ASYMMETRIC 3D/2D FACE RECOGNITION BASED ON LBP FACIAL REPRESENTATION AND CANONICAL CORRELATION ANALYSIS

Di Huang[1], Mohsen Ardabilian[1], Yunhong Wang[2], and Liming Chen[1]
[1]MI Department, LIRIS Laboratory, CNRS 5205, Ecole Centrale de Lyon, 69134, Lyon, France
[2]School of Computer Science and Engineering, Beihang University, 100091, Beijing, China

ABSTRACT

In the recent years, 3D Face recognition has emerged as a major solution to deal with the unsolved issues for reliable 2D face recognition, i.e. lighting condition and viewpoint variations. However, 3D method is currently limited by its registration and computation cost. In this paper, we propose to investigate a solution named asymmetric face recognition scheme, enrolling people in 3D environment but performing authentication in 2D. The goal is to limit the use of 3D data to where it really helps to improve recognition performances. In our approach, Local Binary Patterns (LBP) is used as an efficient facial representation for both 2D texture and 3D range images. A weighted Chi square distance is computed as matching score between 2D LBP facial representations; Canonical Correlation Analysis (CCA) is applied to learn the mapping between the LBP-based range face images (3D) and LBP facial texture images (2D). Both matching scores are further fused to obtain the final result. Compared with traditional 2D/2D algorithms, the proposed asymmetric face recognition solution scheme achieves better accuracy; while avoiding the high cost of data acquisition and computation in 3D/3D approaches.

Index Terms— Face recognition, Local Binary Patterns (LBP), Canonical Correlation Analysis (CCA), asymmetric face recognition, face matching and fusion scheme

1. INTRODUCTION

Machine-based face recognition is an important and popular issue in computer vision for its wide application potential and scientific challenges. As compared to other biometrics, such as fingerprint or iris, face offers advantages that place it potentially as the best choice for person identification task. Unfortunately, in despite of the great progress made within the field [1], face remains a biometrics still not very reliable and one needs to deal with variations of lighting condition, pose, facial expressions, etc.

3D face recognition has emerged in the recent years as a major solution to handle the unsolved issues for reliable 2D face recognition, i.e., lighting condition and viewpoint variations [2, 3]. However, 3D approach is currently limited by its registration and computation cost. Generally, for face recognition, data of gallery and probe set are desired to have the similar properties: 2D or 3D, color or gray, and even demanded to be captured by the same sensors. However, the fact that relationship between different categories of data can be known in [4, 5], makes it possible to learn the mapping between 2D face images and 3D face models.

Very few works in the literature has addressed such an asymmetric face recognition problem so far. Rama *et al.* [6] proposed to project the 3D texture information in cylindrical coordinate, and apply Partial Principle Component Analysis (P^2CA) for feature extraction, but there is a high correlation between gallery and probe feature vectors, since the probe of the 2D face image in P^2CA subspace only contains partial information of the face. Riccio *et al.* [7] proposed to utilize control points to compute several geometrical invariants for recognition, but it is not easy to locate these control points accurately. More recently, in [8], Yang *et al.* applied CCA directly to learn the mapping between 2D face image and 3D face data; only 3D range images are used for enrollment, both of which leave the space to improve the accuracy.

In this paper, we propose to investigate an asymmetric face recognition solution, i.e., enrolling people in 3D while performing authentication in 2D using, for instance, 2D face texture image acquired by a simple webcam. The goal is to limit the use of 3D data to where it really helps to improve the face recognition performances. In our 3D/2D approach, 3D face models, from each of which one range image and one texture image is extracted, are exploited for enrollment while only 2D face texture image is utilized for probe. Our approach splits the face recognition task into two steps, a matching step respectively processed in 2D/2D and 3D/2D then a fusion step combining two matching scores. Since all the human faces are quite similar and the major challenge in recognition is how to tolerate within-class variations whilst discriminate different classes well, Local Binary Patterns (LBP) is introduced for facial representation as it highlights the local structures of an image in 2D/2D matching, which is proved effective for one image face recognition (only one image in gallery for each person) [9]. For 3D/2D matching, LBP is also applied as a preprocessing technique [10] not only to reduce illumination changes for 2D texture images, but also amplify the detail of 3D range images; Canonical Correlation Analysis (CCA) method is exploited to learn the mapping between LBP faces of range and texture images.

