FACE PHOTO RECOGNITION USING SKETCH

Xiaoou Tang and Xiaogang Wang

Department of Information Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong
Xtang@ie.cuhk.edu.hk

ABSTRACT

Automatic retrieval of face images from police mug-shot databases is critically important for law enforcement agencies. It can help investigators to locate or narrow down potential suspects efficiently. However, in many cases, the photo image of a suspect is not available and the best substitute is often a sketch drawing based on the recollection of an eyewitness. In this paper, we present a novel photo retrieval system using face sketches. By transforming a photo image into a sketch, we reduce the difference between photo and sketch significantly, thus allowing effective matching between the two. Experiments over a data set containing 188 people clearly demonstrate the efficacy of the algorithm.

1. INTRODUCTION

Due to growing demands in such application areas as law enforcement, video surveillance, banking and security system access authentication, automatic face recognition has attracted great attention in recent years. The advantages of facial identification over alternative methods, such as fingerprint identification, are based primarily on user convenience and the cost, since face recognition can be corrected in uncertain cases by people without extensive training.

An important application of face recognition is to assist law enforcement. For example, automatic retrieval of photos of suspects from police mug-shot databases can help police narrow down potential suspects quickly. However, in most cases, the photo image of a suspect is not available. The best substitute available is often an artist drawing based on the recollection of an eyewitness. Searching an image database by using a sketch drawing is potentially very useful. It will not only help the police to locate a group of potential suspects, but may also help the witness and the artist to modify the sketch drawing of the suspect interactively based on similar images retrieved.

Despite the great need of such a sketch-based photo retrieval system, little research can be found in this area, probably due to the difficulties in building a large data set of facial sketches. We recently constructed a database of face photos and sketches of 188 people. Some examples are shown in Fig.1. Based on such a database, we develop a novel photo to sketch transformation method for face sketch recognition. The method is shown to be much more effective than directly using geometrical measures and the traditional eigenface method.

2. SKETCH RECOGNITION

2.1. Conventional eigenface method

One of the most successful methods for face image recognition is the eigenface method [3][6]. It has been ranked among the most effective methods by the comprehensive FERET test [4][5], confirming similar findings in the survey by Chellappa et. al. [1] and the comparison study by Zhang et. al. [7]. Even though eigenface method is sensitive to illumination, expression, and rotation changes, it is not an issue for our application given our focus on mug-shot photo identification.

The eigenface approach uses the Karhunen-Loeve Transform (KLT) for the representation and recognition of faces [3][6]. Once a set of eigenvectors, also called eigenfaces, is computed from the ensemble face covariance matrix, a face image can be approximately reconstructed using a weighted combination of the eigenfaces. The weights that characterize the expansion of the given image in terms of eigenfaces constitute the feature vector. When a new test image is given, the weights are computed by projecting the image onto the eigenface vectors. The classification is then carried out by comparing the distances between the weight vectors of the test image and the images from the database.

To compute the Karhunen-Loeve Transform, let Q_i be a sample face image with the mean face computed as $Q_\mu = \frac{1}{M}\sum_{i=1}^{M} Q_i$, where M is the number of training samples. Removing the mean face from each image, we have $P_i = Q_i - Q_\mu$. The photo training set then forms an N by M matrix $P = [\vec{P}_1, \vec{P}_2, ..., \vec{P}_M]$, where $\vec{P}_i$ is a column vector representation of the face image P_i, and N is the total number of pixels in the image. The sample covariance matrix can be estimated by

$$W = PP^T. \quad (1)$$

Given the large size of a photo image, direct computation of the eigenvectors of W is not practical. The dominant eigenvector estimation method [2] is generally used. Because of the relatively small sample image number M, the rank of W will only be $M-1$. So the eigenvector of the smaller matrix P^TP can be computed first,

$$(P^TP)V_p = V_p\Lambda_p, \quad (2)$$

where V_p is the eigenvector matrix and Λ_p is the diagonal eigenvalue matrix. Multiplying both sides by P, we have

$$(PP^T)PV_p = PV_p\Lambda_p. \quad (3)$$

Therefore, the orthonormal eigenvector matrix of $W = PP^T$ is,

$$E_p = PV_p\Lambda_p^{-\frac{1}{2}}. \quad (4)$$

For a new face photo $\vec{P}_k$, its projection coefficients in the eigenvector space form the vector $\mathbf{b}_p = E_p{}^T\vec{P}_k$, which is used as a feature vector for the classification.

