dots, k)$ from $\Omega'(I_{input})$ and composite them with I_{input} using $\alpha_j^T (j = 1,2, \dots, k)$.

The design of feature vector V is crucial to obtain good results, as well as to quickly searches through the training examples. We will discuss this in detail in the next section.

3 Feature Vector Design

Active Shape Models (ASM) [5] is one of the most popular techniques for detecting the geometric features of faces from images. Instead of using ASM directly, we want to have a more compact feature vector space, which can successfully model the relationship between facial shapes and hairstyles. As seen in Fig 2(a), it is known that human faces can be roughly classified into four categories by shape: *oval, round, triangular* and *home base*.



(a)Shapes of face (b) Hairstyles for round faces

Fig. 2. Facial shapes and their relationships to hairstyles

A hairstyle giving an impression of an oval shaped face is likely to be suitable [13]. For example, as we can see from Fig 2(b), for the round face, the styles with long bangs flowing smoothly toward both sides or the back make the face look longer and hence are suitable, while the one with bangs cut straight across and a thick volume on top of the head is not suitable because it further emphasizes the impression of roundness. To find these kinds of relationships between facial shapes and hairstyles, we first compute six line segments (Fig 3(b)):

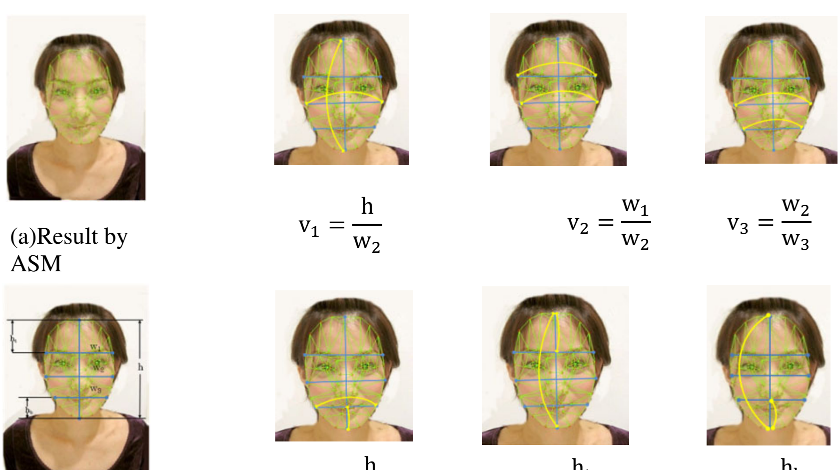
h	:	the center vertical line segment
w_1	:	the horizontal line segment at the height of the eyebrows
w_2	:	the horizontal line segment at the widest position of the facial area
w_3	:	the horizontal line segment at the height of the mouth
h_t	:	the vertical line segment from the top of the face to the cross-section of h and w_1
h_b	:	the vertical line segment from the bottom of the face to the cross-section of h and w_3

From the six line segments above, we define a six-dimensional feature vector $V(v_1,v_2,v_3,v_4,v_5,v_6)$ with $v_j(j=1,2,..6)$ being the ratio of two line segments, normalized to be in (0,1] (Fig 3(c)). To automatically compute the feature vector, we trained the ASM model with 81 points so that it includes the end points of the six line segments as the feature vector (Fig 3(a)).

The distance function for non-parametric sampling is defined as

$$\mathrm{d}\left(\mathrm{I}^{\mathrm{T}}, \mathrm{I}_{\mathrm{input}}\right) = \sum_{i=1}^{6} k_i \left(v_i^T - v_i^{I_{input}}\right)^2. \tag{4}$$

Coefficient $k_j(j = 1,2, \dots ,6)$ controls the weight of each dimension in determining the hairstyle. For example, if we use a large k_1 and a small k_2, then the system would treat the aspect ratio of the face as being more important and the shape of the cheek as being less important in determining a suitable hairstyle. Two faces with different shapes for cheeks but a similar aspect ratio may be suggested with similar hairstyles. We set all six coefficients to be equal to treat all dimensions homogenously in our current implementation. We plan to explore the best coefficients through machine learning in future work. A. Kagian et al used neural network to model the attractiveness of human face based on a 3240($_{81}C_2$) dimension feature vector [14]. Each dimension of the feature vector is the length of a line segment connecting two feature points of ASM model and the 3240 dimension corresponds to all the possible combination among 81 feature points. They trained the neural network through subjective an experiment. As the future work, we plan to adopt a similar approach explore the best feature vector to model the relationship between the geometric features of faces and hairstyles



