

ILLUMINATION TOLERANT FACE RECOGNITION USING A NOVEL FACE FROM SKETCH SYNTHESIS APPROACH AND ADVANCED CORRELATION FILTERS

Yung-hui Li, Marios Savvides, and Vijayakumar Bhagavatula

Carnegie Mellon University
{yhli, msavvid}@cs.cmu.edu; kumar@ece.cmu.edu

ABSTRACT

Current state-of-the-art approach for performing face sketch recognition transforms all the test face images into sketches, and then performs recognition on sketch domain using the sketch composite. In our approach we propose the opposite; which has advantages in a real-time system; we propose to generate a realistic face image from the composite sketch using a Hybrid subspace method and then build an illumination tolerant correlation filter which can recognize the person under different illumination variations from a surveillance video footage. We show how effective proposed algorithm works on the CMU PIE (Pose Illumination and Expression) database.



Figure 1: examples of face and their sketch images.

sketch images using various approaches, for example, eigenspace [3] and LLE [5]. This phase can be called “synthesis phase”. Second, they perform all the recognition tasks on sketch image domain using a different approach, for example, distance metric [3], Bayesian, LDA, PCA and KNDA [5].

1. INTRODUCTION

Face recognition has attracted much attention in recent years and many different methods have been proposed[1]. However, most of the proposed methods focus on the problem where both training and testing images are face images. In practical law-enforcement scenarios, we may encounter the situation where only a police-sketch of a suspects face image is available. And from this sketch image, we would like to retrieve or synthesize the real face image. In this paper we propose a novel approach to deal with this problem.

1.1. Problem Definition

Available Training data: A collection of photos of face images from many different people (generic data).

Given: A face sketch image, drawn by police sketch artist.

Goal: Find the identity of the person whose face sketch is given. Examples of face-sketch image pair are shown in Figure 1.

1.2. Previous Work

Wang & Tang [2-5] have tackled this approach in a different way. They divide the task into two phases: At the first phase, they transform all the face images in their database into

2. PROPOSED METHOD

Although previous work has shown to achieve good recognition results, we propose an alternative method to approach this problem. First, given the fact that sketch images have less information compared to original face image, by transforming all face images into their sketch counterpart, we lose information which may be useful for the recognition phase. More importantly if we want to make this algorithm into a fully automated face recognition system, synthesizing sketch faces from live footage will not be accurate sketch representations due to the effects of illumination variation in the ambient background which can lead to distorted face sketches due to artefacts introduced by cast-shadows. This is the key difference from our approach and others; while the previous work shows promise to perform sketch using a well collected enrolment mug-shot database, it is not practical for face recognition from sketch in a real-time system.

We propose a novel approach to address this problem. We propose the exact opposite, we would like to synthesize a realistic face image from the sketch, this has real world applications for law-enforcement officers looking for criminals, but more importantly from this synthesized facial image we can then use that to build an illumination tolerant correlation filter which can be used in real-time to match against live-video footage.

2.1. Face from Sketch Synthesis

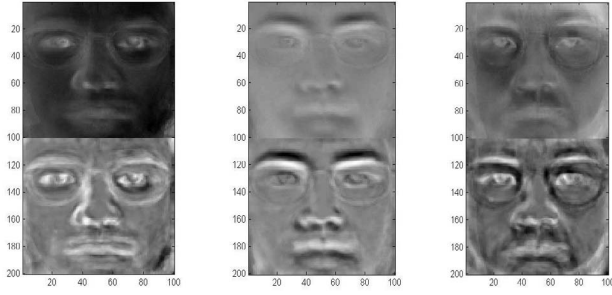


Figure 2: the first 3 eigenfaces from the face-sketch hybrid space.

In synthesis phase, we want to transform the sketch image back to its face image counterpart. Given a training database consisting of N face images and their synthesized sketch image counterparts, we form a hybrid space by appending sketch image vector and the face image vector. Specifically, suppose $T_f = [T_{f1} \ T_{f2} \dots T_{fN}]$ is a matrix which represents all our face image data, and T_{fi} (i^{th} column of T_f) is formed by vectorizing i^{th} face image in our face database. Similarly, $T_s = [T_{s1} \ T_{s2} \dots T_{sN}]$ is a matrix which represents all our sketch image data, and T_{si} (i^{th} column of T_s) is formed by vectorizing the i^{th} sketch image corresponding to i^{th} face image in our database. We form a new hybrid space $H = [h_1 \ h_2 \dots h_N]$, where h_i (the i^{th} column of H) is formed by appending the matrices T_{si} and T_{fi} .

