Photo-Sketch Recognition: Eigentransformation Method

M. A. A. Silva, G. Cámara-Chávez, D. Menotti
Computer Science Department, Federal University of Ouro Preto
Ouro Preto, MG, Brazil

Abstract—*Automatic systems for matching facial sketch images from the police mug shot database are very important for law enforcement agencies. These systems can help to locate or narrow down potential suspects. This paper deals the problem of face recognition through forensic sketches with focus on the eigentransformation method developed by Tang & Wang. This method is based on the Eigenface method. In this method a sketch is transformed into a photo using a global linear transformation, reducing significantly the difference between them, allowing an effective matching.*

Keywords: Face Recognition, Sketch Recognition, Eigenface, Eigentransformation.

1. Introduction

A new face recognition problem that has recently emerged is the association between sketches and photos. The consequence of this problem is the development of robust algorithms for security agencies. When a crime is observed by an eyewitness, often a verbal description of the features of the offender is employed by a police artist to draw a sketch of the suspect. Many criminals have been apprehended when identified by such sketches [1].

Automating this process helps the police to reduce the number of suspects, making the identification faster and less tiring. Besides making the search easier, this method can also help witnesses and designers to modify the design of sketch interactively [2].

The last two decades have witnessed tremendous advances in facial recognition. The research of Turk & Pentland [3], [4] has served as the foundation for the modern mechanisms of facial recognition [1].

However, due to the large difference between sketches and photos, in addition to the lack of knowledge about the psychological mechanisms of sketch generation, recognizing suspects through sketch becomes a much more difficult task than facial recognition [5].

Most of the researches in photo-sketch recognition in the last ten years have been developed by Tang and Wang [6], [7], [2], [5]. The first approaches developed by Tang and Wang (2002, 2003, 2004) [6], [7], [2] use global linear transformations, based on eigenface method [3], [4], to convert a photo into a sketch.

In [5], the authors propose a new method for photo-sketch synthesis and recognition based on a multiscale Markov random fields (MRF). They use a multi-scale MRF model to learn face structures at different scales. Local patches in different regions and structures are learned jointly. Another characteristic of this approach is that it can also synthesize face photos given sketches. The solution to the MRF was estimated using the belief propagation algorithm [8]. Solution patches are stitched together and form a synthetic photograph. The transformation of a photo into sketch (or the reverse) significantly reduces the difference between them. After the synthetic image generation, in principle, most of the algorithms for facial recognition may be applied directly.

Klare and Jain [1] proposed a Scale Invariant Feature Transform (SIFT) based local feature approach. The method consists in sampling the SIFT feature descriptors uniformly across all the sketch and photo images, then both are matched directly. The recognition proceeds by computing the distance of the SIFT representation between the sketch and photo, or using a dictionary composed by training pairs.

Most recent researches focus on identifying sketches that were drawn while a viewing a photograph of the person, this type of sketches are known as viewed sketches [7], [2], [9], [5], [1]. Unfortunately real-world scenarios only involve sketches that were drawn by interviewing a witness to gain a description of the suspect, known as forensic sketches. As we can see in Figure 1.

In [11], the authors presents a framework called LFDA where photo and sketch images are represented by descriptors SIFT and MLBP Multiscale Local Binary Patterns) features. Local feature-based discriminant analysis (LFDA) is used to compute the minimum distance matching between sketches and photos. Which creates a projection function based on vertical slices of sketches and photos. The LFDA attempts to maximize inter-class distances while the intra-class distances are minimized.

In this paper we evaluate with more details the eigentransformation method [2], varying the interocular distance, using five-fold cross-validation, and training and testing with ”real” database, obtained from [10].

The paper is organized as follows. In Section 2 we briefly review the Eigentransformation method. The results of our experiments using this method can be found in Section 3. And the conclusions can be found in Section 4.

(a)

(b)

Fig. 1: Difference between sketch made by artist looking at the photo (a) and through the description of witnesses (b). Images from CUHK database[2] and *Forensic Art and Illustration*[10].

2. Eigentransformation

The Eigentransformation method is based on the eigenface method, developed by Turk and Pentland (1991) [3], [4], which is one of the classic methods for face recognition. Because of the structural similarity across all face images, there exists a strong correlation between them. The eigenface method takes advantage of this and produce a highly compressed representation of face images. Unfortunatey, this can not be extended to face photos and sketches. Direct application of the eigenface method for sketch-based face identification may not work. This occurs because the distance between a photo and a sketch is much larger than the distance between two photos of two different people. In order to overcome this problem, the authors developed a photo-sketch transformation to either project a sketch image into a photo space, or to project a photo image into a sketch subspace, see Figure 2.

