Lighting and Pose Robust Face Sketch Synthesis

Wei Zhang[1], Xiaogang Wang[2], and Xiaoou Tang[1,3]

[1] Department of Information Engineering, The Chinese University of Hong Kong
{zw009,xtang}@ie.cuhk.edu.hk
[2] Department of Electronic Engineering, The Chinese University of Hong Kong
xgwang@ee.cuhk.edu.hk
[3] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

Abstract. Automatic face sketch synthesis has important applications in law enforcement and digital entertainment. Although great progress has been made in recent years, previous methods only work under well controlled conditions and often fail when there are variations of lighting and pose. In this paper, we propose a robust algorithm for synthesizing a face sketch from a face photo taken under a different lighting condition and in a different pose than the training set. It synthesizes local sketch patches using a multiscale Markov Random Field (MRF) model. The robustness to lighting and pose variations is achieved in three steps. Firstly, shape priors specific to facial components are introduced to reduce artifacts and distortions caused by variations of lighting and pose. Secondly, new patch descriptors and metrics which are more robust to lighting variations are used to find candidates of sketch patches given a photo patch. Lastly, a smoothing term measuring both intensity compatibility and gradient compatibility is used to match neighboring sketch patches on the MRF network more effectively. The proposed approach significantly improves the performance of the state-of-the-art method. Its effectiveness is shown through experiments on the CUHK face sketch database and celebrity photos collected from the web.

1 Introduction

Automatic face sketch synthesis has drawn a great deal of attention in recent years [1][2][3][4][5] due to its applications in law enforcement and digital entertainment. For example, in law enforcement, it is useful to develop a system to search photos from police mug-shot databases using a sketch drawing when the photo of a suspect is not available. By transferring face photos to sketches, intermodality face recognition is made possible [2]. In the movie industry, artists can save a great amount of time on drawing cartoon faces with the assistance of an automatic sketch synthesis system. Such a system also provides an easy tool for people to personalize their identities in the digital world, such as through the MSN avatar.

Computer-based face sketch synthesis is different from line drawing generation [7][8]. Line drawings without texture are less expressive than sketches with both contours and shading textures. Popular sketch synthesis methods are mostly

K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part VI, LNCS 6316, pp. 420–433, 2010.

(a) (b) (c) (d)

Fig. 1. Examples of synthesized sketches from web face photos. (a) Test photos; (b) Sketches synthesized by [5]; (c) Sketches synthesized by [5] with luminance remapping [6]; (d) Sketches synthesized by our method. Note that luminance remapping refers to zero-mean unit-variance normalization of the luminance channel of all photos in our implementation. This simple technique was found to be better than non-smooth mappings in image style transformation, such as histogram matching/equalization [6]. The results are best viewed on screen.

example-based, which generates a sketch with rich textures from an input face photo based on a set of training face photo-sketch pairs [1][3][4][5]. These approaches can synthesize sketches of different styles by choosing training sets of different styles. Tang and Wang [1] proposed to apply the eigentransform globally to synthesize a sketch from a photo. However, such a global linear model does not work well if the hair region is included, as the hair styles vary significantly among different people. To overcome this limitation, Liu *et al.* [3] proposed patch-based reconstruction. The drawback of this approach is that the patches are synthesized independently, ignoring their spatial relationships, such that some face structures cannot be well synthesized. In addition, face sketch synthesis through linear combinations of training sketch patches causes the blurring effect.

Following this line of work, a state-of-the-art approach using a multiscale Markov random field (MRF) model has been proposed recently [5] and achieved good performance under well controlled conditions (i.e. the testing face photo has to be taken in the frontal pose and under a similar lighting condition as the training set). This approach has some attractive features: (1) it can well synthesize complicated face structures, such as hair, which are difficult for previous methods [1]; (2) it significantly reduces artifacts, such as the blurring and aliasing effects, which commonly exist in the results of previous methods [1][3]. In spite of the great improvement compared with previous methods, this approach often fails if the testing face photo is taken in a different pose or under a different lighting condition (even if the lighting change is not dramatic) than the training set. Some examples are shown in Fig. 1. Due to the variations of lighting and pose, on the synthesized sketches by [5] some face structures are lost, some dark regions are synthesized as hair, and there are a great deal of distortions

