Markov Model based Face Photo-Sketch Synthesis

A. Ezhilarasi
M.E. (communication systems) Saranathan college of Engineering Trichy, Tamilnadu.

Dr. M. Padmaa Ph.D.,
Professor,
Saranathan college of Engineering
Trichy, Tamilnadu.

***Abstract*—In recent years, Face photo-sketch synthesis become popular due to its applications in law enforcement and digital amusement. The objective is to synthesize a sketch for an input photo image. An image can be represented by multiple features. Image patch intensities, Speeded Up Robust Features and Multiscale Local Binary Pattern are the features which are obtained from the image patches of the original face image and face images processed using multiple filters. The above said features describe the different characteristics of the face image precisely. The filters used are Difference of Gaussian filter, Center-Surround Divisive Normalization filter and the Gaussian smoothing filter. In this method, the multiple representations are adaptively combined to represent an image patch. Markov model is used to utilize the relationships among the neighbouring patches. An alternating optimization is used to solve this framework.**

Keywords—face photo-sketch synthesis, multiple representations, markov model

I. INTRODUCTION

Many biometrics are used for recognition. Among them, face images can be captured in a convenient manner. A number of applications have been developed based on face images. Face photo-sketch synthesis is used in law enforcement. When a crime occurs in a place, the police use the mug shot database to find the suspects. But the photo of the suspect may not be available always. In such situations the police get the help of the artists. Artists draw the sketches of the suspects based on the remembrance of the eyewitness. But it is difficult to find the suspect by comparing the sketch drawn by artists and the photo in the mug-shot database because they are not in same modality. The solution to the above problem is either transforming photos to sketches and compares it with an inquiry sketch or transforming inquiry sketches to photos and compare it with photos in databases. Face photo-sketch synthesis is also used in drawing cartoons. In recent years, facebook is getting popular among people. The people are interested in using their face sketches as their display pictures.

In existing methods, the face images are acquired under some conditions like illumination must be even, backgrounds are to be similar. Also the existing methods consider only the single representation of the image patch, that is, the image patch intensity. But multiple representations of an image can be used to extract features. These features describe different characteristics of the face image precisely. Multiple representations of an image can be combined by the direct concatenation of all feature vectors into a single vector and existing methods can be applied to the vector. By doing so the diversity of multiple representations are lost and it leads to sub optimality.

The above said problems have dealt in the proposed method. The image is passed through three filters. The three features are obtained from the image patches of the original face image and face images processed using multiple filters. Totally, 12 features are extracted. The features are adaptively combined to represent an image patch. Markov model is used to utilize the relationships among the neighbouring patches. An alternating optimization is used to solve the framework.

The paper is organized as follows. Section II describes the related work. Section III describes the proposed markov model based face photo-sketch synthesis. The results are shown in section IV.

II. RELATED WORK

Liu et al. [1] proposed a method which contains 2 fundamental elements. The 2 elements were pseudo sketch synthesis and sketch recognition. The face images were divided into patches. Given an input photo, the photo patches in the training set were linearly combined to represent a patch of the input photo. The sketch patches were synthesized by linearly combining the corresponding candidate sketch patches. In this method, the relationships between the neighbouring patches were neglected. It leads to poor synthesis of sketches.

Wang and Tang [2] used a multiscale Markov Random Fields (MRF) model. They have assumed the face images were in frontal pose, having normal lighting, neutral expression, with no occlusions. Direct learning of the face structure globally was too complicate. So they target at local patches. The face image was divided into patches. To approximate a sketch patch for the photo patch, they found a photo patch which is similar in the training set and uses the corresponding sketch patch.

Zhou et al. [3] used Markov Weight Fields (MWF) model which can synthesize new sketch patches that may not present in the training set. They formulated the model into a convex quadratic programming (QP) problem. They used a cascade decomposition method (CDM) to solve the problem practically. It performed better even when the target patch did not find a proper sketch patch in the training set. When the input images were not formed under controlled environment, the MWF model did not work well.