First, create an eigenspace for photos and another for the sketches, using training pairs formed by their corresponding photos and sketches. Then, a new entering photo is projected in the photo space, and represented by a vector, where the coefficients represents the contribution of an existing face of the training samples. Thus, a new face can be approximately reconstructed from the training set and the coefficients of the projection of the photo space are used to build a sketch in the sketch face. An example can be seen in Figure 3.

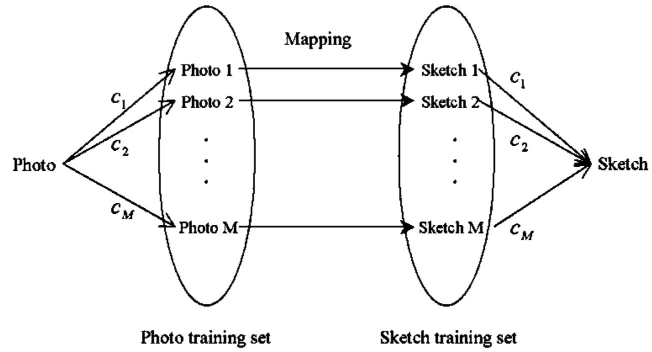


Fig. 2: Graphic representation of the projection method. Tang & Wang (2004) [2].

Fig. 3: Results of the synthesis of sketch from the projection method of coefficients of eigenspace.

Below is described the process of Tang & Wang's approach [2]:

$$\vec{m}_p = 1/M \sum_{i=1}^{M} \vec{Q}_i$$ Среднее значение

$$\vec{P}_i = \vec{Q}_i - \vec{m}_p$$ Центрирование

$$A_p = [\vec{P}_1, \vec{P}_2, ..., \vec{P}_M]$$ Матрица данных

$$W = A_p A_p^T$$ Матрица ковариации

Матрица Гр-Шм

$$(A_p^T A_p) V_p = V_p \Lambda_p$$ Задача на собственные значения

$$(A_p A_p^T) A_p V_p = A_p V_p \Lambda_p$$ $A_p(A^T_pA_p)V_p = A_pV_{p\ p}$

$$U_p = A_p V_p \Lambda_p^{-1/2}$$ Ортонормальный вектор для W

$$\vec{b}_p = U_p^T \vec{P}_k$$

Проекция для нового изображения в форме вектора (например, скетча)

Реконструкция скетча

$$\vec{P}_r = U_p \vec{b}_p$$

With the vector $\vec{b}$ we can reconstruct the image in the same domain.

$$U_p = A_p V_p \Lambda_p^{-1/2}$$

$$\vec{P}_r = A_p V_p \Lambda_p^{-1/2} \vec{b}_p = A_p \vec{c}_p$$

With the vector $\vec{c}$ we can reconstruct the image from the contributions of the images of the training set. The proposed method, synthesize an image on another domain.

$$\vec{c}_p = V_p \Lambda_p^{-1/2} \vec{b}_p = [c_{p_1}, c_{p_2}, ..., c_{p_M}]^T$$

$$\vec{P}_r = A_p \vec{c}_p = \sum_{i=1}^{M} c_{p_i} \vec{P}_i$$

Tang & Wang [2] method follows the next steps:

- Compute the average images $\vec{m}_p$ for the training set of photos and $\vec{m}_s$ for sketches.
- Compute the photo eigenspace U_p and sketch eigenspace U_s .
- Remove the photo mean $\vec{m}_p$ from the input photo image $\vec{Q}_k$ to get $\vec{P}_k = \vec{Q}_k - \vec{m}_p$.
- Project $\vec{P}_k$ in the eigenspace U_p to compute the eigenface weight vector $\vec{b}_p$.
- Found de contribution vector $\vec{c}_p = V_p \Lambda_p^{-1/2} \vec{b}_p$
- Reconstruct the pseudo-sketch by: $\vec{S}_r = A_s \vec{c}_p = \sum_{i=1}^{M} c_{p_i} \vec{S}_i$
- Finally, add back the average sketch: $\vec{T}_r = \vec{S}_r + \vec{m}_s$.