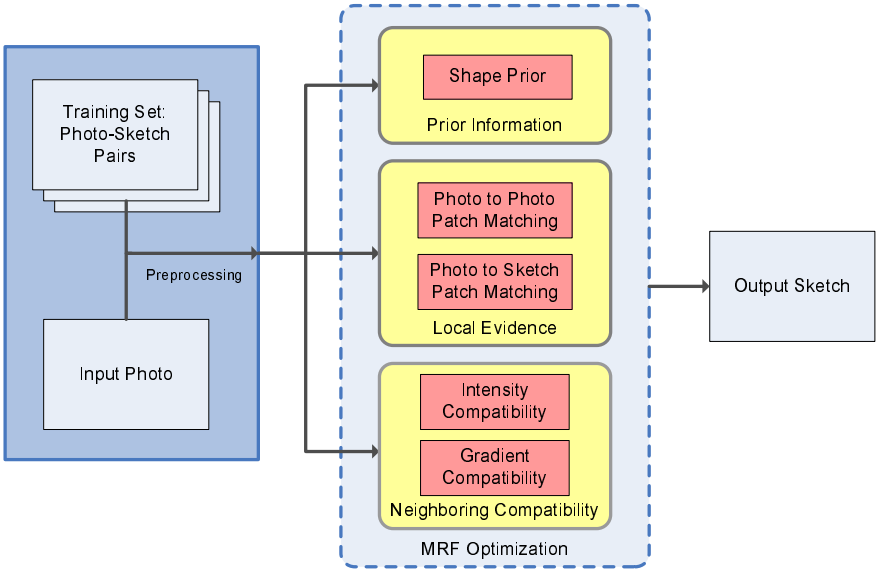


Fig. 2. Illustration of our framework

and artifacts. This is also a serious problem not addressed by other approaches [1][3][4]. It limits their applications to real-world problems.

In face recognition studies, some preprocessing techniques such as histogram equalization, and features such as Local Binary Patterns (LBP) [9], were used to effectively recognize face photos under lighting variations. In the area of nonphotorealistic rendering, luminance remapping was introduced to normalize lighting variations [6]. However, experiments show that simply borrowing these techniques is not effective in face sketch synthesis. See examples in Fig. 1.

In this paper, we address this challenge: *given a limited set of photo-sketch pairs with frontal faces and normal lighting conditions, how to synthesize face sketches for photos with faces in different poses (in the range of* $[-45^o+45^o]$*) and under different lighting conditions.* We adopt the multiscale MRF model whose effectiveness has been shown in face sketch synthesis [5] and many low-level vision problems [10]. In order to achieve the robustness to variations of lighting and pose, some important improvements are made in the design of the MRF model as summarized in Fig. 2. Firstly, a new term of shape priors specific to face components are introduced in our MRF model. It effectively reduces distortions and artifacts and restores lost structures as shown in Fig. 1. Secondly, patch descriptors and metrics which are more robust to lighting variations are used to find candidates of sketch patches given a photo patch. In addition to photo-to-photo patch matching, which was commonly used in previous approaches [3][5], our "local evidence" term also includes photo-to-sketch patch matching, which improves the matching accuracy with the existence of lighting and pose variations. Lastly, a smoothing term involving both intensity compatibility and

gradient compatibility is used to match neighboring sketch patches on the MRF network more effectively.

The effectiveness of our approach is evaluated on the CUHK face sketch database which includes face photos with different lightings and poses. We also test on face photos of Chinese celebrities downloaded from the web. The experimental results show that our approach significantly improves the performance of face sketch synthesis compared with the state-of-the-art method [5] when the testing photo includes lighting or pose variations.

2 Lighting and Pose Robust Face Sketch Synthesis

In this section, we present our algorithm for face sketch synthesis. For ease of understanding, we use the single-scale MRF model in the presentation, instead of the two-scale MRF model in our implementation[1].

