Research Article

A Simple Framework for Face Photo-Sketch Synthesis

Xuewei Li and Xiaochun Cao

School of Computer Science and Technology, Tianjin University, Tianjin 300072, China

Correspondence should be addressed to Xiaochun Cao, xcao@tju.edu.cn

This paper proposes a simple framework for face photo-sketch synthesis. We first describe the shadow details on faces and extract the prominent facial feature by two-scale decomposition using bilateral filtering. Then, we enhance the hair and some unapparent facial feature regions by combining the edge map and hair color similarity map. Finally, we obtain the face photo sketch by adding the results of the two processes. Compared with current methods, the proposed framework demands non feature localization, training or iteration process, creating vivid hair in sketch synthesis, and process arbitrary lighting conditions of input images, especially for complex self-shadows. And more importantly, it can be easily expanded to natural scene. The effectiveness of the presented framework is evaluated on a variety of databases.

1. Introduction

Face sketching is a simple yet expressive representation of faces. It depicts a concise sketch of a face that captures the most essential perceptual information with a number of strokes [1]. It has useful applications for both digital entertainment and law enforcement.

In recent years, two kinds of representative methods for computer-based face sketching have been presented: (1) line drawing [1–5] and (2) Eigen transformation based [6, 7]. Line drawing-based methods are expressive to convey 3D shading information, at the cost of losing sketch texture. The performances of these approaches largely depend on the shape extraction and facial feature analysis algorithms, such as active appearance model [5]. Other line drawing methods use compositional and-or graph representation [8, 9] or the direct combined model [10] to generate face photo cartoon. Eigen-transformation-based approaches use complex mathematical models to synthesize face sketches, such as PCA, LDA, E-HMM, and MRF. Gao et al. [6] use an embedded hidden Markov model and a selective ensemble strategy to synthesize sketches from photos. However, the hair region is excluded in PCA/LDA/E-HMM-based methods [11]. Wang and Tang [2] use a multiscale Markov random fields (MRFs) model to synthesize face photo-sketch and recognize it. The face

region is first divided into overlapping patches for learning, and the size of the patches decides the scale of local face structures to be learned. From a training set, then the joint photo-sketch model is learned at multiple scales using a multiscale MRF model. This method requires modeling both face shape and texture and can provide more texture information. However, the current approaches have three disadvantages: (1) both line drawing and Eigen-transformation-based methods require complex computing; (2) human face sketch is unnecessarily exaggerated, which depicts the facial feature with distortion; (3) most of the existing methods can only sketch the human face but fail to apply to natural scene images.

In this paper, we present a novel and simple face photo-sketch synthesis framework. The hair is synthesized using a two-scale decomposition and a color similarity map. The proposed framework is very simple, without any iteration, or facial feature extraction. Specially, the proposed method can easily be applied to natural scene for sketching.

A schematic overview of our framework is shown in Figure 1. Firstly, for an input face image, a two-scale image decomposition by bilateral filtering is used to describe the shading texture and the prominent feature shape, while the color similarity map-based hair creating can generate the hair texture and the unapparent facial feature. Then, the edge map is computed by edge detector from skin color similarity map. As a result, the hair, eyes, and thus mouth regions are enhanced by multiplying the edge map and hair color similarity map. Finally, the face photo sketch is synthesized by combining the results of former two processes using addition operation.

2. Bilateral Filter

2.1. Bilateral Filter

The bilateral filter is an edge-preserving filter developed by Tomasi and Manduchi [12]. It is a normalized convolution in which the weighting for each pixel p is determined by the spatial distance from the center pixel q, as well as its relative difference in intensity. The spatial and intensity weighting functions f and g are typically Gaussian [13, 14]. The spatial kernel increases the weight of pixels that are spatially close, and the weight in the intensity domain decreases the weight of pixels with large intensity differences. Therefore, bilateral filter effectively blurs an image while keeping sharp edges intact. For input image I, output image J, and a window Ω neighboring to q, the bilateral filtering is defined as follows:

$$J = \frac{\sum_{p \in \Omega} f(p-q) g(I_p - I_q) I_p}{\sum_{p \in \Omega} f(p-q) g(I_p - I_q)},$$

(2.1)

where

$$f(p-q) = \exp\left[-\frac{\left(\|p-q\|^2\right)}{2\sigma_S^2}\right],$$

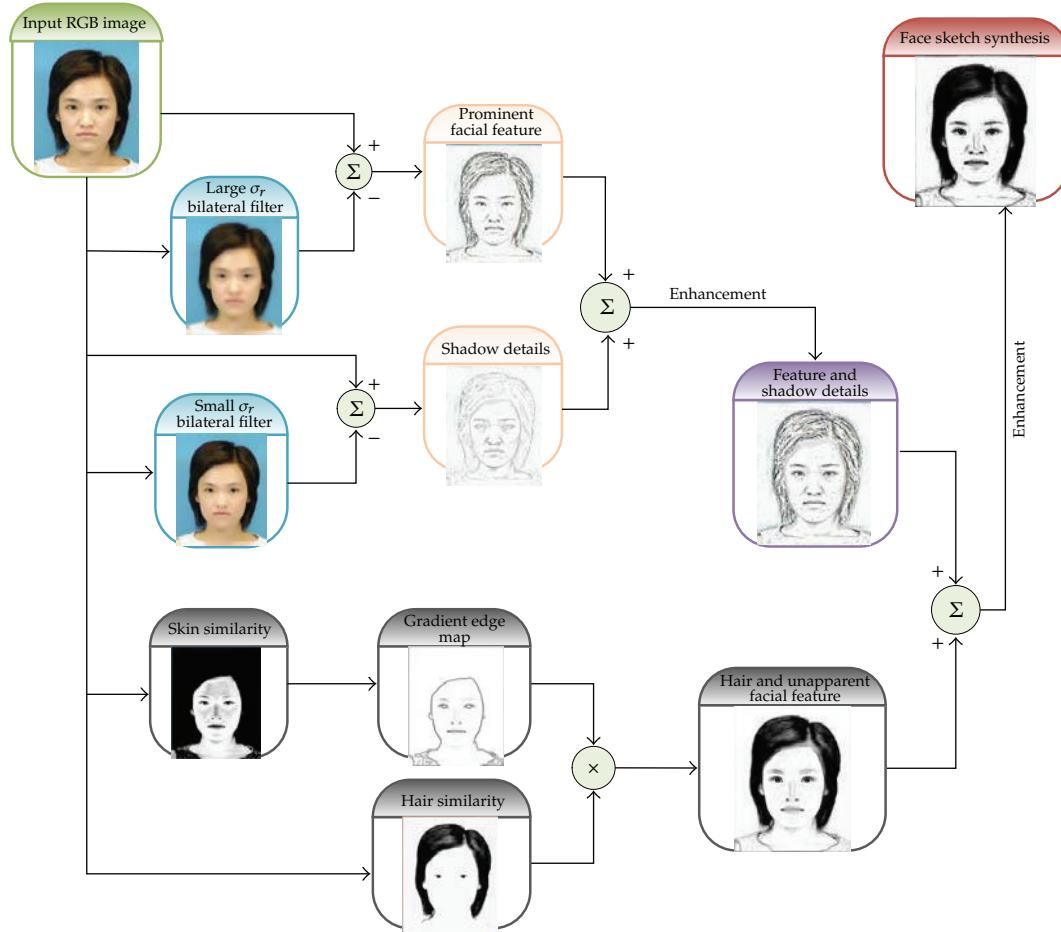$$g(I_p - I_q) = \exp\left[-\frac{(I_p - I_q)^2}{2\sigma_r^2}\right].$$

(2.2)

Figure 1: The system framework.

σ_s and σ_r are the size of spatial kernel and the range kernel, corresponding to the Gaussian functions f and g. When σ_s increases, the larger features in image will be smoothed; when σ_r increases, the bilateral filter will become closer to Gaussian blur.