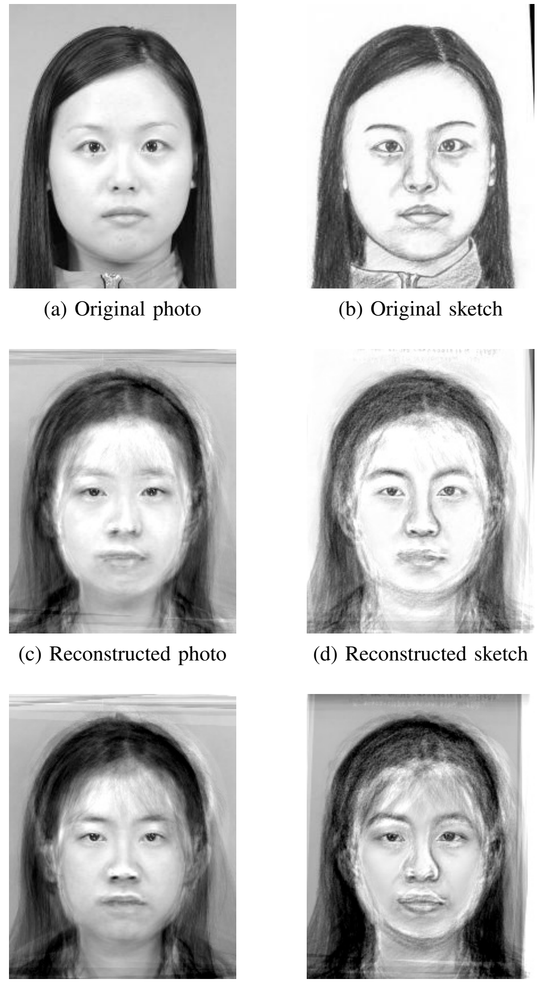
Then we can use the next metrics for recognition:

- $d_1 = ||\vec{c}_p - \vec{c}_s||$
  Direct distance
- $d_2 = ||\vec{b}_r - \vec{b}_s||$
  $\vec{b}_r$ is a pseudo-sketch
- $d_3 = ||\vec{b}_r - \vec{b}_p||$
  $\vec{b}_r$ is a pseudo-photo

3. Results and Discussion

In this section are shown the obtained results. We use Tang *et al.* data set, available in `http://mmlab.ie.cuhk.edu-.hk/facesketch.html`, composed by *viewed sketches* and *forensic sketches* from [10] composed by real sketches.

The implemention was done using the library OpenCV (Open Source Computer Vision Library) on the C++ language, since the methods require high computational performance.



(a) Original photo (b) Original sketch

(c) Reconstructed photo (d) Reconstructed sketch

(e) Photo reconstructed from the sketch (f) Sketch reconstructed from the photo

Fig. 4: Results of reconstructions using the projections of the coefficients between eigenspaces.

In Figure 4 is presented the results of reconstructions of a photograph and a sketch.

The tests were done by changing the interocular distances. Different values were tested, the best results were achieved with 66 pixels of interocular distance, as we can see in Figures 5, 6 and 7. We keep the original window size of 200×250 pixels.

Table 1 and Figure 8 are shown the results obtained from the implementation of eigentransformation training and testing using the CUHK database [2]. The result from [2] is shown in the Table 2. In our experiments we observed that the variation in interocular distance modifies the results. The background reduction is the main factor to improve the results. Because background information acts as noise in this case, since the method uses global transformations. Then find a correct ratio to cropping is important.

We evaluate the method through five-fold cross-validation using 66 pixels of interocular distance in the CUHK database [2], the results are shown in Table 3.

We test the eigentransformation with real sketches, obtained from [10]. We trained with CUHK database(188 pairs of viewed sketches and photo) and tested with 58 pairs of forensic sketches and photos, the results are shown in Figure 9 and Table 4. We can see that the results were low, since the test database consisted of only 58 pairs. The ideal would be to have a training base with forensic sketches too, but this requires more sketches, which is a difficult task.

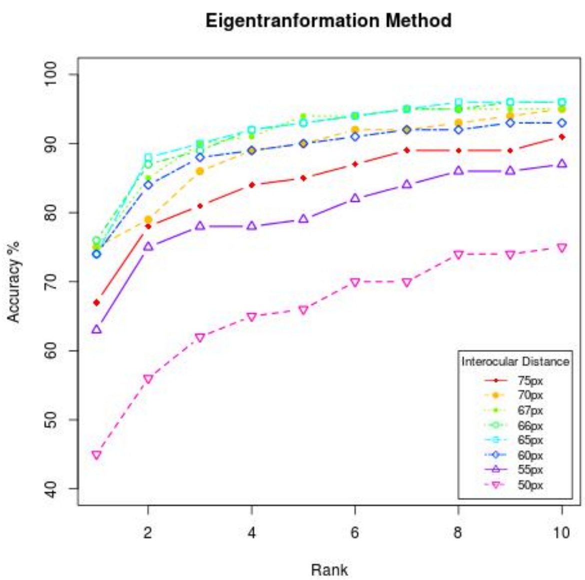


Fig. 5: Comparison of the results using distance method d_1.