Zhang et al. [4] proposed a method to synthesize a face sketch for a given face photo. It works for the face photos which were taken in a different pose and in different lighting conditions when compared with the training set. Distortion is reduced by introducing shape priors. The sketch patch candidates were found using descriptors. Matching of neighbouring sketch patches can be done by a smoothing term.

Tang and Wang [5, 6] modeled an Eigen sketch transformation method after considering the face sketch synthesis as a linear method. The projection coefficients were obtained by projecting the input test photo onto the photos in the training set. The related training sketches and the formerly obtained projection coefficients were combined linearly and so the target sketch was obtained.

III. Proposed Work

The block diagram of the proposed method is shown in Fig.1.

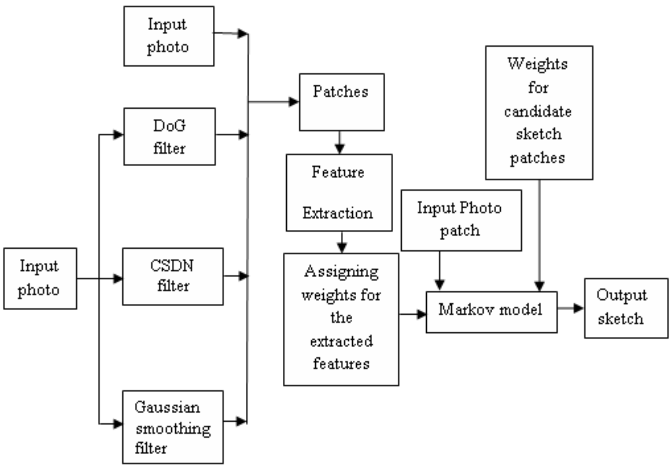

Fig. 1. Block diagram of the proposed method

Consider M face photo-sketch pairs in the training set where $x_1, x_2, \ldots, x_M$ are the photos and $y_1, y_2, \ldots, y_M$ are the corresponding sketches. The face photo and sketch images are divided into patches. The input photo is also divided into patches. The image patch intensity, Speeded Up Robust Feature and Multiscale Local Binary Pattern are obtained from all the image patches of the original image and the images which are processed using three filters. The different features are adaptively combined using weighting. For every input photo patch, K candidate photo patches are found from the training set and the corresponding K candidate sketch patches are obtained. The weights of multiple representations and the candidate sketch patches are learned adaptively using Markov model. An alternating optimization is used to solve the framework.

A. Features extracted

Image patch intensity: Intensity refers to the amount of light or numerical value of a pixel.

Speeded Up Robust Feature: The SURF descriptors are extracted on the centre of image patches.

Multiscale Local Binary Pattern: It is a texture descriptor. The 3x3 neighbourhood is threshold with the centre pixel value and the subsequent pattern is considered as a binary number.

B. Filters

Difference of Gaussian (DoG) filter: It is used in the removal of lighting variations effect and it is an edge enhancement algorithm.

Center Surround Divisive Normalization filter: The intensity gradients which are occurred by shadows are eliminated by CSDN.

Gaussian smoothing filter: Helps in the removal of noise which is present in high spatial frequencies in the input test photo image.

C. Markov model

An image patch can be represented by combining the multiple representations adaptively. Markov model is developed to learn the weights of multiple representations and candidate patches adaptively.

Graphical representation of the Markov model is shown in Fig. 2.