2.2. Facial Feature Detail Detection by Two-Scale Image Decomposition

Multiscale image decomposition (or multiscale retinex, MSR) is developed by Jobson et al. [15–17], in an attempt to bridge the gap between images and the human observation of scenes. It is widely used in HDR image rendering for color reproduction and contrast reduction [18, 19], color enhancement, and color constancy processing. In HDR image rendition, the smallest scale is strong on detail and dynamic range compression but weak on tonal and color rendition. The reverse is true for the largest spatial scale. Multiscale retinex combines the strengths of each scale and mitigates the weaknesses of each [15].

Durand and Dorsey [19] only use two-scale decomposition to decompose the input image into a "base" and a "detail" image. The base layer has its contrast reduced and contains only large-scale intensity variations, which is obtained by bilateral filtering. The detail layer

is the division of the input intensity by the base layer, while the magnitude is unchanged, thus preserving detail.

Then, we will describe how two-scale decomposition can be used for lighting conditions and facial feature detail detection in face photo-sketch synthesis. According to image decomposition, we can get the detail image by subtracting the base layer from the input image. And the detail image can preserve the important details of the input image, such as edges, texture, and shadows, depending on the smoothing degree. In theory, the input image is smoothed more heavily; the details will be preserved more. Figures 2 and 3 illustrate the phenomenon.

On the other hand, facial feature and shadow details are very important to face photo sketch. The difference between sketches and photos mainly lies in two aspects: texture and shape, which are often exaggerated by the artist in sketch. The texture contains hair texture and shadow texture [2]. In this paper, we perform good shading effects near the interest features from detail images by two-scale image decomposition, proposed by Durand and Dorsey [19]. And the shape of obvious facial features is also obtained from the detail images.

The two-scale decomposition is performed on the logs of pixel intensities using piecewise-linear bilateral filtering and subsampling. On the one hand, the use of logs of intensities is because image can be considered as a product of reflectance and illumination component. So the decomposition can be viewed as an image separated into intrinsic layers of reflectance and illumination [20–22], while base layer corresponds to illumination component and detail corresponds to reflectance [23]. Therefore, we can obtain the facial feature from the detail layer because of the distinct reflectance difference in facial feature and skin region. In fact, human vision is mostly sensitive to the reflectance rather than to the illumination conditions. Even more, the logarithm function deals with the low intensity far better than those high-intensity pixels because of its function character. On the other hand, the piecewise bilateral-linear filtering in the intensity domain and a subsampling in the spatial domain can efficiently accelerate bilateral filtering.

As is described in (2.1), when the scale σ_s of the spatial kernel or/and the scale σ_r of intensity domain increases, the input image will be smoothed more. Although the scale σ_s of the spatial kernel has little influence on the result, it plays an important role in facial feature detection. Several conclusions can be observed. (1) For the same small scale σ_r of intensity domain, the increase of scale σ_s of the spatial domain will smooth more pixels near the edges of the facial feature, while it results in the heavier shadow in detail layer without change in edges. The results are shown in Figure 2 by using different spatial scales. (2) For the same scale σ_s of spatial domain, the increase of scale σ_r of the intensity domain will directly smooth the input image more than (1), because the larger intensity scale will smooth more pixels on the edges [24, 25]. Figure 3 illustrates the different results with the intensity-scale changes.

Further research demonstrates that facial feature, such as hair, eyebrows, eyes, mouth, and nose, which have low intensity than skin, can be obtained in detail image by two-scale decomposition using piecewise bilateral filtering. With the small scale σ_r in intensity domain and the larger scale σ_s in spatial domain, we can get more reflectance component responses of the low contrast area, such as the shadow near the nose and the chin, as shown in Figure 2. While using the small scale σ_s in spatial domain and larger scale σ_r of intensity domain, lower intensity pixel in small area regions can appear in detail image, including eyes and eyebrows, as shown in Figure 3. When the scales in both intensity and spatial domain are set to be largest values, we can get the lowest intensity region, such as black hair, as shown in Figure 4. However, it is noticeable that this works only for the dark hair with fuscous color. To create good hair for arbitrary color, we propose a new method in Section 3. In addition,
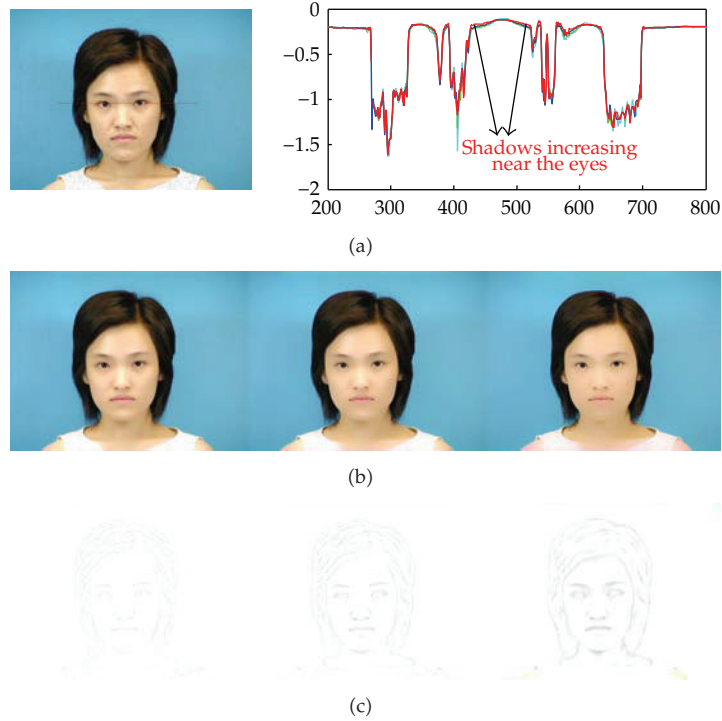
(a)

(b)

(c)

Figure 2: The influence of fixed intensity scale and varied spatial scale in decomposition. Row 1 is input image with scanline, and the scanplots of the log intensity along a row of input images (cyan) were represented using a floating point. Green, blue, and red are the corresponding scanplots of log intensity along the second row. Rows 2 and 3 are the decompositions of input images using different spatial scale, while row 2 is base images and row 3 is the detail images. From left to right, σ_s = 1%, 2%, and 5% the size of the input image, respectively, while σ_r = 0.08.
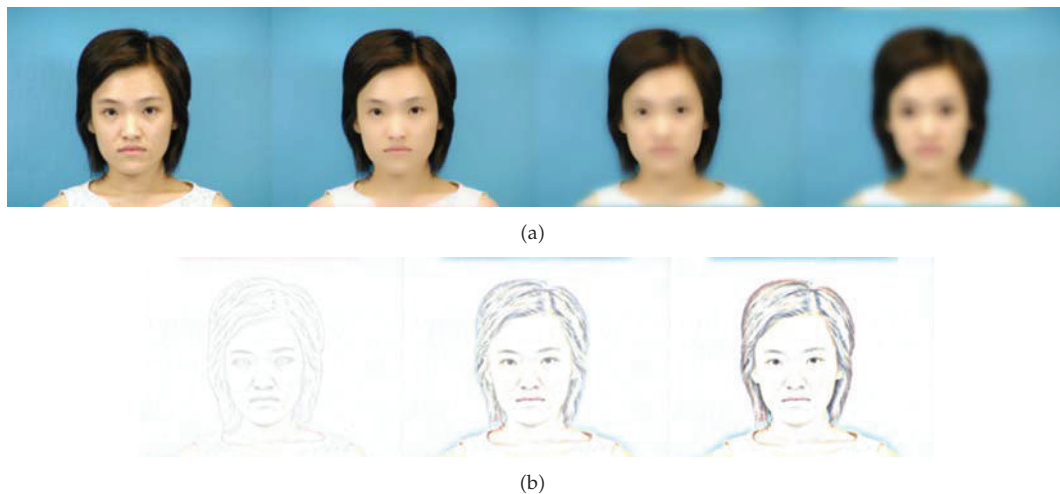

(a)

(b)