2.1 Overview of the Method

A graphical illustration of the MRF model is shown in Fig. 3. A test photo is divided into N overlapping patches with equal spacing. Then a MRF network is built. Each test photo patch x_i^p is a node on the network. Our goal is to estimate the status $y_i = (y_i^p, y_i^s)$, which is a pair of photo patch and sketch patch found in the training set, for each x_i^p. Photos and sketches in the training set are geometrically aligned. y_i^p is a photo patch and y_i^s is its corresponding sketch patch. If patches i and j are neighbors on the test photo, nodes y_i and y_j are connected by an edge, which enforces a compatibility constraint. The sketch of the test photo is synthesized by stitching the estimated sketch patches $\{y_i^s\}$. Based on the MRF model, our energy function is defined in the following form,

$$E(\{y_i\}_{i=1}^N) = \sum_{i=1}^{N} E_L(x_i^p, y_i) + \sum_{i=1}^{N} E_{Pi}(y_i) + \sum_{(i,j)\in\Xi} E_C(y_i^s, y_j^s), \qquad (1)$$

where Ξ is the set of pairs of neighboring patches, $E_L(x_i^p, y_i)$ is the local evidence function (Subsection 2.2), $E_{Pi}(y_i)$ is the shape prior function (Subsection 2.3), and $E_C(y_i^s, y_j^s)$ is the neighboring compatibility function (Subsection 2.4). The shape prior function is specific to face components, which means that different location indicated by i has different E_{Pi}. The above MRF optimization problem can be solved by belief propagation [10] [11].

A MRF model was also used in [5], however, with several major differences with ours. It has no shape prior function which is effective in sketch synthesis. Its local evidence function only computes the sum of the squared differences (SSD) between x_i^p and y_i^p and is sensitive to lighting variations. Our local evidence function uses new patch descriptors which are more robust to lighting variations.

[1] We do find that the two-scale MRF model performs better. The details of multiscale MRF can be found in [5]. However, it is not the focus of this paper.

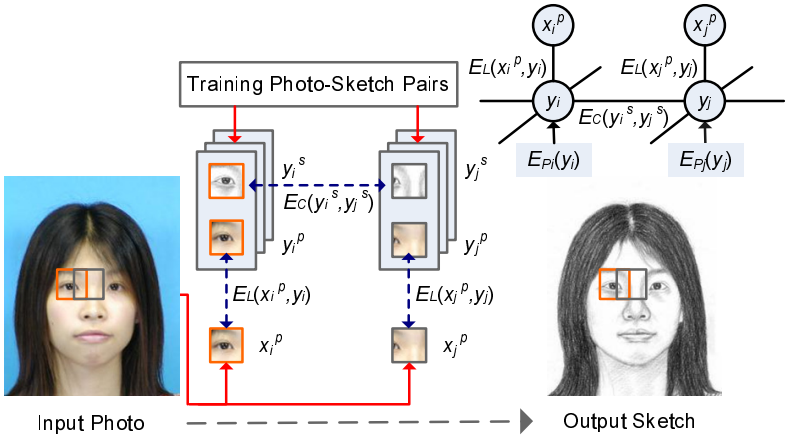


Fig. 3. Illustration of the MRF model for face sketch synthesis

Our method includes not only photo-to-photo patch matching (between x_i^p and y_i^p) but also photo-to-sketch patch matching (between x_i^p and y_i^s) to improve the robustness. The neighboring compatibility function in [5] is to minimize SSD between neighboring estimated sketch patches (y_i^s and y_j^s) in their overlapping region, while ours also minimizes the difference of gradient distributions. Details will be explained in the following subsections.

2.2 Local Evidence

The goal of the local evidence function is to find a sketch patch y_i^s in the training set best matching the photo patch x_i^p in test. However, since photos and sketches are in different modalities, it is unreliable to directly match them. So the training photo patch y_i^p corresponding to a training sketch patch y_i^s is involved. It is assumed that if y_i^p is similar to x_i^p, it is likely for y_i^s to be a good estimation of the sketch patch to be synthesized. We propose to match a testing photo patch with training photo patches and also with training sketch patches simultaneously, i.e. we define the local evidence function as the weighted sum of squared intra-modality distance d_{L1}^2 and squared inter-modality distance d_{L2}^2,

$$E_L(x_i^p, y_i) = d_{L1}^2(x_i^p, y_i^p) + \lambda_{L2} d_{L2}^2(x_i^p, y_i^s), \tag{2}$$

where λ_{L2} is the weight to balance different terms in the energy function E and it is chosen as 2 in our experiments.