After we have this matrix H in hybrid space, we perform eigen-analysis [6] to form a global subspace for face/sketch representation. In the end we will have a hybrid eigenface/sketch representation $V_h = [V_{h1} \ V_{h2} \dots V_{hN-1}]$, where the i^{th} column V_{hi} represents the i^{th} eigenvector of H , as shown in Figure 2.

If we dissect each column vector V_{hi} of V_h to two parts, V_{fi} and V_{si} , and form $V_f = [V_{f1} \ V_{f2} \dots V_{fN-1}]$ and $V_s = [V_{s1} \ V_{s2} \dots V_{sN-1}]$, then we can treat the newly derived V_f and V_s as pseudo-eigenface matrix and pseudo-eigensketch matrix, respectively. Note that, theoretically speaking, V_f and V_s are not eigenvector matrices of their sub-space and each column vectors in V_f (and V_s) is not orthogonal to each other. However we call them pseudo-eigenvectors. Note that if we train eigensketch solely from the sketch images, the outcome eigensketch matrix U_s will be different than pseudo-eigensketch matrix V_s which we get from lower half of hybrid eigenface matrix V_h . Similarly is the case for pseudo-eigenface matrix V_f . Furthermore, by closely comparing the hybrid eigenfaces, we find that the images in upper half (which are pseudo-eigenfaces) are highly correlated to the corresponding images in the lower half (which are pseudo-eigensketches). But if we compute eigenfaces U_f solely from face images, and eigensketches U_s solely from sketch images, and compare corresponding

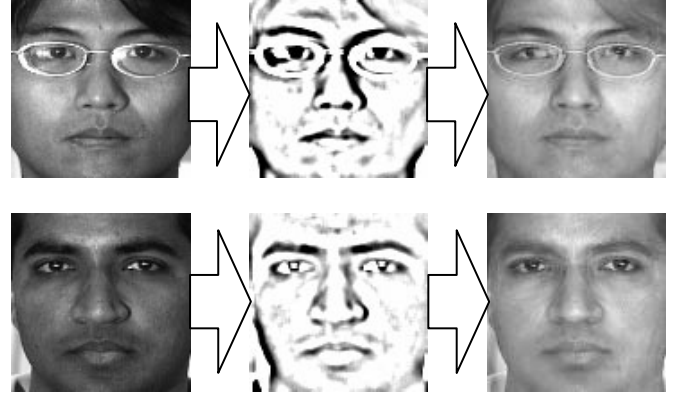


Figure 3: Examples of original face and corresponding reconstructed face when we only have sketch images. The images in first column are original face; the images in second column are sketch of the face images, and the images in third column are reconstructed face images from sketches.

column of U_f and U_s , there's no similarity or correlation between these two.

The reason is that the maximal variation in subspace of covariance matrix of face images must be different from the maximal variation in subspace of covariance matrix of sketch images. Therefore, when doing PCA, the eigenfaces we found are not necessarily correlated to eigensketches, when training separately.

Figure 3 demonstrates the original images vs. their reconstructed counterpart which are built from the hybrid space transform method. From Figure 3 we can see that the reconstruction is pretty successful in terms of the resemblance of facial characteristics. However, one can also notice that the reconstructed images are a little lighter than the original ones. So, if we perform recognition with distance-based metric, like one nearest neighbour (1NN), the distance between original image and reconstructed ones will be very large. To attack this problem we choose advanced correlation filter as our recognition approach.

2.2. Recognition Phase – Using Advanced Correlation Filters

Advanced correlation filters [7] when correlated with an image result in a correlation plane C which measures the correlation between the filter and the image. Correlation of a class-specific filter with authentic and impostor data yield very different correlation planes. Figure 4 demonstrates this difference. These advanced correlation filters optimize specific criteria to obtain sharp correlation peak outputs as shown below; this is very different from matched filters or normalized correlation approaches which are more common in the literature.