Figure 3: The influence of fixed spatial scale and varied intensity scale in decomposition. Row 1 is the base image obtained using piecewise bilateral filtering, and row 2 is the detail image. From left to right, σ_r = 0.08, 0.35, and 0.95, respectively, while σ_s equals 2% the size of the input image.
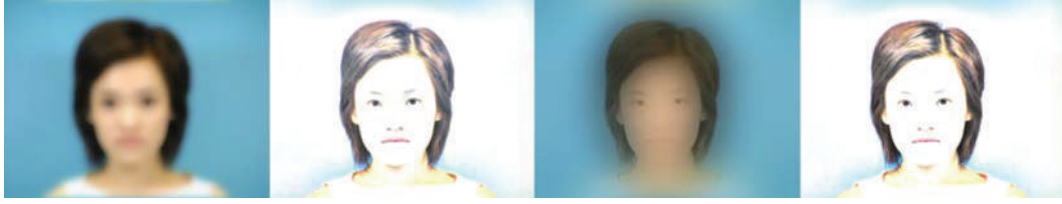
Figure 4: Two-scale decomposition with large spatial scale and large intensity scale. The input image is the same with Figure 3. From left to right, $\sigma_s = 10\%$, while $\sigma_r = 0.60$ and 0.95, respectively.
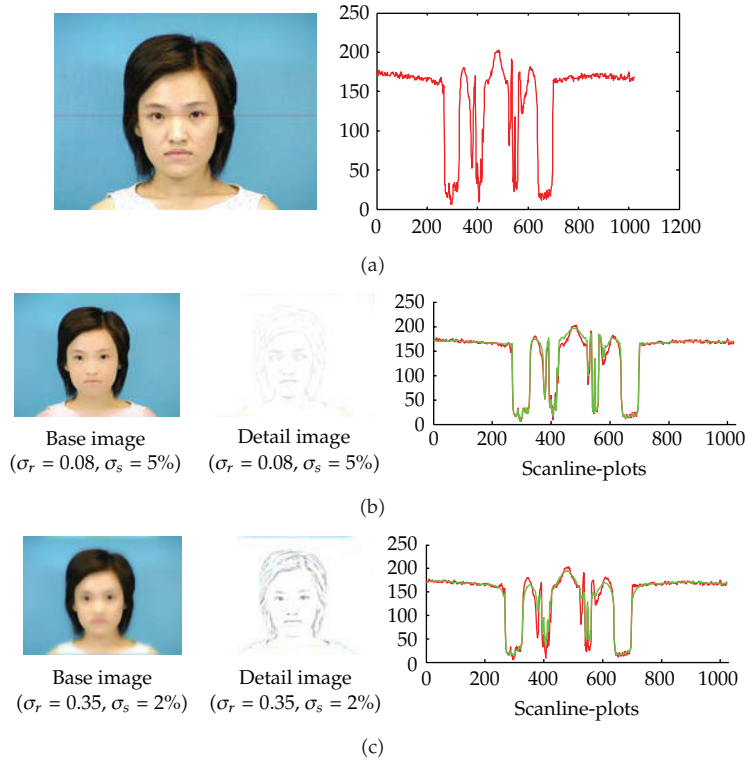

(a)

Base image	Detail image
$(\sigma_r = 0.08, \sigma_s = 5\%)$	$(\sigma_r = 0.08, \sigma_s = 5\%)$

Scanline-plots

(b)

Base image	Detail image
$(\sigma_r = 0.35, \sigma_s = 2\%)$	$(\sigma_r = 0.35, \sigma_s = 2\%)$

Scanline-plots

(c)

Figure 5: The input intensities (red) along a row and its decomposition into base (green) and detail by bilateral filtering. Row 1 is the input image and its scanline plots of intensity. Row 2 and row 3 are the base image and detail image using different scales of intensity and spatial domain. Row 2 shows that weak edges and reflecting details are preserved in detail image at large spatial scale and small intensity scale. Row 3 shows that the sharp edge of facial features is preserved in detail image at small spatial scale and large intensity scale.

because of the highlight on some human mouth, especially when its color and intensity are similar to skin on human face, the mouth will not be extracted at this step, and unapparent mouth extraction is dealt with in Section 3.

In this paper, we define the two-scale decomposition of input image as follows [16]:

$$R_{\mathrm{MSR}_i}(x, y) = \sum_{n=1}^{N} w_n \cdot \left\{ \left(\log(I_i(x, y)) - \log\left[F_n(x, y) * I_i(x, y) \right] \right) \right\}. \tag{2.3}$$

(a) Input (b) Skin color similarity (c) Hair color similarity (d) Multiply enhancement
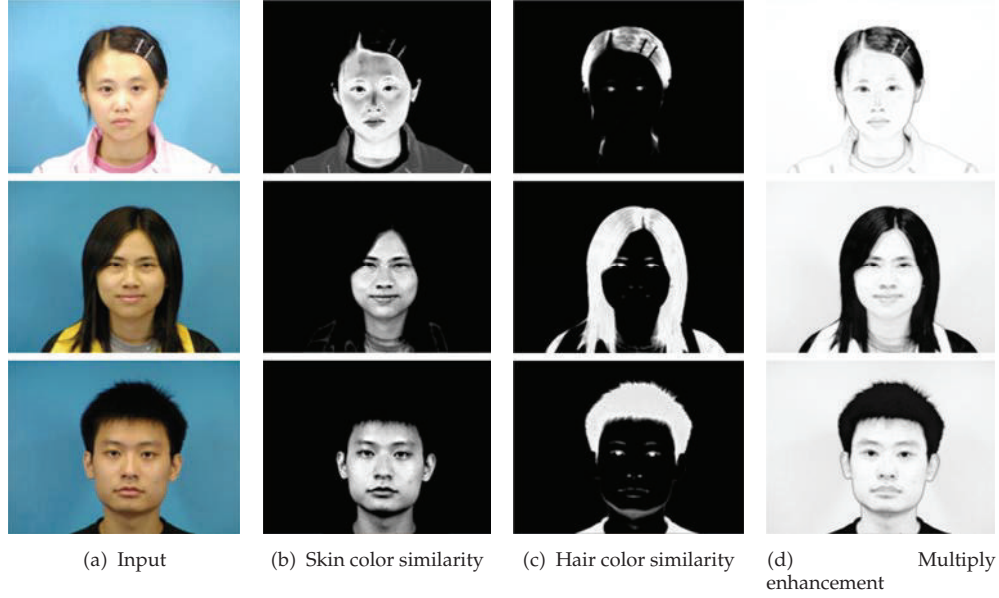
Figure 6: Hair and unapparent facial feature creating.

R_{MSR_i} is the ith color component of the MSR output, $i \in (R, G, B)$, N is the number of scales, w_n is the weighting factor for the nth scale, $I_i(x, y)$ is the image distribution in the ith color band, "$*$" denotes the convolution operation, and $F_n(x, y)$ is the weighting function in the nth bilateral filtering, that is, the $F_n(x, y)$ is given by,

$$F_n(x, y) = f(p - q)g(I_p - I_q) = \exp\left[\frac{-\left(\|p - q\|^2\right)}{2\sigma_{s_n}^2}\right] \cdot \exp\left[-\frac{(I_p - I_q)^2}{2\sigma_{r_n}^2}\right]. \qquad (2.4)$$