Photo-to-Photo Patch Matching. A straightforward choice of E_L is the Euclidean distance between x_i^p and y_i^p as used in [5]. However, it does not perform well when the lighting condition varies. Noticing that most of the sketch contours correspond to edges in the photo, we use a difference-of-Gaussians (DoG) filter to process each photo, i.e. convolving each photo with the difference of two Gaussian

Fig. 4. Compare the results with/without DoG filtering under a normal lighting condition. (a) Test photos which are under the same lighting as the training set. (b)Synthesized sketch by the method in [5] without DoG filtering. (c) Synthesized sketches by our method with DoG filtering. To evaluate the effectiveness of DoG filtering, other parts, such as shape priors and photo-to-sketch patch matching, in our framework are not used in these examples.

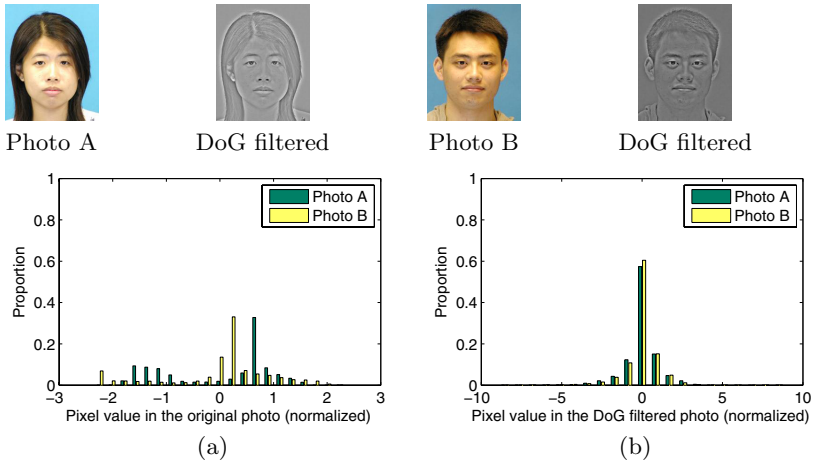


Fig. 5. Examples of DoG filtering with $(\sigma_0, \sigma_1) = (0, 4)$. Photo A is from the training set taken under the normal lighting condition, and Photo B is from the testing set taken under a different lighting condition. The pixel values of DoG filtered photos are scaled to $[0, 1]$ for visualization. (a) Histograms of pixel values of the two photos after luminance remapping. They do not match well. (b) Histograms of pixel values of the two photos after DoG filtering and normalization. They match well.

kernels with standard deviations σ_0 and σ_1, and normalize all pixel values to zero-mean and unit-variance. In our experiments, we find that $(\sigma_0, \sigma_1) = (0, 4)$ or $(1, 4)$ performs the best. DoG filtering has two advantages. First, it can detect and enhance the edges, and thus the synthesized sketch has better facial details. As shown in Fig. 4, even for normal lighting, the DoG filtering can improve facial details. Second, subtracting low-frequency component reduces the effect of lighting variations, e.g. shading effects. The example in Fig. 6 shows that DoG filtering improves synthesized facial details, especially on the nose and the

(a) (b) (c) (d) (e) (f)

Fig. 6. Sequential illustration of the roles of each part in our framework. (a) Test photo under a different lighting condition than the training set; (b) Sketch by the method in [5] with luminance remapping as preprocessing [6]; (c) Sketch by our method with P2P+IC; (d) Sketch by our method with P2P+P2S+IC; (e) Sketch by our method with P2P+P2S+prior+IC; (f) Sketch by our method with P2P+P2S+prior+IC+GC. P2P, P2S, prior, IC and GC represent photo-to-photo patch matching, photo-to-sketch patch matching, shape priors, intensity compatibility and gradient compatibility, respectively. The results are best viewed on screen.

eyebrows, when there are lighting variations. Luminance remapping [6], which normalizes the distribution of pixel values in an image to zero-mean and unit-variance, is commonly used for lighting normalization. However, its improvement is limited in this application. An example is shown in Fig. 5. After luminance remapping, the distributions of pixel values in two photos taken under different lighting conditions still do not match. On the contrary, their distributions after DoG filtering match well. In some cases, photo-to-photo patch matching is not enough and the mismatching problem, such as the hair and profile regions shown in Fig. 6 (c), still exists. Thus, photo-to-sketch patch matching is introduced.