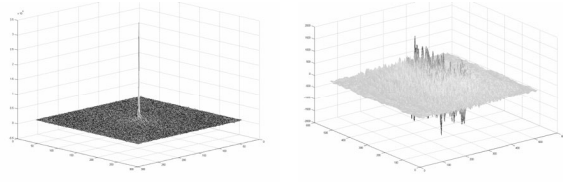


Figure 4: Left: correlation plane of an authentic sample. Right: Correlation plane of an imposter sample

To quantify the difference between the two types of correlation planes, we define a measure of recognition called **Peak to Correlation Energy (PCE)**. This will measure the sharpness of the largest peak in the correlation plane with respect to the rest of the correlation plane.

$$PCE(C) = \frac{\max(|C|) - \text{mean}(|C|)}{\text{std.dev}(|C|)} \quad (1)$$

The **Minimum Average Correlation Energy (MACE) Filter** [8] is designed to minimize the average energy E in the correlation plane or **Average Correlation Energy (ACE)**.

$$h_{MACE} = D^{-1} X (X^+ D^{-1} X)^{-1} u \quad (2)$$

where u is the constrained peak values (vector of ones), X contains the 2D Fourier transforms of training images along the columns, and D is a diagonal matrix, of which each diagonal element represents the power spectrum of each training image x_i .

The **Unconstrained MACE (UMACE) Filter** [9] removes the constraint on the peak value. By removing this constraint, more solutions to the minimization problem are available. We also try to maximize the average value of the peaks or **Average Correlation Height (ACH)**. The closed form solution to the UMACE filter h_{UMACE} :

$$h_{UMACE} = D^{-1} m \quad (3)$$

where m is the average Fourier transform of the training images.

We will consider generalizations of the MACE and UMACE filters called the **Optimal Tradeoff Synthetic Discriminant Function (OTSDF)** filter [10] and the **Unconstrained OTSDF (UOTSDF)** filter respectively. The filter designs h_{OTSDF} and h_{UOTSDF} are:

$$h_{OTSDF} = T^{-1} X (X^+ T^{-1} X)^{-1} u \quad (4)$$

$$h_{UOTSDF} = T^{-1} m \quad (5)$$

where T is defined as:

$$T = \alpha D + \sqrt{1 - \alpha^2} C \quad \text{given } 0 \leq \alpha \leq 1 \quad (6)$$

where C is the Gaussian white noise matrix (identity matrix). The primary difference between MACE and OTSDF is the replacement of D with T . In this paper we use OTSDF (with alpha close to 1) throughout all the recognition experiments.

3. EXPERIMENTS

3.1. Database

The database we used in our experiment is CMU-PIE database, which can be divided into two sub-sets: which we refer to as Light(captured with ambient background lighting) and NoLight(captured with no ambient background lighting). There are 1430 images in Light database, and 1365 in NoLight. CMU-PIE database has following characteristics:

- Contains both male and female faces
- Contains people from different race and color
- Contains images of people with and without glasses
- Contains illumination variation across images of each person.

To generate corresponding sketch images for each of the face, we used one of the non-linear sketch functions found in Adobe PhotoShop® 7.0 and manually repeated this for all face images. We choose image index 7 and 8 of each person as training data for making the mug-shot sketches, and all other images are left to be testing data (unknown harsh illumination). We performed experiments on Light and NoLight separately to compare the performance of proposed algorithm in different illumination condition. In experiment of Light database, we have 130 training images and 1300 testing images. In experiment of NoLight database, we have 130 training images and 1235 testing images.

3.2. Procedures

In training phase, assume we have a set of face images $T_f = [T_{f1} \ T_{f2} \ \dots \ T_{fN}]$, and a set of corresponding sketch images $T_s = [T_{s1} \ T_{s2} \ \dots \ T_{sN}]$, where T_{fi} and T_{si} is column vector, derived from the i^{th} face and sketch images, respectively. We do following steps:

- (1) Form a new hybrid space $H = [h_1 \ h_2 \ \dots \ h_N]$, and Calculate the mean of h_i for $i=1..N$, at the end we will get M_h . Then we get the mean of sketch space M_s by keeping only the lower half of M_h .
- (2) Let $X_i = h_i - M_h$, then form the covariance matrix of hybrid space $X_h = [X_1 \ X_2 \ \dots \ X_N]$
- (3) Assume:

$$(X_h^T X_h) V = V \Lambda \quad (7)$$

Then the eigenvector matrix V_h of $X_h X_h^T$ can be computed from

$$V_h = X_h V \Lambda^{-1} \quad (8)$$

- (4) The eigensketch matrix V_s can be obtained from keeping only the lower half of V_h

In recognition phase, for every testing sketch image Q_i , we do the following steps:

- (1) Subtract the mean M_s from Q_i to get S_i
- (2) We compute the partial dimension projection coefficients of S_i into pseudo-eigenspace using pseudo-inverse approach.
$$P_i = (V_s^T V_s)^{-1} V_s^T S_i \quad (9)$$
- (3) Use the partial projection coefficient P_i to reconstruct face images R_h in hybrid space, i.e:
$$R_h = V_h \cdot P_i + M_h$$
- (4) Retrieve the reconstructed face image R_f by only keeping the upper half of R_h
- (5) Build an OTSDF correlation filter F_i from R_f
- (6) Match each training images with F_i , and get a correlation plane, from this plane, a PCE score is calculated
- (7) Classify the sketch image Q_i as the person with whom the resulting PCE score is the highest.