2.2. Eigensketch transformation

Because of the structural similarity across all face images, strong correlation exists among face images. Through the KLT, the eigenface method takes advantage of such a high correlation to produce a highly compressed representation of face images, thus greatly improves the face classification efficacy.

However, because of the great difference between face photos and sketches, direct application of the eigenface method for sketch based photo identification does not work. The distance between a photo and a sketch of the same person is in general much larger than the distance between two photos of two different people. In order to overcome such a difference, we develop an eigensketch transformation algorithm to convert a photo into a sketch first then perform the classification using eigensketch features.

For the conventional eigenface method, a face image can be reconstructed from the eigenfaces by,

$$\vec{P}_r = E_p\mathbf{b}_p. \quad (5)$$

Since $E_p = PV_p\Lambda_p^{-\frac{1}{2}}$, we can represent the reconstructed photo by,

$$\vec{P}_r = PV_p\Lambda_p^{-\frac{1}{2}}\mathbf{b}_p = P\mathbf{c}_p, \quad (6)$$

where $\mathbf{c}_p = V_p\Lambda_p^{-\frac{1}{2}}\mathbf{b}_p = [c_{p_1}, c_{p_2}, \cdots, c_{p_M}]^T$ is a column vector of dimension M. We can rewrite equation (6) in a summation form,

$$\vec{P}_r = P\mathbf{c}_p = \sum_{i=1}^{M} c_{p_i}\vec{P}_i. \quad (7)$$

This shows that the reconstructed photo is in fact the best approximation of the original image using an optimal linear combination of the M training sample images. The coefficients in $\mathbf{c}_p$ describe the contribution weight of each sample image. If we now replace each sample photo image $\vec{P}_i$ in equation (7) by its corresponding sketch $\vec{S}_i$, we get,

$$\vec{S}_r = \sum_{i=1}^{M} c_{p_i}\vec{S}_i. \quad (8)$$

Given the structural similarity between photos and sketches, it is reasonable to expect that $\vec{S}_r$ is a reconstructed sketch resembling the real sketch. For such a reconstruction, a sample sketch contributes more weight to the reconstruction if its corresponding photo sample contributes more weight to the reconstructed face photo. For an extreme example, if a reconstructed photo $\vec{P}_r$ has a unit weight $c_{p_k} = 1$ for a particular sample photo $\vec{P}_k$ and zero weights for all other sample photos, i.e. the reconstructed photo looks exactly like the sample photo $\vec{P}_k$, then the reconstructed sketch $\vec{S}_r$ is simply reconstructed by replacing it with the corresponding sketch $\vec{S}_k$.

Through such a substitution, we successfully transform a photo image into its sketch. Figure 2 shows the comparison between the real sketch and the reconstructed sketch. We can clearly see the similarity between the two. After such an photo to sketch transformation, normal eigensketch recognition can be applied to the probe sketch and the photo transformed sketches. We first compute the eigenvectors using the sketch training samples. Then the probe sketch and the photo-transformed sketches are projected onto the eigensketch vectors. The projection coefficients are then used as feature vectors for final classification.

3. EXPERIMENTS

To demonstrate the effectiveness of the new algorithm, we conduct an experimental comparison with the simple geometrical measure method and the conventional eigenface method. A database containing 188 pairs of photo and sketch of 188 different people is used for the experiment. We use 88 photo-sketch pairs as training data, and the other 100 photo-sketch pairs for testing.

For the geometrical method, we use 26 measures of geometrical distances between key fiducial points. They include the sizes of nose, eyes, mouth, eyebrows, face contour and their relative positions. For the traditional eigenface method, we simply treat a probing sketch as if it is a regular photo.