Photo-to-Sketch Patch Matching. The intra-modality distance between photo patches does not always work for selecting a good sketch patch. Similar photo patches under the Euclidean distance may correspond to very different sketch patches. Interestingly, people have the ability to directly match photos with sketches. Inspired by this, we propose to use inter-modality distance between testing photo patches and training sketch patches to enhance the selection ability. As the visual appearances of photo and sketch patches are different, it is difficult to directly match them. However, there exists some similarity of gradient orientations between a photo and its sketch. We choose to use the dense SIFT descriptor [12] from the family of histogram-of-orientations descriptors. Our strategy is to assign each patch a dense SIFT descriptor, and use the Euclidean distance between SIFT descriptors of photo patches and sketch patches as the inter-modality distance. To capture structures in large scales, we extract the descriptors in larger regions than patches. For each patch, we extract a region of size 36×36 centered at the center of the patch (the size of patch is 10×10), and divide it into 4×4 spatial bins of the same size. 8 orientations bins are evenly spaced over 0°-360°. The vote of a pixel to the histogram is weighted by its gradient magnitude and a Gaussian window with parameter $\sigma = 6$ centered at the center of the patch. So the descriptor is 128 dimensional. The descriptor is normalized by its $L2 - norm$, clipped by a threshold 0.2 and renormalized

as reported in [12]. The synthesis result with photo-to-sketch patch matching is shown in Fig. 6 (d). It restores the hair and partial profile lost in Fig. 6 (c).

2.3 Shape Prior

Face images are a special class of images with well regularized structures. Thus shape priors on different face components can be used to effectively improve the synthesis performance. The loss of some face structures, especially the face profile, is a common problem for the patch-based sketch synthesis methods without referring to global structures. When this happens, the contours of some face components are replaced by blank regions. This problem becomes much more serious when there are variations of lighting and pose. See examples in Fig. 1. However, it can be effectively alleviated by using the prior information on different face components to guide the selection of sketch patches. In our approach, a state-of-the-art face alignment algorithm [13] is first utilized to detect some predefined landmarks on both the training sketches and the testing photo. The chosen landmarks locate in regions where loss of structures often happens, especially on the face profile. Shape priors are imposed to these regions but not in other regions. If a landmark f falls into patch i on the test photo, a prior distribution is computed via kernel density estimation,

$$E_{Pi}(y_i) = \lambda_P \ln \left[\frac{1}{\sqrt{2\pi} N_t} \sum_{k=1}^{N_t} \exp\left(-\frac{(\beta(y_i^s) - \beta_{k,f})^2}{h_f^2} \right) \right]. \tag{3}$$

N_t is the number of sketches in the training set. $\beta(y_i^s)$ is some statistic on the sketch patch y_i^s. $\beta_{k,f}$ is the statistic on a sketch patch centered at landmark f in sketch image k. h_f is the bandwidth of landmark f and is set as three times of the standard deviation of $\{\beta_{k,f}\}$. The weight $\lambda_P = 0.01$ is to normalize the metric scale of the shape prior term and the performance of our algorithm is robust to λ_P in a fairly large range.

We test several kinds of patch statistics, such as mean gradient magnitude, variance of pixel values, proportion of edge pixels, and find that mean gradient magnitude performs the best and it is chosen as $\beta(\cdot)$. It can well solve the problem of losing structures, as shown in Fig. 6 (e).

2.4 Neighboring Compatibility

The goal of the neighboring compatibility function is to make the neighboring estimated sketch patches smooth and thus to reduce the artifacts on the synthesized sketch. In our model it is defined as

$$E_C(y_i, y_j) = \lambda_{IC} d_{IC}^2(y_i^s, y_j^s) + \lambda_{GC} d_{GC}^2(y_i^s, y_j^s), \tag{4}$$

where the intensity compatibility term d_{IC}^2 is the SSD in the overlapping region between two neighboring sketch patches y_i^s and y_j^s, and the gradient compatibility term d_{GC}^2 is the squared Euclidean distance between the dense SIFT

descriptors of y_i^s and y_j^s. The intensity compatibility term is for the smoothness of the output sketch. However, only using this term tends to lose some face structures since two blank regions in neighbors have high intensity compatibility. Thus, we further add the gradient compatibility constraint, which requires that the neighboring patches have similar gradient orientations. The use of gradient compatibility can further alleviate the structural loss, an example of which is given in Fig.s 6 (e) and (f) (the region in the red box). We set the weights $\lambda_{IC} = 1$ and $\lambda_{GC} = 0.1$.

2.5 Implementation Details

All the photos and sketches are translated, rotated, and scaled such that the two eye centers of all the face images are at fixed position. We crop the images to 250×200 and the two eye center positions are $(75, 125)$ and $(125, 125)$. All color images are converted to grayscale images for sketch synthesis.