(a)Result by ASM

$v_1 = \frac{h}{w_2}$ $v_2 = \frac{w_1}{w_2}$ $v_3 = \frac{w_2}{w_3}$

(b)Feature lines

$v_4 = \frac{h}{w_3}$ $v_5 = \frac{h_t}{h}$ $v_6 = \frac{h_b}{h}$

(c)Components of feature vector

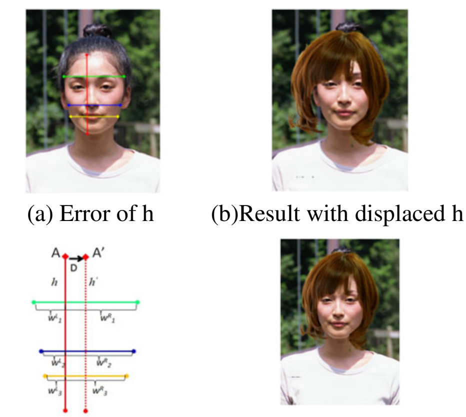
Fig. 3. Feature vector

4 Hairstyle Image Synthesis

Before we can superimpose obtained hairstyle image S over input face image I_{input}, position alignment and size adjustment between the two images are required. This is achieved by scaling H with $w_2^{I_{input}}/w_2$ in width, $h^{I_{input}}/h^H$ in height, followed by translation aligning the upper end point of h^H with that of $h^{I_{input}}$. As shown in Fig 4(a), due to fitting error with the ASM model, we may fail to obtain centered h^H or $h^{I_{input}}$, and this may result in an unnatural composition like the one in Fig 4(b). To correct error, we translate h horizontally by displacement D (Fig 4 (c)):

$$D = \frac{1}{2} \cdot \frac{1}{3} \cdot ((w_1^L - w_1^R) + (w_2^L - w_2^R) + (w_3^L - w_3^R). \tag{5}$$

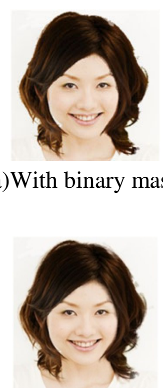
Fig 4(d) shows an improved result obtained by using the new position of h^H and $h^{I_{input}}$ for position alignment.



(a) Error of h (b)Result with displaced h

(c)Correction of h (d)Result with corrected h

Fig. 4. Position alignment



(a)With binary mask

(b) With α-matte

Fig. 5. Compare composition results using binary mask and α-matte

Finally, the α-matte of S is used to composite S and I_{input} *to ob*tain output image I_{output}

$$O_p = (1 - \alpha) \times H_p + \alpha \times I_p. \tag{6}$$

Here, p denotes a pixel, and O_p, H_p, and I_p, correspond to the pixel values for the output hairstyle image, input image, and suitable hairstyles suggested by the system, respectively. Fig 5 compares the results of binary mask based composition with our α-matting based method. We can see our method produces more realistic images, especially at regions near the boundary of hair.

5 Implementation and Evaluation

5.1 Implementation and Result

We built a training data set consisting of 84 hairstyle images collected through the courtesy of hair stylists from three hair salons in our current implementation. Through a preliminary user study, we found that in many cases users wanted to specify the length of their hair before searching for the best hairstyles, and hence it would be helpful if we could advise them of the best candidates for different lengths. Currently, both the hairstyle examples and the input face need to be frontal photos and the facial area in the input image should not be covered by strands of hair. In addition to the six dimensions representing the shape of a face, three additional dimensions, representing the length, hardness and volume of hair, are also added and a nine-dimensional feature vector is computed for each example. Each hair can take one of three values in the three newly added dimensions:

Length : long, medium, short.
Volume : large, medium, small.
Hardness : hard, medium, soft.

The necessity to consider the properties (hardness and volume) of hair arose from the fact that the hairstyle one can actually have is constrained by the properties of his/her hair, even if one can find the best hairstyle making her look virtually attractive. Therefore, with the additional dimensions characterizing the properties of hair, we can constrain sampling to only hairstyles with similar hair properties. The hair properties of the input face are specified by the user.

Fig. 8 shows two examples of results. For the input face images at the left, the nearest face in the feature vector space for each of the long, medium and short hairstyle training sets are shown in the upper row and below it are the resulting hairstyles.

5.2 Evaluation

We conducted two experiments to validate the effectiveness of our approach. The first was aimed at investigating how the hairstyles recommended by our system to a person would look for other people, while the second experiment was aimed at how satisfactory the result would be for the person herself.

Ten female college students participated in the first experiment. We prepared nine sets of hairstyle images, each consisting of ten hairstyles with two of them recommended by the system as suitable hairstyles. At each trial, a subject was presented with one set of images and asked to mark the top three most suitable hairstyles out of the 10 hairstyles. Fig 6 shows an example of the hairstyle image set used in the experiment. Since hair color can largely affect the impression of hairstyle, we used the monochrome picture in the experiment to exclude the effect of color. There were a total number of 90 trials (9 sets×10 people) and we evaluated the probability of the occurrence of the following three cases.

Case 1 : The hairstyles recommended by the system were marked as the top suitable hairstyle.
Case 2 : At least one hairstyle recommended by the system was included in the top two suitable hairstyles.
Case 3 : At least one hairstyle recommended by the system was included in the top three suitable hairstyles.

We used a binomial test and our null hypothesis was that all 10 hairstyles in an image set would be marked with the same probability. Table 1 summarizes the experimental results. Out of the 90 trials, there were 20 trials for Case 1, 58 trials for Case 2 and 80 trials for Case 3. The probability for the number of occurrences above those observed ones upon the null hypothesis is listed at the rightmost column of Table 1. The null hypothesis was rejected at a significance level lower than 0.05 for the latter two cases. In other words, the experimental results suggested that the hairstyles recommended by our system at least could be viewed as the second best hairstyle even though it might not be the best.