We adopt the recognition test protocol used in FERET [5]. So our gallery set consists of 100 face photos. The probe set consists of 100 face sketches. The cumulative match score is used to evaluate the performance of the algorithms. It measures the percentage of "the correct answer is in the top n matches", where n is called the rank. Table 1 shows the cumulative match scores of the first ten ranks for the three methods.

Both the geometrical method and the eigenface method perform poorly in the experiment. Only around 30% accuracy rate is obtained for the first match. The accuracy for the tenth rank match is 70%. The poor performance of eigenface method can be expected given the large differences between photo and sketch. As for the geometrical measure, the results show that the reason that photo and sketch look alike is not because of the geometrical similarity of the facial components. In fact, like caricature, a sketch exaggerates the sizes of facial components. If a person has a larger than average nose, the sketch will depict an even larger nose. On the contrary, if a person has a smaller than normal nose, he will be drawn with a nose with further reduced size.

The eigensketch transform method greatly improves the recognition accuracy to 96% for the top ten candidates. The first match more than doubles the other two methods. This clearly shows the advantage of the new approach. It should be pointed out that the absolute accuracy of the algorithm should not be given too much emphasis, given the small size of the database. The results also depend on the quality of the sketch drawing. As shown in Fig. 1, not all sketches look exactly like the original photo. The first row of sketches in Fig. 1 are quite similar to their corresponding photos, yet sketches in the second row are much less so. It is hard even for human beings to give one hundred percent recognition accuracy.

4. SUMMARY

In this paper, we present a novel face recognition approach using drawn sketches. By transforming a photo image into a sketch, we reduce the difference between photo and sketch significantly, thus allowing effective matching between the two. For future study, we plan to construct a larger database to further verify the results. We will also compare the performance of our algorithm to human performance on a larger data set. This will allow us to study how the algorithm responds if the quality (judged by human observers) of the sketch drops.

ACKNOWLEDGMENT

We thank Dr. Qiumei Yang for many constructive comments. The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region. (Project no. CUHK 4190/01E).

REFERENCES

[1] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey", *Proceedings of the IEEE*, Vol. 83, pp. 705-741, May 1995.

[2] K. Fukunaga, *Introduction to statistical pattern recognition*, New York: Academic Press, 1972.

[3] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces", *IEEE Trans. on PAMI*, Vol. 12, No. 1, pp. 103-108, Jan. 1990.

[4] H. Moon and P. J. Phillips, "Analysis of PCA-based face recognition algorithms", in *Empirical Evaluation Techniques in Computer Vision*, K. W. Bowyer and P. J. Phillips, Eds., IEEE Computer Society Press, Los Alamitos, CA, 1998.

[5] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET Evaluation", in *Face Recognition: From Theory to Applications*, H. Wechsler, P.J. Phillips, V. Bruce, F.F. Soulie, and T.S. Huang, Eds., Berlin: Springer-Verlag, 1998.

[6] M. Turk and A. Pentland, "Eigenfaces for recognition", *J. of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.

[7] J. Zhang, Y. Yan, and M. Lades, "Face recognition: eigenface, elastic matching, and neural nets", *Proceedings of the IEEE*, Vol. 85, pp.1423-1435, Sept. 1997.

Table 1. Cumulative match score for the three methods.

Rank	1	2	3	4	5	6	7	8	9	10
Geometry Method	30	37	45	48	53	59	62	66	67	70
Eigenface Method	31	43	48	55	61	63	65	65	67	67
New Method	71	78	81	84	88	90	94	94	95	96