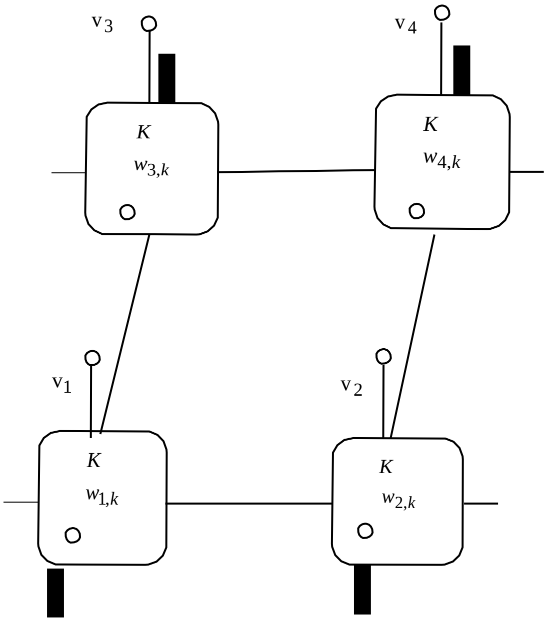

Fig. 2. Graphical representation of the Markov model

v_1, v_2, v_3, v_4 are the photo patches. $w_{1,k}, w_{2,k}, w_{3,k}, w_{4,k}$ are the lists of weights for the sketch patches.

For an input photo patch v_i, the corresponding sketch patch u_i can be synthesized by finding K candidate photo patches. $\{v_{i,1}, v_{i,2}, \ldots, v_{i,K}\}$ are the candidate photo patches, where $v_{i,k}$ stands for the kth candidate photo patch for the ith input photo patch v_i. The linear combination of the corresponding K candidate sketch patches $\{u_{i,1}, u_{i,2}, \ldots, u_{i,K}\}$ which is weighted by the K-dimensional vector w_i, gives the sketch patch u_i

$$u_i = \sum_{k=1}^{K} w_{i,k} u_{i,k} \tag{1}$$

$w_{i,k}$ - Weight of the kth candidate sketch patch

The joint probability of the input photo patches and the corresponding target sketch patches is equal to that of the input photo patches and the weights because the target sketch patches depend on the weights.

$$p(u_1,\dots,u_N,v_1,\dots,v_N)$$
$$= p(W_1,\dots,W_N,v_1,\dots,v_N)$$
$$= \prod_i \Phi\left(f(v_i), f(W_i)\right) \prod_{(i,j)\in\Xi} \Psi\left(W_i, W_j\right) \qquad (2)$$

$(i,j)\in\Xi$ means jth image patch is the neighbour of ith image patch.

$\Phi(f(v_i), f(W_i))$ is the Local evidence function.

$\Psi(W_i, W_j)$ is compatibility function.

$\Phi(f(v_i), f(W_i))$ can be given as

$$\Phi(f(v_i), f(W_i))$$
$$\propto \exp\left\{-\sum_{l=1}^{L}\mu_{i,l}\left\| f_l(v_i) - \sum_{k=1}^{K} w_{i,k} f_l(v_{i,k})\right\|^2 \Big/ 2\delta_\Phi^2\right\} \qquad (3)$$

$f(v_i) = [f_1(v_i), f_2(v_i), \dots, f_L(v_i)]$ where $f_l(v_i)$ means the lth representation of the photo patch v_i

$f(W_i) = [f_1(W_i), f_2(W_i), \dots, f_L(W_i)]$ where $f_l(W_i) = \sum_{k=1}^{K} w_{i,k} f_l(v_{i,k})$

$\mu_{i,l}$ represents the weight of the distance of the lth representation between the ith photo patch $f_l(v_i)$ and the combination of its candidates.

Ψ (w_i, w_j) can be given as

$$\Psi(W_i, W_j) \propto \exp\left\{-\left\|\sum_{k=1}^{K} w_{i,k} - \sum_{k=1}^{K} w_{j,k}\right\|^2 \Big/ 2\delta_\Psi^2\right\} \qquad (4)$$

δ_Φ and δ_Ψ are two balancing parameters.

$$p(u_1,\dots,u_N,v_1,\dots,v_N)$$
$$= p(W_1,\dots,W_N,v_1,\dots,v_N)$$
$$= \prod_{(i,j\in\Xi)} \Psi\left(W_i, W_j\right) \prod_i \Phi\left(f(W_i), f(v_i)\right) \prod_i \exp\left\{-\lambda_i \|\mu_i\|^2\right\} \qquad (5)$$

$\exp\left\{-\lambda_i \|\mu_i\|^2\right\}$ is a regularization term which avoids the overfitting of the weight of multiple representations to one representation. $\mu_i = [\mu_{i,1}, \mu_{i,2}, \dots, \mu_{i,L}]$ and λ_i balances the two terms with the regularization term. The joint probability in equation (5) is maximized to get the optimal weights for sketch synthesis.