- **Preprocessing on Test Photos.** Empirically, when lighting is near frontal, our algorithm can work well without the preprocessing step. However, for side light, we need to use Contrast Limited Adaptive Histogram Equalization (CLAHE) [14] for preprocessing.[2] We use the setting that the desired histogram shape is Rayleigh distribution (parameter $\alpha = 0.7$).
- **Candidate Selection.** In order to save computational cost, a step of candidate selection as suggested in [10] is used before optimizing the MRF model. For each test photo patch x_i^p, top K ($K = 20$) photo-sketch pairs with the smallest energy of $E_L(x_i^p, y_i) + E_{Pi}(y_i)$ are selected from the training set as candidates. In order to take the advantage of face structures, candidates are searched within a 25×25 local region around patch i instead of in the entire images. The final estimation y_i on node i is selected as one of the K candiates through joint optimization of all the nodes on the MRF network.
- **Two-scale MRF.** We use two-scale MRF with the same setting as in [5]. Patch sizes at the two layers are 10×10 and 20×20, respectively. MAP estimate is used in the belief propagation algorithm [10].
- **Stitching Sketch Patches.** To avoid blurring effect, we use a minimum error boundary cut between two overlapping patches on their overlapped pixels as what is usually done for texture synthesis [15].

3 Experimental Results

We conduct experiments on the CUHK database [5] commonly used in face sketch synthesis research, and a set of celebrity face photos from the web. In all the experiments, 88 persons from the CUHK database are selected for training, and each person has a face photo in a frontal pose under a normal lighting condition, and a sketch drawn by an artist while viewing this photo. In the first

[2] CLAHE improves the method in [5] little and deteriorates its performance in some cases. So we choose to report their results without the preprocessing.

(a) (b) (c) (d)

Fig. 7. Representative results on the baseline set. (a) Test photo; (b) Sketch drawn by the artist while viewing the normal lighting photo; (c) Sketch by the method in [5]; (d) Sketch by our method. The results are best viewed on screen.

experiment, 100 other persons are selected for testing. We have three data sets: the baseline set, the lighting variation set, and the pose variation set. The baseline set includes 100 face photos taken in a frontal pose under the same lighting condition as the training set. The lighting variation data set includes three photos with faces in a frontal pose with three different lightings (dark frontal/dark left/dark right) for each person. And the pose variation set includes two photos with faces in left and right poses (with 45 degrees) under a normal lighting condition for each person. In the second experiment, some face photos of Chinese celebrities with uncontrolled lighting conditions and poses are downloaded from the web.[3] All photos are with a neutral expression. Parameters are fixed throughout the experiments. It takes about 2 minutes to synthesize a sketch running our MATLAB implementation on a computer with 3.20 GHz CPU. Due to the paper length, only a limited number of examples are shown in this paper.

3.1 Lighting and Pose Variations

We first investigate the effect of lighting and pose variations separately on the CUHK database. A preliminary test is on the baseline set. Our algorithm performs as well as the method in [5]. On some photos, our algorithm can produce even better face sketches as shown in Fig. 7. To give a quantitative evaluation of the performance, we test the rank-1 and rank-10 recognition rates when a query sketch synthesized from a test photo is used to match the sketches drawn by the artist. The results are shown in Table 1.[4] Our algorithm slightly beats the previous method by 3%.

[3] The CUHK database cannot be used as a training set for photos of people from other ethnic groups, partially due to the human perception.

[4] Recognition rates cannot completely reflect the viual quality of synthesized sketches. It is used as an indirect measurement to evaluate the performance of sketch synthesis since no other proper quantitative evaluation methods are available.

Table 1. Rank-1 (Rank-10) recognition rates using whitened PCA [16]. The whitened PCA model is trained on the 100 sketches drawn by the artist while viewing the baseline set. It performs better than standard PCA without whitening on all the tasks. The reduced number of dimension is 99, and it is the best for all the tasks.