Table 1: Results of the three methods of calculating the distance, running with the same database with 66 pixels of interocular distance.

Rank	1	2	3	4	5	6	7	8	9	10
d_1	76	87	89	92	93	94	95	95	96	96
d_2	84	93	96	97	98	98	99	99	99	100
d_3	71	78	83	84	85	88	90	90	91	91

Table 2: Results of the three methods of calculating the distance, shown in the original paper, running with CUHK database [2] (88 pairs for training and 100 pairs for testing.

Rank	1	2	3	4	5	6	7	8	9	10
d_1	20	49	59	65	69	73	75	76	81	82
d_2	71	78	81	84	88	90	94	94	95	96
d_3	57	70	77	79	83	84	85	86	87	88

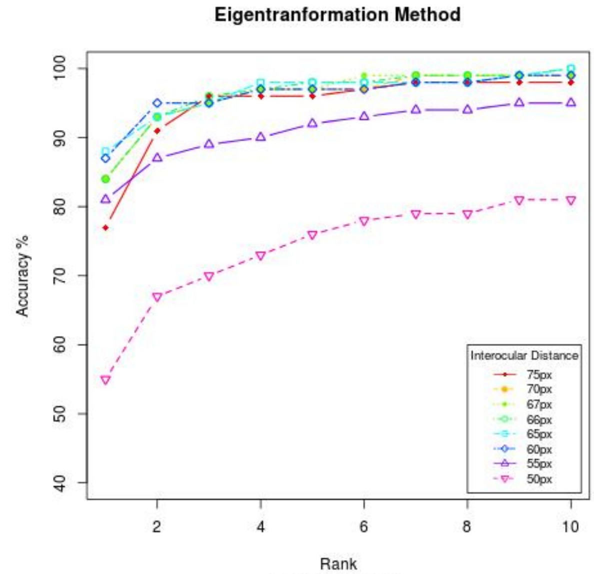


Fig. 6: Comparison of the results using distance method d_2.

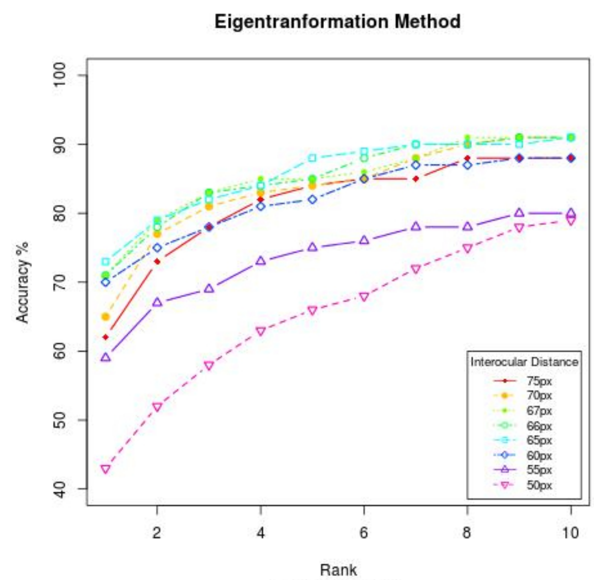


Fig. 7: Comparison of the results using distance method d_3.

4. Conclusions

The results demonstrated that how the images are cropped influences the method, since the method uses global transformation, then all information is computed. As we can see the method could have been explored to improve the results. We can not define if the method can be applied to forensic sketches, because is necessary a larger database for training and testing.

Table 3: Results of the three methods of calculating the distance, running with CUHK database [2] with five-fold cross-validation with 66 pixels of interocular distance.

Rank	1	2	3	4	5	6	7	8	9	10
d_1%	91.4±6.2	95.1±5.2	98.4±2.4	98.4±2.4	98.4±2.4	98.4±2.4	99.5±1.2	99.5±1.2	99.5±1.2	99.5±1.2
d_2%	93.5±2.4	98.4±2.4	98.4±2.4	98.4±2.4	98.4±2.4	98.4±2.4	98.4±2.4	98.4±2.4	98.4±2.4	99.5±1.2
d_3%	88.6±3.5	92.4±4	94.6±3.8	96.2±4.1	96.2±4.1	96.2±4.1	96.2±4.1	97.3±2.7	97.3±2.7	97.3±2.7

Table 4: Results of the three methods of calculating the distance, running with CUHK database [2] (188 pairs) for training and 58 pairs of photo and real sketch obtained from [10] for testing, both with 66 pixels of interocular distance.