So the base image is the output of bilateral filtering, and the detail image D is

$$D(x, y) = \min\{R_{\mathrm{MSR}_i}(x, y)\}. \qquad (2.5)$$

As a conclusion, to ensure the speed and effect, we set the relative scale in two-scale decomposition as follows. (1) For creating the shadows near the facial features, the small intensity scale σ_r is set to constant value of 0.05~0.08, and the associated spatial scale σ_s is set to constant value of 5% the image column size. (2) For creating the clear facial features, including eyes, eyebrows, and sometimes mouth, the large intensity scale σ_r is set to a constant value of 0.35, and the spatial scale σ_s is set to constant to a value of 2% the image column size. Experimental results demonstrate that the above fixed-scale values perform consistently well for all our face images. The results and analysis process are shown in Figure 5.
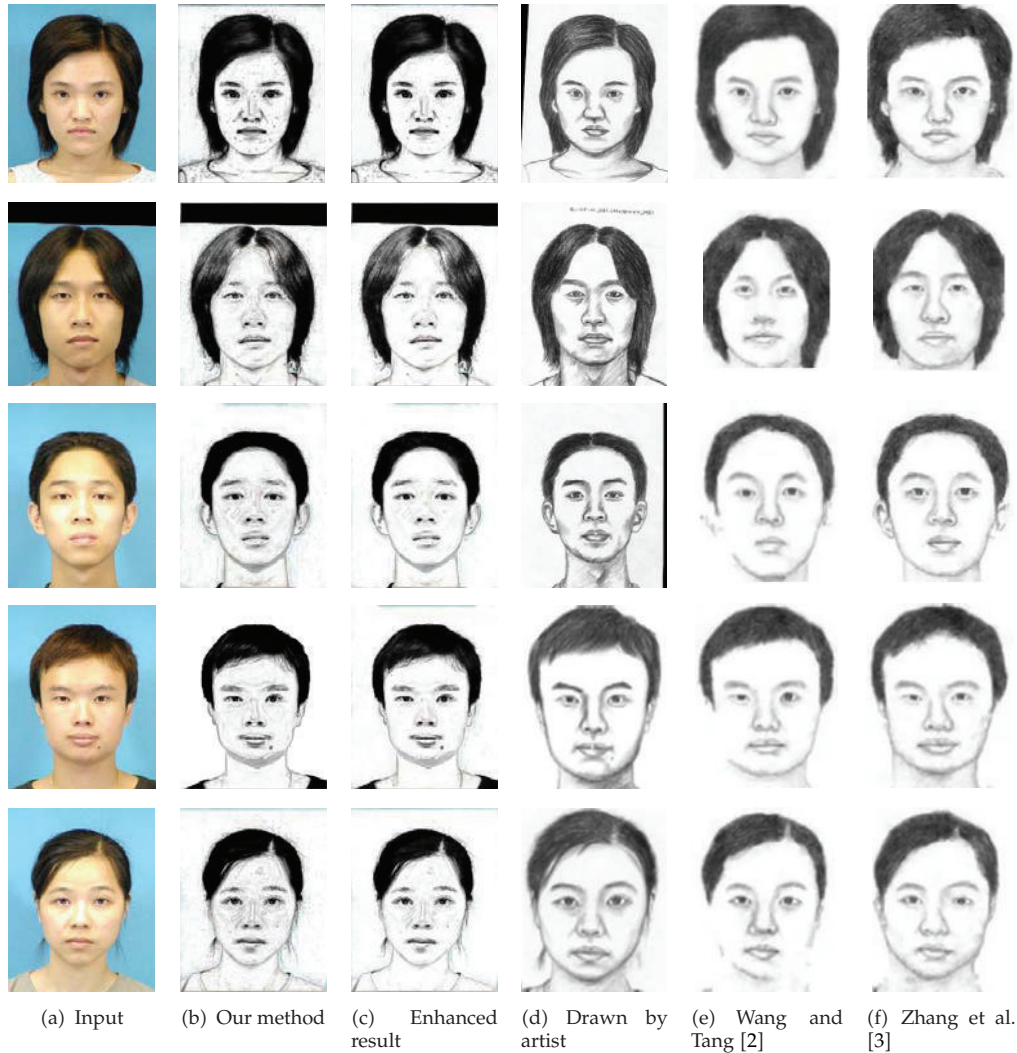
| (a) Input | (b) Our method | (c) Enhanced result | (d) Drawn by artist | (e) Wang and Tang [2] | (f) Zhang et al. [3] |

Figure 7: Face photo-sketch synthesis results: (a) the background of the cropped images in CUHK database is simple; (b) our method illustrates the hair and facial feature more accurately, especially the profile of the face; (c) in order to get the better effect, we smoothed (b) by simple enhancing method; (d) the artist draws the face sketches with some exaggerations, such as nose bridge and mouth edge; (e) in the synthesis results of Tang's method, obviously, face profile is defective, and some of the important marks, such as the mole in the fourth man's face, are lost; (f) Zhang's method draws the facial features, eyes, and mouth and so forth, with some distortion in size and shape.

3. Hair and Facial Feature Creating

3.1. Color Similarity Map

In this section, we discuss two problems: (1) computing color similarity map for input image, which is used to select the skin region and hair region separately and (2) hair and unapparent facial feature creating.

To detect the skin color region, we propose a skin/hair classification method based on color similarity. A Gaussian similarity measuring function is defined to compare
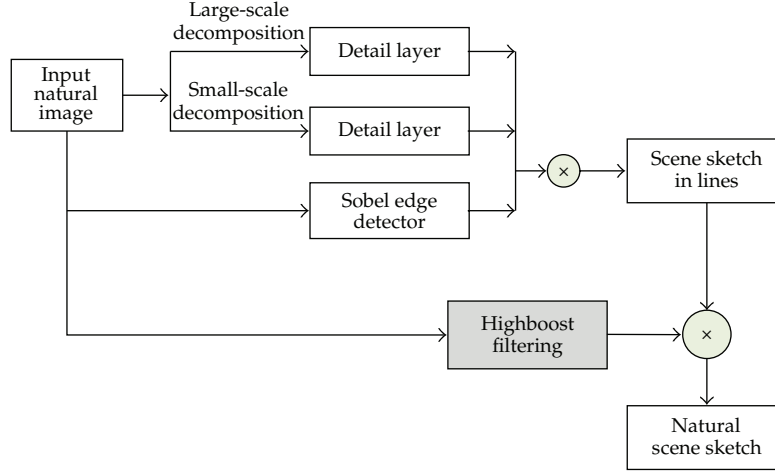
Figure 8: Flowchart of natural scene sketch.

the similarity between two colors. Gaussian convolution performs directly on every pixel of the image in the same way, which achieves the similarity of all the pixels to specific skin color by the concept of color difference to compute it. If the difference between current pixel's color with the specific color is larger, the probability of this pixel of skin/hair is lower. Let $G(x)$ denote the Gaussian masks, and E_b is the specified color (known color), color similarity, function can be defined as

$$g(E_p) = e^{-\|E_p - E_b\|^2 / 2\sigma_c^2}, \tag{3.1}$$

E_p is the color of pixel p in CIE Lab color space, and $\|E_p - E_b\|^2$ is the color difference between the specified color and the pixel p in input image, whose value is determined by CIELAB color difference equation. σ_c is the threshold of color difference, which determines whether the current pixel belongs to the same kind as the known color. Generally, we keep σ_c constant to a value of 30.

To exactly compute the similarity of any color to the specified color, the first step is to confirm a benchmark color. It can be set by two methods: specified by user interaction or the program produces an average color automatically as benchmark while using a known color to select a certain percent of the most similar pixels. The time complexity of this method on confirming a skin color is lower than any face detection-based algorithm. Consequently, hair similarity map and skin similarity map can be achieved by different benchmark colors.