Testing set	[5]	[5] with LBP	[5] with HE	[5] with LR	Ours
Baseline	96% (100%)	-	-	-	99% (100%)
Front Light	58% (87%)	58% (87%)	70% (95%)	75% (96%)	84% (96%)
Side Lights	23.5% (56%)	25.5% (75.5%)	38% (80.5%)	41.5% (78.5%)	71% (87.5%)

Lighting. Although the previous method performs well on the normal lighting set, their performance degrades dramatically when the lighting changes. Our method performs consistently well under different lighting conditions. To make a fair comparison, we also report the results of [5] with several popular illumination normalization methods, including histogram equalization (HE) and luminance remapping (LR) [6], and with LBP [9], an illumination invariant feature.

On the recognition rate, our method beats all the others, as shown in Table 1. The method in [5] performs very poorly without any preprocessing. LR and HE improve the method in [5], but LBP improves little. LR performs better than HE and LBP. As hair and background are included in face photos, previous illumination normalization methods, such as HE, do not perform well. By converting a patch to its LBP feature, information to distinguish different components, which is important for sketch synthesis, may be lost and thus mismatching often occurs. In addition, we find that dark side lighting conditions are more difficult than dark frontal lighting, and under dark side lightings, our method beats all the others by a large amount on the rank-1 recognition rate.

On the visual quality, LR improves the method in [5], but as shown in Fig.s 8 and 9, the facial details and profile are still much worse than those given by our method. Under dark frontal lighting, their results usually have incorrect blank regions and noisy details. Under dark side lightings, the preprocessing helps only a little as it processes the photos globally. See the failed results shown in Fig. 9.

Pose. To test the robustness of our method to pose variations, we use the pose set with the similar lighting condition as the training set. As shown in Fig. 10, our method performs better than the method in [5].[5] With pose variations, the major problem of the results by [5] is to lose some structures especially on the profile. This problem can be efficiently alleviated by the shape priors, photo-to-sketch patch matching and gradient compatibility designed in our model.

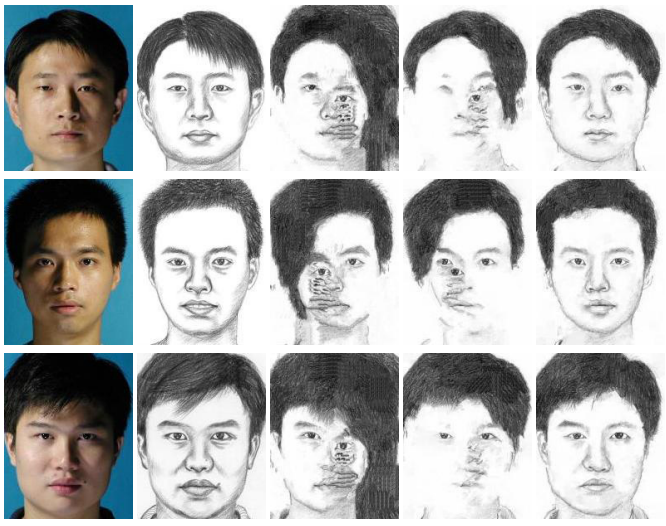
3.2 Celebrity Faces from the Web

The robustness of our method is further tested on a challenging set of face photos of Chinese celebrities with uncontrolled lighting and pose variations from

[5] As we do not have the sketches drawn by the artist for different poses, the recognition rates are not tested.

(a) (b) (c) (d) (e)

Fig. 8. Representative results on photos under the dark frontal lighting. (a) Test photo; (b) Sketch drawn by the artist while viewing a normal lighting photo; (c) Sketch by the method in [5]; (d) Sketch by the method in [5] with luminance remapping [6]; (e) Sketch by our method. The results are best viewed on screen.



(a) (b) (c) (d) (e)

Fig. 9. Representative results of photos under dark side lightings. The notations (a)–(e) are the same as Fig. 8. The results are best viewed on screen.

Fig. 10. Representative results of photos with pose variations. (a) Photo; (b) Sketch by the method in [5]; (c) Sketch by our method. The results are best viewed on screen.