Rank	1	2	3	4	5	6	7	8	9	10
d_1	3.4%	8.6%	10.3%	10.3%	12.1%	12.1%	17.2%	20.7%	20.7%	20.7%
d_2	1.7%	12.1%	17.2%	22.4%	24.1%	27.6%	34.5%	39.7%	43.1%	43.1%
d_3	5.2%	6.9%	8.6%	10.3%	17.2%	20.7%	22.4%	25.9%	25.9	27.6%

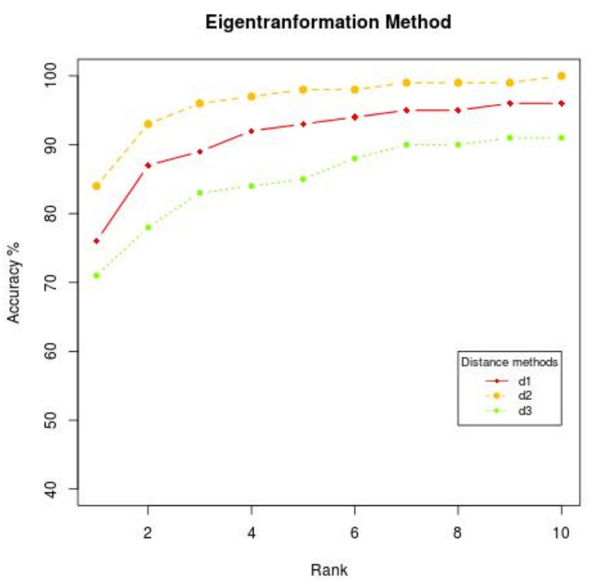


Fig. 8: Comparison of the three methods of distance with 66 pixels of interocular distance.

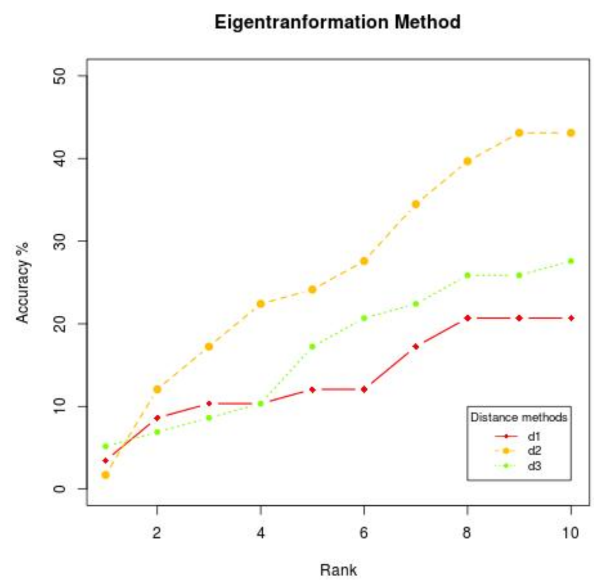


Fig. 9: Results of the method using the distance d_2. Training with CUHK database (188 pairs) and testing with 58 pairs of photo and real sketch.

References

[1] B. Klare and A. Jain, "Sketch to photo matching: A feature-based approach," *Proc. SPIE, Biometric Technology for Human Identification VII*, vol. 7667, pp. 766 702–766 702, 2010.

[2] X. Tang and X. Wang, "Face sketch recognition," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 14, no. 1, pp. 50–57, 2004.

[3] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*. IEEE, 1991, pp. 586–591.

[4] ——, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

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

[6] X. Tang and X. Wang, "Face photo recognition using sketch," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 1. IEEE, 2002, pp. I–257.

[7] ——, "Face sketch synthesis and recognition," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 687–694.

[8] J. Yedidia, W. Freeman, and Y. Weiss, "Generalized belief propagation," *Advances in neural information processing systems*, pp. 689–695, 2001.

[9] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 1005–1010.

[10] K. T. Taylor, *Forensic art and illustration*. CRC Press, 2010.

[11] B. Klare, Z. Li, and A. Jain, "Matching forensic sketches to mug shot photos," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 3, pp. 639–646, 2011.