3.2. Hair and Unapparent Facial Feature Creating

After the skin color similarity computing, we can get the color similarity map, as shown in Figure 6. Because the color and intensity of skin distribute uniformly, the gradient is very small in skin region, while the gradient is large in hair region due to the irregular hair colors. In hair color similarity map, the hair region's value is the largest. Then the product image of gradient skin color similarity map and the negative of the hair color similarity map will strengthen the hair region by hair value minimization.
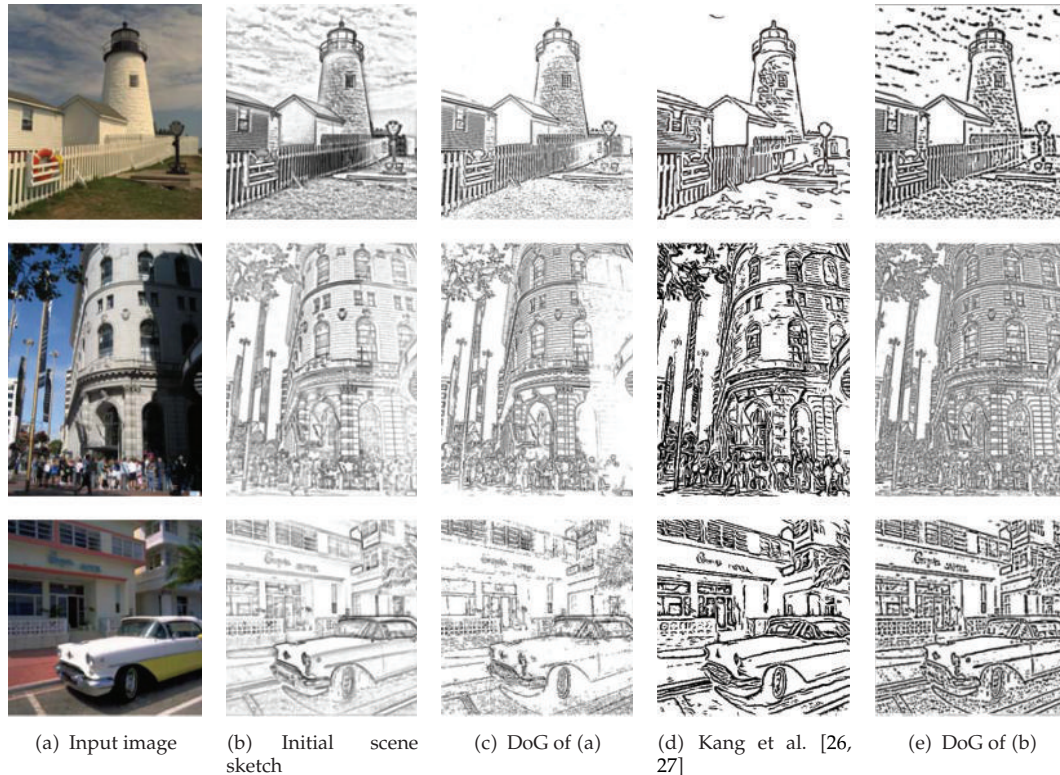
(a) Input image   (b) Initial scene sketch   (c) DoG of (a)   (d) Kang et al. [26, 27]   (e) DoG of (b)

Figure 9: Comparisons of our line extraction based on initial scene sketch with existing methods: difference of Gaussian, Kang's line drawing [26, 27].

The hair region enhancement is realized by multiplying the gradient edge map of skin color similarity and the hair color similarity map, which is proposed in Section 3.1. Since facial feature regions, such as nose and mouth regions, have obvious color difference, they will be also enhanced by the multiplication. On the other hand, the color of eyes and eyebrows either is similar to hair color or is distinct from skin color. Both of the cases will be enhanced by the above operation. Figure 6 shows the process of hair and facial feature creating. The proposed method can extract good hair texture under different lighting conditions. It is noticeable that the hair creating method can be applied to other scenes, for example, image abstraction, cartoon making, and wig wearing, in which hair region is necessary.

4. Experimental Results

We tested the two-scale decomposition-based face sketch synthesis framework on CUHK face photo-sketch database, which contains 606 faces totally. All the input images' size is 1024 × 768 pixels. On average for the uncropped CUHK student face image, one decomposition of an input image for RGB component took about 15.5 seconds (24 s with large spatial scale and 6.9 s with small spatial scale) using two-scale image decomposition based on piecewise bilateral filtering. And the time for hair and unapparent feature creating based on color similarity computing took about 3 seconds. So the total synthesis time is
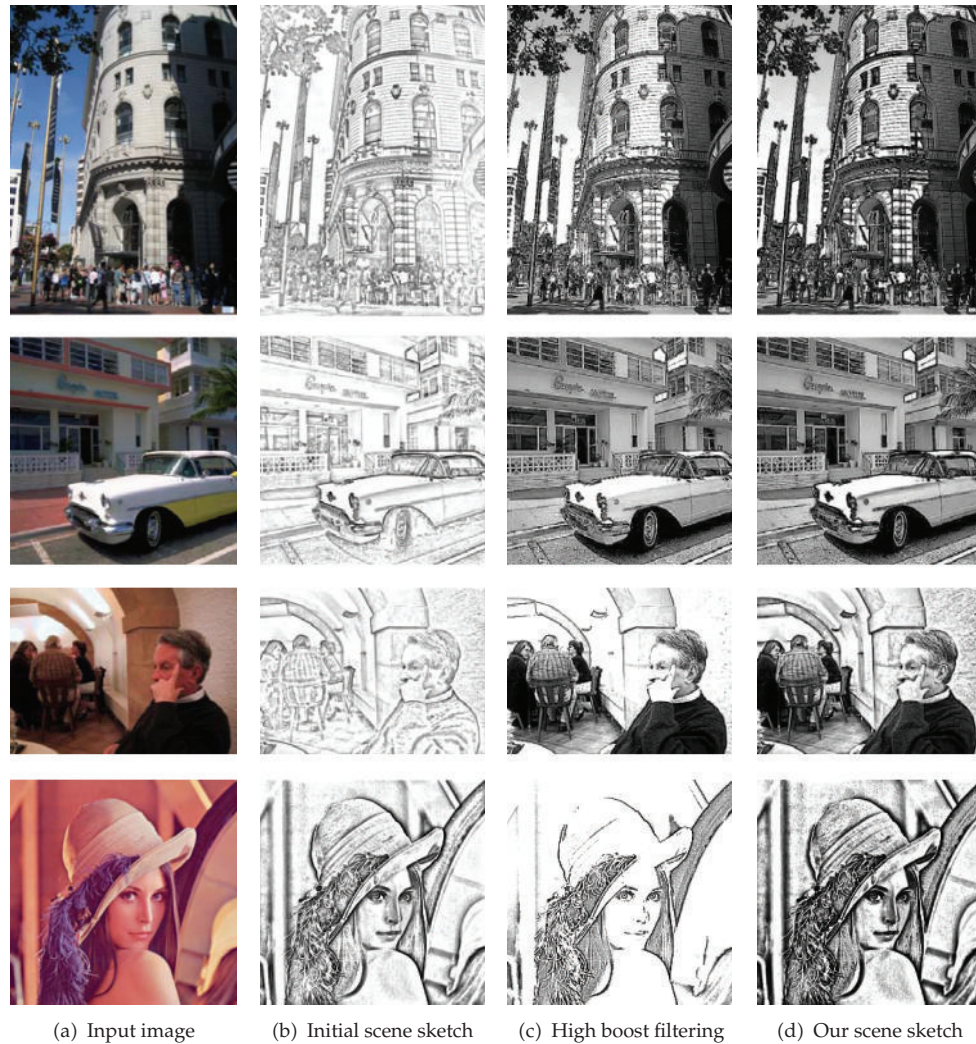
(a) Input image (b) Initial scene sketch (c) High boost filtering (d) Our scene sketch

Figure 10: Arbitrary scene sketch results: our method sketches the images vividly and fast. And the test images in the second and third rows were cited in DêCarlo's paper [28].

Table 1: The computing speed of the proposed method and others.

Method	Wang and Tang [2]	Zhang et al. [3]	The proposed
Speed	>3 mins	>2 mins	<35 s