After substituting (3) and (4) in (5), maximizing the joint probability becomes equal to minimizing the problem given below:

$$\min_{W,\mu} \frac{1}{2\delta_\Psi^2} \sum_{(i,j)\in\Xi} \left\| \sum_{k=1}^{K} w_{i,k} - \sum_{k=1}^{K} w_{j,k} \right\|^2$$
$$+ \frac{1}{2\delta_\Phi^2} \sum_{i=1}^{N}\sum_{l=1}^{L} \mu_{i,l} \left\| f_l(v_i) - \sum_{k=1}^{K} w_{i,k} f_l(v_{i,k}) \right\|^2 + \sum_{i=1}^{N} \lambda_i \|\mu_i\|^2$$
$$\sum_{k=1}^{K} w_{i,k} = 1, 0 \le w_{i,k} \le 1$$
$$\sum_{l=1}^{L} \mu_{i,l} = 1, 0 \le \mu_{i,l} \le 1 \qquad (6)$$

where $i = 1,2,\dots,N$, $k=1,2,\dots,K$ and $l=1,2,\dots,L$.

The alternating optimization strategy is used to optimize the above problem. All the synthezised sketch patches are generated using the weighed combination of candidate patches. The target sketch is obtained by stitching them together.

IV. Results And Discussion

The experiments are performed on windows 7 processor with an Intel(R) Core™ i3-2328M Machine with 2.20 GHz CPU and 2GB RAM. All the programs are written and compiled on MATLAB version 7.14.0.739 (R2012a). The sizes of the input images are 768x1024.

Fig. 3. Input images

The input images are passed through three filters.

Fig. 4. Filtered images

The images are divided into patches which is shown in Fig. 5.

Fig. 5. Extraction of patches

The patches are shown in Fig. 6.

Fig. 6. Patches

The sketch images are shown in Fig. 7.

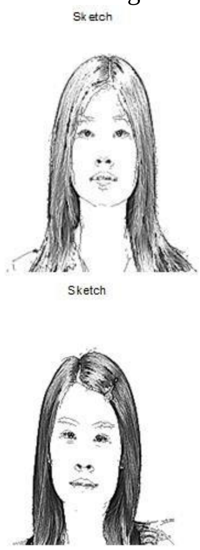

Fig. 7. Sketch images

V. CONCLUSION

In this paper, a face sketch synthesis method is proposed which uses multiple representations of an image patch. Unlike the existing methods, the features are extracted from the multiple representations and are adaptively combined using weighting. The weights of multiple representations and the candidate sketch patches are learned adaptively using markov model. An alternating optimization is used to solve the framework. By this, the sketch of a photo image is synthesized efficiently.

REFERENCES

[1] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, “A nonlinear approach for face sketch synthesis and recognition,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., San Diego, CA, USA, Jun. 2005, pp. 1005–1010.

[2] X. Wang and X. Tang, “Face photo-sketch synthesis and recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 11, pp. 1955–1967, Nov. 2009.

[3] H. Zhou, Z. Kuang, and K.-Y. K. Wong, “Markov weight fields for face sketch synthesis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Providence, RI, USA, Jun. 2012, pp. 1091–1097.

[4] W. Zhang, X. Wang, and X. Tang, “Lighting and pose robust face sketch synthesis,” in Proc. 11th Eur. Conf. Comput. Vis., Heraklion, Greece, Sep. 2010, pp. 420–433.

[5] X. Tang and X. Wang, “Face sketch recognition,” IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 1, pp. 50–57, Jan. 2004.

[6] X. Tang and X. Wang, “Face photo recognition using sketch,” in Proc. IEEE Int. Conf. Image Process., Rochester, NY, USA, Sep. 2002, pp. I-257–I-260.

[7] C. Peng, X. Gao, N. Wang, D. Tao, X. Li and J. Li, "Multiple Representations-Based Face Sketch–Photo Synthesis," in IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 11, pp.2201-2215,sep.2015