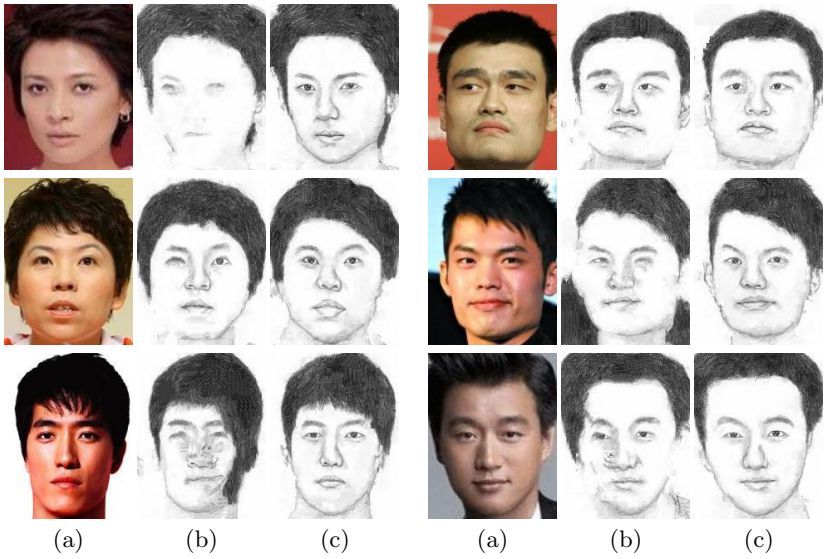


Fig. 11. Results of Chinese celebrity photos. (a) Photo; (b) Sketch by the method in [5] with luminance remapping [6]; (c) Sketch by our method. The results are best viewed on screen.

the web. They even have a variety of backgrounds. As shown in Fig. 11, the method in [5] usually produces noisy facial details and distortions, due to the uncontrolled lightings and the large variations of pose and face shape. However, our method performs reasonably well.

4 Conclusion

We proposed a robust algorithm to synthesize face sketches from photos with different lighting and poses. We introduce shape priors, robust patch matching, and new compatibility terms to improve the robustness of our method. Our method is formulated using the multiscale MRF. It significantly outperforms the

state-of-the-art approach. In the future work, we would like to further investigate face sketch synthesis with expression variations.

Acknowledgement

The authors would like to thank Dr. Wei Zhang for his generous help in experiments.

References

1. Tang, X., Wang, X.: Face sketch synthesis and recognition. In: ICCV (2003)
2. Tang, X., Wang, X.: Face sketch recognition. IEEE Trans. CSVT 14, 50–57 (2004)
3. Liu, Q., Tang, X., Jin, H., Lu, H., Ma, S.: A nonlinear approach for face sketch synthesis and recognition. In: CVPR (2005)
4. Gao, X., Zhong, J., Li, J., Tian, C.: Face sketch synthesis algorithm based on E-HMM and selective ensemble. IEEE Trans. CSVT 18, 487–496 (2008)
5. Wang, X., Tang, X.: Face photo-sketch synthesis and recognition. IEEE Trans. PAMI 31, 1955–1967 (2009)
6. Hertzmann, A., Jacobs, C., Oliver, N., Curless, B., Salesin, D.: Image analogies. In: SIGGRAPH (2001)
7. Koshimizu, H., Tominaga, M., Fujiwara, T., Murakami, K.: On KANSEI facial image processing for computerized facialcaricaturing system PICASSO. In: Proc. IEEE Int'l. Conf. on Systems, Man, and Cybernetics (1999)
8. Freeman, W.T., Tenenbaum, J.B., Pasztor, E.C.: Learning style translation for the lines of a drawing. ACM Trans. Graphics 22, 33–46 (2003)
9. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. IEEE Trans. PAMI 28, 2037 (2006)
10. Freeman, W., Pasztor, E., Carmichael, O.: Learning low-level vision. IJCV 40, 25–47 (2000)
11. Yedidia, J., Freeman, W., Weiss, Y.: Understanding belief propagation and its generalizations. Exploring artificial intelligence in the new millennium 8, 236–239 (2003)
12. Lowe, D.: Distinctive image features from scale-invariant keypoints. IJCV 60, 91–110 (2004)
13. Liang, L., Xiao, R., Wen, F., Sun, J.: Face alignment via component-based discriminative search. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part II. LNCS, vol. 5303, pp. 72–85. Springer, Heidelberg (2008)
14. Pizer, S., Amburn, E., Austin, J., Cromartie, R., Geselowitz, A., Greer, T., Romeny, B., Zimmerman, J., Zuiderveld, K.: Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing 39, 355–368 (1987)
15. Efros, A.A., Freeman, W.T.: Image quilting for texture synthesis and transfer. In: SIGGRAPH (2001)
16. Yang, J., Zhang, D., Yang, J.: Is ICA significantly better than PCA for face recognition? In: ICCV (2005)