less than 35 seconds. In fact, if the decomposition is for gray image, the corresponding time consumption will be reduced to about 30 percent. All the above experiments were implemented on a 2.66 G PC. If realized on the cropped image of CUHK (URL: http://mmlab.ie.cuhk.edu.hk/facesketch.html), whose size is $250 * 200$, the time cost is 2.4 seconds for two-scale decomposition and 0.45 seconds for hair and unapparent feature creating. The speed comparison is shown in Table 1. Figure 7 shows the experimental results on CUHK cropped images.
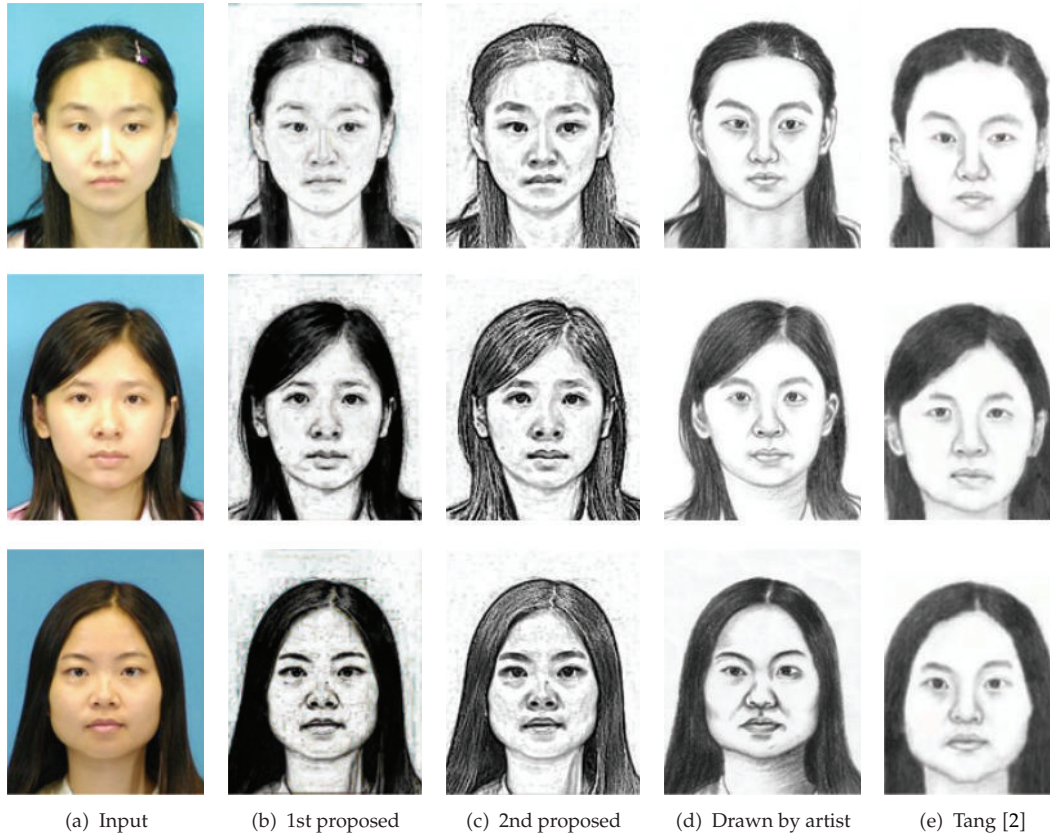
| (a) Input | (b) 1st proposed | (c) 2nd proposed | (d) Drawn by artist | (e) Tang [2] |

Figure 11: Comparisons of our human face sketch results with the artist and Tang's approaches. The first proposed method is introduced in Figure 1, while the second method is in Figure 8. Both the presented methods in this paper successfully depict the facial feature without distortion.

5. Extended Applications

5.1. Natural Scene Sketch and Line Extraction

The method we proposed can be applied in natural scene sketch and line extraction, with a little modification. The modified framework is shown in Figure 8. While illustrating the detail by two-scale bilateral filtering, which is the same with proposed framework, high boost filtering [29] in Figure 8 is used to enhance highlights and shadows in image. Sobel detector can give prominence to the distinct edges. Both the computation cost of high boost filtering and Sobel detector are small, so the natural scene sketch is very fast.

When multiplying the Sobel edge by the two detail layers, we can get the initial natural scene sketch, which has a line look, as shown in Figure 9(b). Then, we can extract the lines using Difference of Gaussian (DoG). Without human skin similarity computing and human hair extraction, the line drawing process is speeded up. The results are shown in Figure 9. Our method performs well on thick subtle edges in input images. Although Kang's method depicts the edges with smooth and coherent lines [26, 27], its speed is very low because of LIC. In addition, it is hard to detect the dense edges in some region, such as the edges in the building in the second image. On the other hand, DoG operator on the input image fails to deal with the edges in shadows and details.

(a) Input image   (b) Scene sketch   (c) DoG of (b)   (d) Image abstraction

Figure 12: Image abstraction based on our line extraction

Table 2: The computing speed of the proposed line extraction method and other.

Image size	Our line extraction method	Kang et al. [26]
512∗512	3.8 s	13.4 s
256∗256	0.64 s	3.3 s

Our line extraction approach takes less than 0.7 second for a $256 * 256$ image to synthesize a sketch. We implement our method using MATLAB and run the codes on a computer with 2.20 GHz CPU. The speed comparison is shown in Table 2.

If we want to get the better sketch of any input images, high boost filtering is preferred. Our approach operates well on any kind of images, such as outdoor natural scenes, animals, plants, building, and human faces. Some of the sketch results are shown in Figure 10. Figure 11 shows the human face sketch results of the CUHK database, which is introduced in Section 4. And all the parameters in bilateral filtering are the same with Section 2.2. The performance of our approach is well when it is used in human face. It can deal with gray images and color images of different races and different lighting conditions. More results on other kinds of images are shown in the appendix.

5.2. Image Stylization

Based on the initial line sketch, we can easily extract the edges by DoG operator. In combination with base layer in two-scale decomposition after color quantization, we can get the simple image abstraction as proposed in [27, 28]. The results are shown in Figure 12.
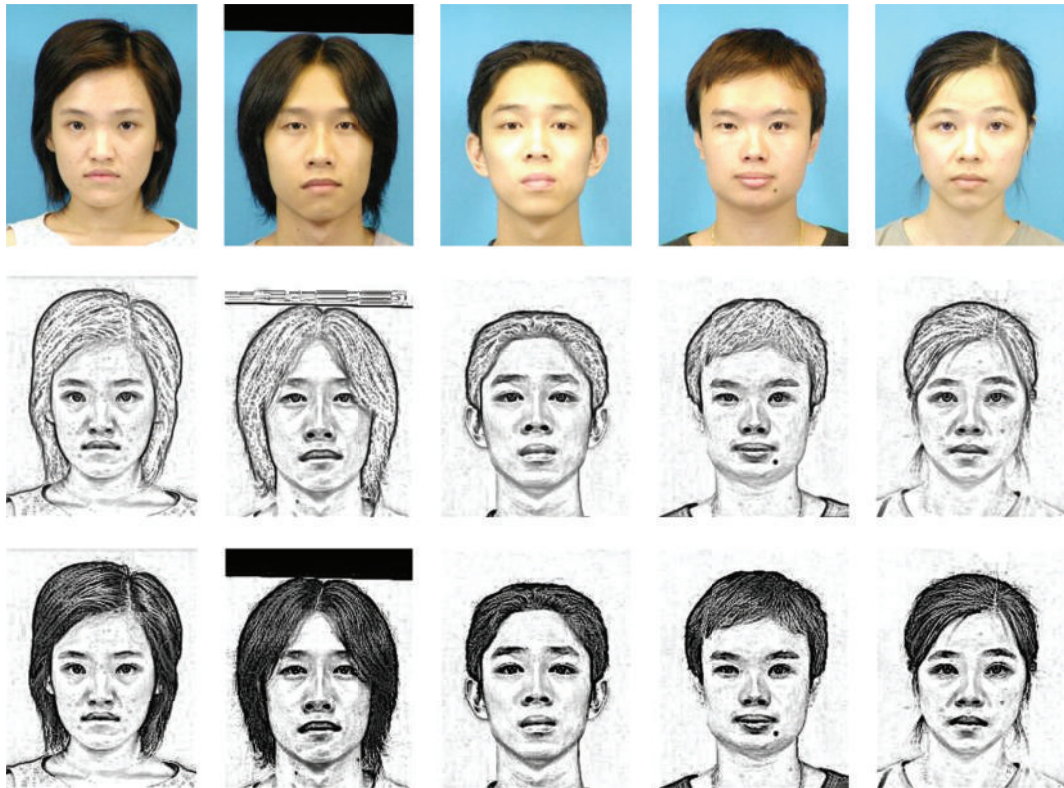
Figure 13: The 1st row is input human face. The 2nd row is initial face sketch results in lines by the proposed approach in Section 5. The 3rd row is the final sketches by the proposed approach in Section 5.

Figure 14: Some sketch results of Olivetti-Att-ORL gray-scale database, which is built for face identification. The first row and the second row are input images and the corresponding sketch results. The ORL database can be downloaded from http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase .html.