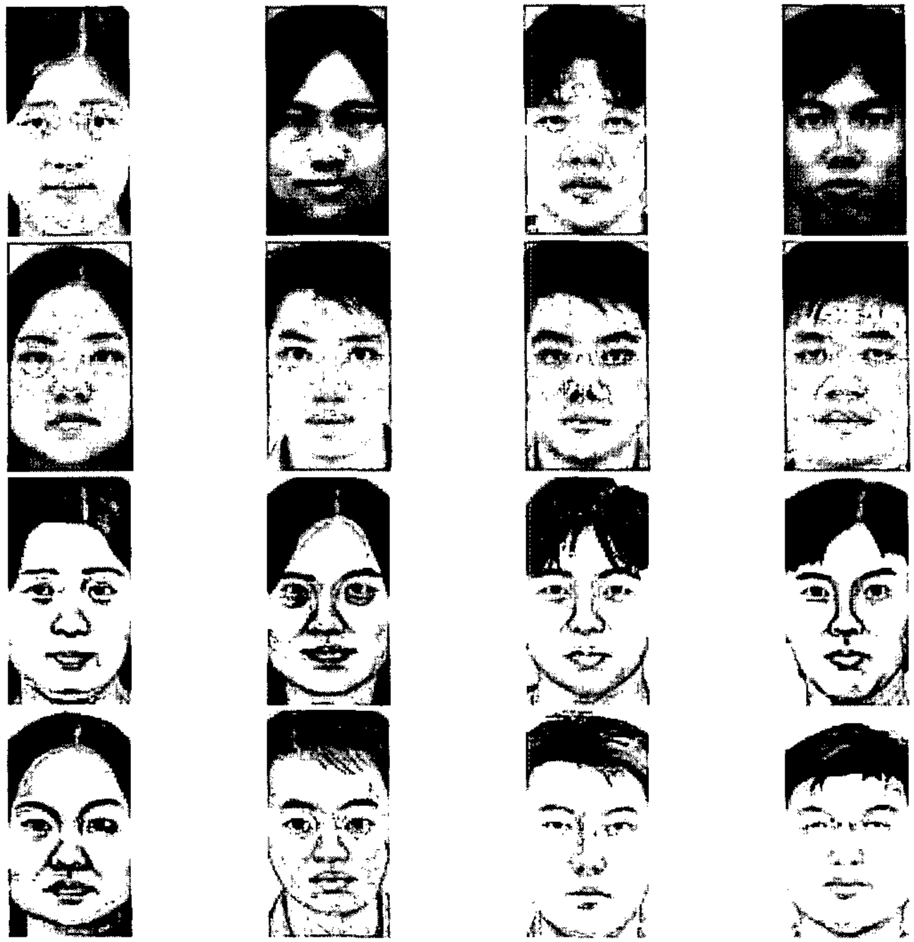
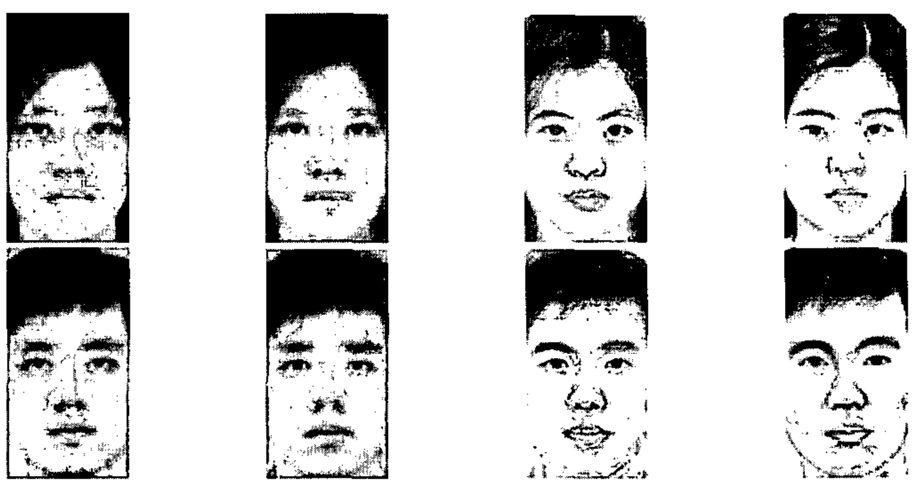


Figure 1. Sample face photos (top two rows) and sketches (bottom two rows).



(a) Photo (b) Reconstructed photo (c) Reconstructed sketch (d) original sketch

Figure 2. Photo to sketch transformation examples.