The second experiment had the same setting as that for the first except that each subject was presented with hairstyle images of herself. The subjects were ten female college students and 3 image sets were prepared for each of them. Therefore, these were a total of 30 trials (3 sets×10 people). The number of trials for the three cases and the corresponding probability upon null hypothesis are summarized in Table 2. We can see the null hypothesis was rejected with a very low level of significance for the latter two cases, which is the same as the results for the first experiment.

Table 1. Result for Experiment 1: Viewed by Others (binomial test)

	Number of occurrence (Out of 90 trials)	Probability upon null hypothesis
Case 1	22%	0.25
Case 2	64%	0.00
Case 3	89%	0.00

Table 2. Results from Experiment 2: Viewed by self (binomial test)

	Number of occurrence (Out of 30 trials)	Probability upon null hypothesis
Case 1	17%	0.57
Case 2	53%	0.02
Case 3	83%	0.00

We can conclude from the experiment results that even though it may not be the best one, our system can advise users of good candidates for hairstyles viewed to be suitable both by themselves and others.

Fig. 6. Example of image set used for experiment

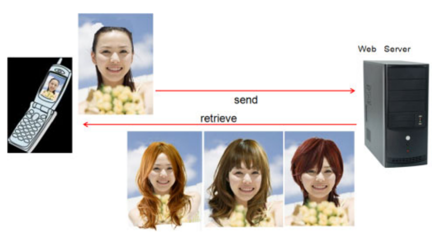


Fig. 7. Example of mobile application

Fig. 8. Example of results

6 Concluding Remarks

We presented a new example-based framework for creating suitable hairstyles for a given face image. Suitability is a perceptual attribute and our evaluation experiments demonstrated the effectiveness of addressing these kinds of problems with an example-based approach. Since the proposed technique is fully automatic, a promising application is to implement it on mobile terminals with camera. For example, shows in Fig. 7, a user can retrieve a suitable hairstyle before going to hair salon simply by taking a photo of herself with the camera on her cell phone and then show it to the hair stylist after arrives the hair salon. In future research work, we want to apply the same framework to other fashion simulation problems, such as advising people of suitable attire by learning the relationships between body shape and successful choice in dress.

References

1. Paris, S., Chang, W., Kozhushnyan, O.I., Jorosz, W., Matusik, W., Zwicker, M., Durand, F.: Hair Photobooth: Geometric and Photometric Acquisition of Real Hairstyles. ACM Transactions on Graphics 27(3), article 30 (2008)
2. Wang, L., Yu, Y., Zhou, K., Guo, B.: Example based hair geometry synthesis. ACM Transactions on Graphics 28(3), Article 56 (2009)
3. Xu, Z., Dong, X., Yang, D.: V-hair studio: An interactive tool for hair design. IEEE Computer Graphics and Applications 21(3), 36–43 (2001)
4. Yu, Y.: Modeling realistic virtual hair styles. In: IEEE Pacific Conference on Computer Graphics and Applications(PG), p. 295. IEEE Computer Society, Washington (2001)
5. Efros, A., Leung, T.K.: Texture synthesis by nonparametric sampling. In: IEEE International Conference on Computer Vision (ICCV), vol. 2, pp. 1033–1038 (1999)
6. Hertzmann, A., Hacobs, C.E., Oliver, N., Curless, B., Salesin, D.H.: Image analogies. ACM Transactions on Graphics 25(4), 1360–1379 (2006)
7. Chang, C., Peng, Y., Chen, Y., Wang, S.: Artistic Painting Style Transformation Using Example-based Sampling Method. Journal of Information Science and Engineering 26(4), 1443–1458 (2010)
8. Chen, H., Xu, Y.Q., Shum, H., Zhu, S.C., Zheng, N.N.: Example-based facial sketch generation with nonparametric sampling. In: IEEE International Conference on Computer Vision (ICCV), pp. 433–438 (2001)
9. Chen, H., Liang, L., Xu, Y.Q., Shum, H.Y., Zheng, N.N.: Example-based automatic portraiture. In: IEEE Asian Conference on Computer Vision, pp. 23–25 (2002)
10. Suo, J., Min, F., Zhu, S., Shan, S., Chen, X.: A multi-resolution dynamic model for face aging simulation. In: IEEE Conference on Computer Vision and Pattern Recognition, Minnesota, USA (June 2007)
11. Wang, J., Cohen, M.F.: Optimized Color Sampling for Robust Matting. In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8 (2007)
12. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. Computer Vision and Image Understanding 61, 38–59 (1995)
13. Best Hairstyle by Face Shape. Edited by Shufunotomo Co., Ltd (2007) (in Japanese)
14. Kagian, A., Dror, G., Leyvand, T., Cohen-Or, D., Ruppin, E.: A humanlike predictor of facial attractiveness. In: Schölkopf, B., Platt, J., Hoffman, T. (eds.) Advances in Neural Information Processing Systems 19, MIT Press (2007)