Figure 15: Some sketch results of MUCT color face database, which contains human faces of different races and different lighting conditions. The 1st row and the 2nd row are input images, and sketch results, respectively. And the database can be downloaded from http://code.google.com/p/muct/downloads/list. The input image size is 640∗480.
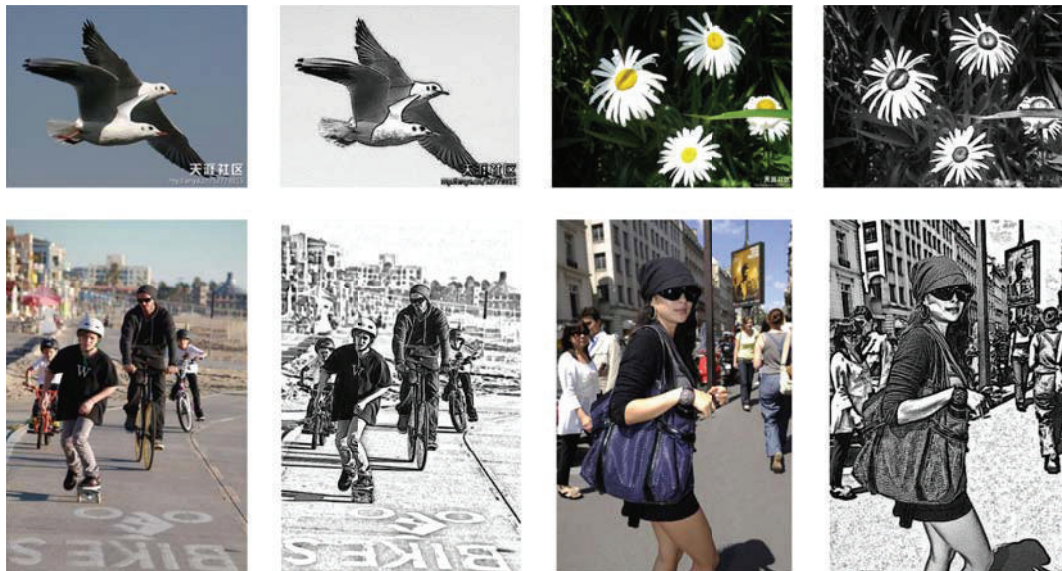

Figure 16: Some sketch results of images under varying lighting conditions downloaded from Internet. The first and the third columns are input images and the other columns are sketch results, respectively.

6. Discussion and Conclusion

We have presented a novel framework for human face photo-sketch synthesis, which is very simple, fast and requires no parameter setting. Firstly, by combing the two-time two-scale image decomposition results, the detail reflecting conditions of facial features and the prominent facial features are obtained. Secondly, based on the color similarity map, we extracted the unapparent facial features and created the vivid hair region. Finally, we
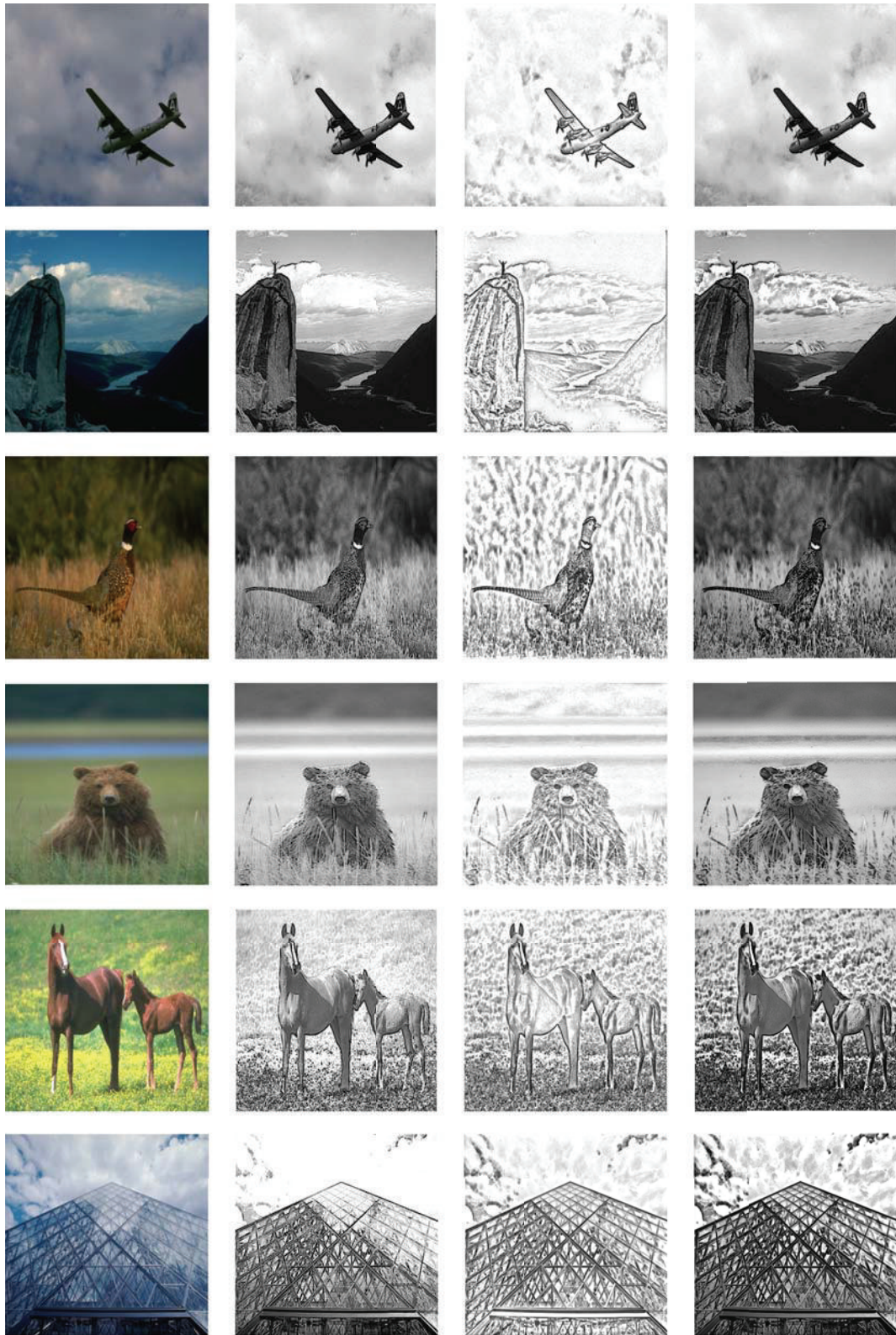
Figure 17: Continued.