For recognition with one-nearest-neighbour method (1NN), we do not build OTSDF for R_f , instead, we compute the Euclidean distance between R_f and every training images, and classify Q_i as the person with whose training images it has the shortest distance.

3.3. Results

In order to see the performance of the proposed method, we compare it with 1NN classifier. Moreover, we also use different number of eigenfaces to see if the proposed method degrades gracefully when the lower quality of reconstructed images. In addition, all the experiment results are based on rank-1 recognition (figure 5).

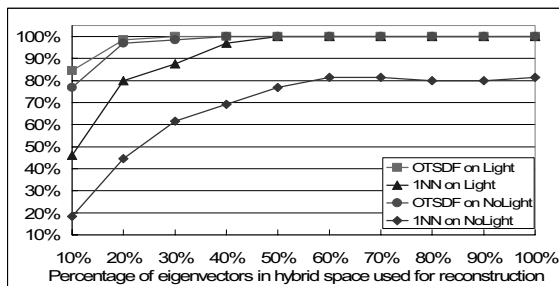


Figure 5: Classification rate with OTSDF and 1NN on PIE-Light and PIE-NoLight datasets.

3.4. Conclusions

When experimenting on PIE Light database, which is a relatively easier task, we can get 100% recognition rate with either the OTSDF or 1NN method. However, OTSDF can achieve 100% even when only 30% of eigenvectors are used, while the 1NN can only achieve recognition rate of 87.69%. When experimenting on the CMU PIE NoLight dataset, which is a much more challenging task due to much harder illumination conditions, the OTSDF approach clearly outperformed in all experiments, showing its capabilities to

perform illumination tolerant face recognition. From these experimental results, we can conclude that our proposed novel face synthesis from sketch approach coupled with advanced correlation filters for face recognition is a successful solution to this problem and is more feasible to work in real world system.

4. FUTURE WORK

We would like to extend our experiment to use bigger database, like FERET [11] and the Notre Dame FRGC [12] to see how the proposed method performs in large scale experiment.

Acknowledgements

We would like to thank the U.S. Technical Support Working Group(TSWG) and for CMU CyLab for supporting this work.

REFERENCES

- [1]. W. Zhao, R. Chellappa, A. Rosenfeld and P. J. Phillips, "Face Recognition: A Literature Survey", CS-Tech Report-4167, University of Maryland, 2000.
- [2]. X. Tang and X. Wang, "Face Photo Recognition Using Sketch", Proc. of Int. Conf. Image Processing, 2002.
- [3]. X. Tang and X. Wang, "Face Sketch Recognition", IEEE Trans. on Circuit System and Video Technology, Vol.14, pp.50-57, 2004.
- [4]. X. Tang and X. Wang, "Face Sketch Synthesis and Recognition", Proc. of Int. Conf. Computer Vision, 2003.
- [5]. Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A Nonlinear Approach for Face Sketch Synthesis and Recognition", Proc. Of. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [6]. M.A. Turk and A. P. Pentland, "Face recognition using eigenfaces", presented at Computer Vision and Pattern Recognition, 1991. Proceeding CVPR '91., IEEE Computer Society Conference on, 1991.
- [7]. B.V.K. Vijaya Kumar, "Tutorial Survey of Composite Filter designs for Optical Correlators," Applied Optics, vol. 31, pp. 4773-4801, 1992
- [8]. A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," Applied Optics, vol. 26:3633-3640, 1987.
- [9]. A. Mahalanobis, B.V.K. Vijaya Kumar, S. Song, S.R.F. Sims, and J.F. Epperson, "Unconstrained correlation filters," Applied Optics, vol. 33:3751-3759, 1994.
- [10]. B. V. K. V. Kumar, D. Carlson, and A. Mahalanobis, "Optimal tradeoff synthetic discriminant function (OTSDF) filters for arbitrary devices," Opt. Lett., vol. 19, pp. 1556-1558, 1994.
- [11]. P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, October 2000.
- [12]. P. J. Phillips, "Face Recognition Grand Challenge", presented at Biometric Consortium Conference, 2004