Face Sketch-Photo Recognition

M. A. A. Silva, G. Cámara-Chávez
Computer Science Department, DECOM
Federal University of Ouro Preto, UFOP
Ouro Preto, MG, Brazil
{marcoantonio.maas,gcamarac}@gmail.com

***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. In this work, we deal the problem of face recognition through forensic sketches with focus on the Eigentransformation method and Local feature-based discriminant analysis (LFDA). The Eigentrasformation method uses a global linear transformation to synthesize a photo from a sketch, and is sensitive to gross characteristics such as face shape and skin color. While LFDA uses local descriptors, which describe structures in the face that are relevant to face recognition.**

***Keywords*-Face Recognition; Matching Photo-Sketch; Forensic Sketches;**

I. 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].

II. Related Work

Most of the researches in photo-sketch recognition in the last ten years have been developed by Tang and Wang. The first approaches developed by Tang and Wang (2002, 2003, 2004) [6], [7], [2] use global linear transformations, based on eigenface method [3], [4], in order 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 or using a dictionary composed by training pairs. The recognition proceeds by computing the distance of the SIFT representation between the sketch and photo.

Most recent researches focus on identifying sketches that were drawn while viewing a photograph of the person [7], [2], [9], [5], [1], this type of sketches are known as viewed sketches. 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 Fig. 1.

In [11], the authors presents a framework called LFDA where photo and sketch images are represented by descriptors SIFT and Multiscale Local Binary Patterns (MLBP) features. Local feature-based discriminant analysis (LFDA) is used to compute the minimum distance matching between sketches and photos. This method 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. This approach obtain excelent results with *viewed sketches* and is the first large-scale published experiment on matching real *forensic sketches*.

III. Theorical Background

A. Eigentransformation

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

- Compute the average images $\vec{m}_p$ for the training set of photos and $\vec{m}_s$ for sketches.

(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].

- 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 the pseudo-sketch
- $d_3 = ||\vec{b}_r - \vec{b}_p||$
  $\vec{b}_r$ is the pseudo-photo

B. *Local feature-based discriminant analysis (LFDA)*

LFDA [11] follows the next steps:

- Divide each sketch and photo in overlapping patches.
- Extracted from each patch SIFT and MLBP feature descriptors.
- Grouping vertical slices of patches together into feature vectors.
- Learn a discriminant projection (LDA) for each slice.
- Combining each projected vector slice into a single vector
- Recognition is done measuring the normed distance between a probe sketch and a gallery photo.

IV. Parcial Results

All sketch and photo images were normalized by rotating the angle between the two eyes to 0 degree, scaling the images to a 75 interocular pixel distance, and cropping the image size to 200 by 250 pixels, with the eyes horizontally centered and vertically placed at row 115, similar to [11].

The implemention was done using the OpenCV library [12] (Open Source Computer Vision Library), since the methods require high computational performance.

In this work we have implemented the Eigentrasformation method [2] and partially LFDA, some details remain to be better investigated. We use 200 pairs of *viewed sketches* for training and the 158 pairs of *forensic sketches* for testing. And 1,000 extra mug shots to populate the galery.

A. *Database*

In our experiments we use CUFSF database [5], available in http://mmlab.ie.cuhk.edu.hk/cufsf/, which is composed by *viewed sketches* , each with a corresponding photograph. And a data set consisting of 158 forensic sketches, each with a corresponding photograph of the subject who was later identified by the law enforcement agency. These sketches were drawn by forensic sketch artists working with witnesses who provided verbal descriptions or working with age progression. The corresponding photographs (mug shots) are the result of the subject later being identified. The forensic sketch data set used here comes from two different sources:

- 110 images from the forensic sketch artist Lois Gibson [10].
- 48 images from the forensic sketch artist Karen Taylor [13].

In addition, we also use a data set of 1,000 mug shot images available in http://www.califoniamugshots.com to populate the gallery.

B. *Experiments*

In our experiments with the Eigentransformation method we observed that the way the images were cropped influences the method, since the method uses global transformations. We also observed slightly better results than those presented in [2] for *viewed sketches*. The variation in interocular distance modifies the results, as we can see in Table I and Table II. The background reduction is the main factor to be improved, because background information acts as noise, since this method uses global transformation. Thus, finding a correct ratio for image cropping is very important task. But the results from from forensic sketches were not good.

TABLE I
RESULTS OF EIGENTRANSFORMATION WITH 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

TABLE II
RESULTS OF EIGENTRANSFORMATION WITH 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

Our implementation of LFDA is currently working with patches size of 16×16 pixels while in [11], they work with patches of 16×16 and 32×32 and after uses a score fusion of the results. The results of our experiments are shown in Table III. And two examples of matching are shown in Fig. 2

TABLE III
RESULTS OF OURS PARTIAL IMPLEMENTATION OF LFDA, RUNNING WITH A 200 PAIRS OF *viewed sketches* TO TRAINING AND 158 PAIRS OF *forensic sketches* TO TESTING. AND 1,000 MUG SHOTS TO POPULATE THE GALLERY.

Rank	1	2	3	4	5	10	50
	3.2%	6.3%	8.2%	8.9%	10.8%	12%	28.5%

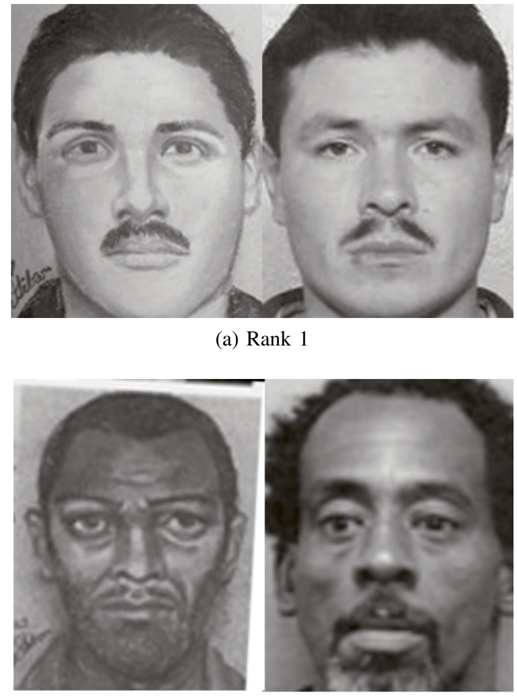
We tested other descriptors like SURF and also change the weights of each slice, but the results were worst.

V. CONCLUSION AND FUTURE WORKS

We reproduce some of the actual state of the art methods and we obtain similar results. Now, we are finishing the implementation of other methods and optimizing the performance of the already implemented. We also pretend to combine different methods in order to improve the recognition rate. We are going to study, implement and test a new method presented in [14].

ACKNOWLEDGMENT

The authors would like to thank CAPES and CNPq for the financial support.



(a) Rank 1

(b) Rank 668

Fig. 2. Examples of (a) *good* matching and (b) *bad* matching.

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.
[12] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.
[13] L. Gibson, *Forensic art essentials: a manual for law enforcement artists*. Academic Press, 2010.
[14] B. Klare and A. Jain, “Heterogeneous face recognition using kernel prototype similarities,” 2012.