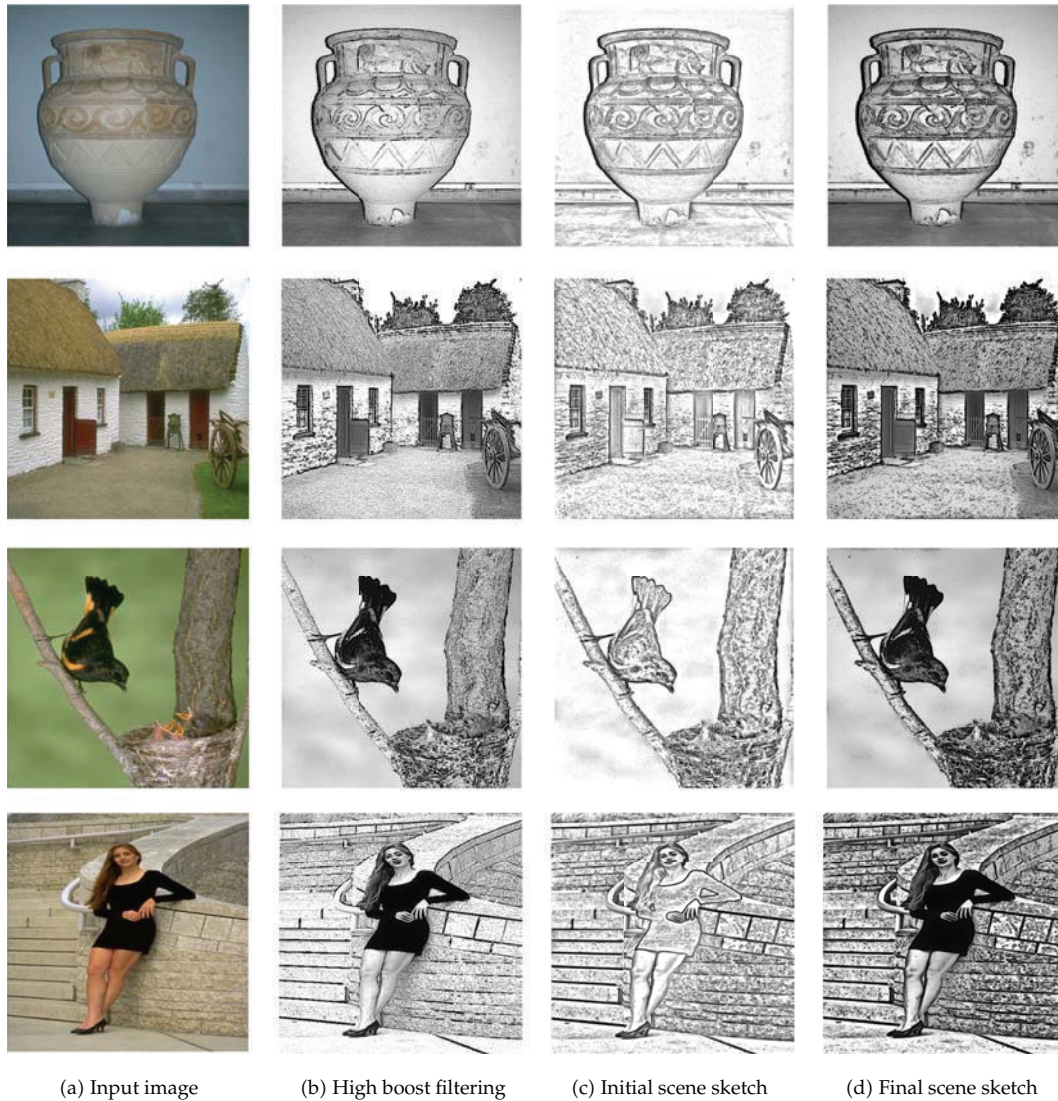| (a) Input image | (b) High boost filtering | (c) Initial scene sketch | (d) Final scene sketch |

Figure 17: Other natural scene sketches of images downloaded from the Berkeley image segmentation database [30].

exploited the framework by the simple addition operation of the former results, which can be applied to other races, such as the white race. Moreover, the framework can also be expanded to other applications, such as natural scene sketch and line extraction. In conclusion, the proposed framework is very simple in that no feature localization algorithm, complex mathematic model, or iteration is needed. In addition, the sketch synthesis result is more vivid than other methods, especially in hair texture creating. Most importantly, the method is easy to be used for other applications, such as line extraction, natural scene sketches, and image abstraction. Sketch recognition is more and more widely used in Sketch-based user interfaces [31]. Since the predominant forces in line extraction of our sketch method, we will try to implement recognition of face photo sketch or other image sketch.

Appendix

See Figures 13, 14, 15, 16, and 17.

Acknowledgment

References

[1] Uhl R. G. and N. D. V. Lobo, "A Framework for recognizing a facial image from a police sketch," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 586–593, 1996.

[2] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, 2009.

[3] W. Zhang, X. Wang, and X. Tang, "Lighting and pose robust face sketch synthesis," in *Proceedings of the European Conference on Computer Vision (ECCV '10)*, pp. 420–433, 2010.

[4] H. Chen, Z. Liu, C. Rose, Y. Xu, H.-Y. Shum, and D. Salesin, "Example-based composite sketching of human portraits," in *Proceedings of the 3rd International ACM Symposium on Non-Photorealistic Animation and Rendering (NPAR '04)*, pp. 95–153, June 2004.

[5] Z. J. Xu, H. Chen, and S. C. Zhu, "A high resolution grammatical model for face representation and sketching," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition 2 (CVPR 2 ' 05)*, vol. 2, pp. 470–477, 2005.

[6] X. Gao, J. Zhong, J. Li, and C. Tian, "Face sketch synthesis algorithm based on E-HMM and selective ensemble," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 4, pp. 487–496, 2008.

[7] X. Tang and X. Wang, "Face sketch synthesis and recognition," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 1, pp. 687–694, 2003.

[8] F. Min, J. L. Suo, S. C. Zhu, and N. Sang, "An automatic portrait system based on and-or graph representation," in *Proceedings of the International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR '07)*, pp. 184–197, August 2007.

[9] Z. Xu, H. Chen, S. C. Zhu, and J. Luo, "A hierarchical compositional model for face representation and sketching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 955–969, 2008.

[10] C.-T. Tu and J.-J. Lien, "Automatic location of facial feature points and synthesis of facial sketches using direct combined model," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 40, no. 4, pp. 1158–1169, 2010.

[11] S. Y. Chen, H. Tong, C. Cattani et al., "Markov models for image labeling," *Mathematical Problems in Engineering*, vol. 2012, Article ID 814356, 18 pages, 2012.

[12] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 839–846, January 1998.

[13] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," in *Proceedings of the European Conference on Computer Vision (ECCV '06)*, pp. 568–580, 2006.

[14] B. Weiss, "Fast median and bilateral filtering," in *Proceedings of the ACM Siggraph*, pp. 519–526, July 2006.

[15] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965–976, 1997.

[16] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Proceedings of the International Conference on Image Processing*, vol. 3, pp. 1003–1006, 1996.

[17] K. Barnard and B. Funt, "Investigations into multi-scale retinex," in *Proceedings of the Colour Imaging in Multimedia*, pp. 9–17, Derby, UK, March 1998.

[18] J. Kuang and D. F. Mark, "iCAM06, HDR, and image appearance," in *Proceedings of the IS& T/SID 15th Color Imaging Conference*, Albuquerque, NM, USA, 2007.

[19] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high dynamic-range image," in *Proceedings of the ACM Transactions on Graphics, ACM Siggraph (ACM '02)*, pp. 257–266, 2002.

[20] W. Zhang, X. Cao, J. Zhang et al., "Detecting photographic composites using shadows," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '09)*, pp. 1042–104, 2009.

[21] B. M. Oh, M. Chen, J. Dorsey, and F. Dorsey, "Image-based modeling and photo editing," in *Proceedings of the ACM Siggraph*, pp. 433–442, 2001.

[22] Q. Liu, X. Cao, C. Deng, and X. Guo, "Identifying image composites through shadow matte consistency," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, 2011.

[23] J. Tumblin and G. Turk, "LCIS: a boundary hierarchy for detail-preserving contrast reduction," in *Proceedings of the Computer Graphics Proceedings (SIGGRAPH '99)*, A. Rockwood, Ed., Annual Conference Series, pp. 83–90, Addison Wesley Longman, Los Angeles, Calif, USA, 1999.

[24] S. Y. Chen and Y. F. Li, "Determination of stripe edge blurring for depth sensing," *IEEE Sensors Journal*, vol. 11, no. 2, pp. 389–390, 2011.

[25] S. Y. Chen, H. Tong, Z. Wang, S. Liu, M. Li, and B. Zhang, "Improved generalized belief propagation for vision processing," *Mathematical Problems in Engineering*, vol. 2011, Article ID 416963, 12 pages, 2011.

[26] H. Kang, S. Lee, and C. Chui, "Coherent line drawing," in *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering (NPAR '07)*, pp. 43–50, August 2007.

[27] H. Kang, S. Lee, and C. K. Chui, "Flow-based image abstraction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 1, pp. 62–76, 2009.

[28] D. DeCarlo and A. Santella, "Stylization and abstraction of photographs," in *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques (ACM SIGGRAPH '02)*, pp. 769–776, 2002.

[29] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, Publishing House of Electronics Industry, Beijing, China, 2nd edition, 2002.

[30] http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/ .

[31] R. Arandjelovic and T. M. Sezgin, "Sketch recognition by fusion of temporal and image-based features," *Pattern Recognition*, vol. 44, no. 6, pp. 1225–1234, 2011.