Compared with traditional 2D/2D face recognition methods, the proposed solution scheme provides better performance, and compared with 3D/3D approaches, it reduce the cost for both data acquisition and computation.

The remainder of this paper is organized as follows: LBP facial representation technique is introduced in section 2, and section 3 presents the CCA methodology. Section 4 describes the framework of the proposed asymmetric fusion scheme. Experimental results are presented and discussed in section 5. Section 6 concludes the paper.

2. LBP FACIAL REPRESENTATION

Local Binary Patterns (LBP), a non-parametric method, that summarizes the local structures of an image efficiently, has received increasing interest for facial representation recently [11]. After first proposed for texture description, it has been widely introduced in many applications. The most important properties of LBP are its tolerance against the variations of monotonic illumination and its computational simplicity.

Specifically, the original LBP operator labels each pixel of one image by thresholding a 3×3 neighborhood with the value of central pixel and considering the result as a binary number, of which the corresponding decimal number is used for labeling. Fig. 1 illustrates such a process. The derived binary number is called Local Binary Pattern or LBP code.

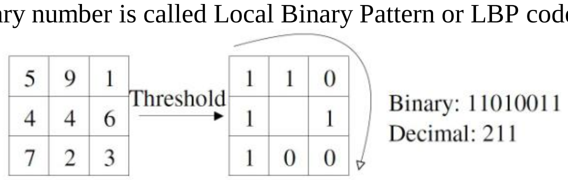

Fig. 1. *An example of the original LBP operator.*

Formally, given a pixel at (x_c, y_c), the resulting LBP can be expressed in decimal form as:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c)2^n \tag{1}$$

where n runs over 8 neighbors of the central pixel, i_c and i_n are gray-level values of central pixel and surrounding pixels, and the function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1 & if \quad x \geq 0 \\ 0 & if \quad x < 0 \end{cases} \tag{2}$$

According to Eqn. (1) and Eqn. (2), the LBP operator is invariant to the monotonic gray-scale transformations which preserve pixel intensity order in local neighborhoods. The histogram of LBP labels calculated over a region is used as a texture descriptor.

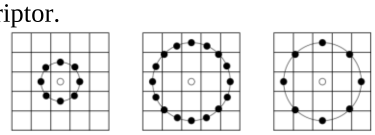

Fig. 2. *Examples of operators: circular (8, 1), (16, 2), and (8, 2).*

To deal with the texture at different scales, the original LBP operator was extended to neighborhoods of different sizes. Local neighborhood is defined as a set of sampling points evenly spaced on a circle centered at the pixel to be labeled, and the sample points that do not fall in the pixels are expressed using bilinear interpolation, thus allowing any radius and number of sampling points in the neighborhood. Fig. 2 shows some examples of the extended LBP operators; the notation *(P, R)* denotes a neighborhood of *P* sampling points on a circle of radius of *R*.

The LBP operator $LBP_{(P,R)}$ produces 2^p different output values, corresponding to 2^p different binary patterns formed by the *P* pixels in the neighborhood. It has been shown that certain patterns contain more information than others [12]. It is possible to use only a subset of the 2p binary patterns to describe the texture of the images. Ojala *et al*. named these patterns as uniform patterns, denoted $LBP^{U2}_{(P,R)}$. A local binary pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vise verse when the corresponding bit string is considered circular.

The general idea for LBP facial representation is that a face image can be seen as a composition of micro-patterns which are described by the LBP operator. But the histogram of LBP computed over the whole face image encodes only occurrences of micro-patterns without any indication about their locations. To also consider shape information of faces, the images are proposed to be divided into a certain number of local regions, from which LBP histograms are extracted. These LBP histograms are then concatenated into a single, spatially enhanced histogram, containing both local texture and global shape information of the face images. Since the psychophysical findings demonstrated some facial features play more important roles than the others, it can be expected that some facial regions contribute more than the others to extra-personal variance. Hence, different regions in the face image can be further weighted according to the importance of inside information. The similarity between two faces can be calculated by comparing the final LBP features.

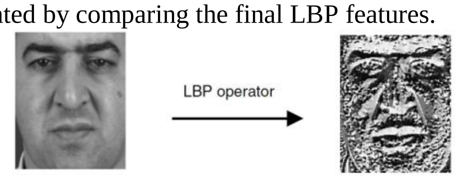

Fig. 3. *Face image (left) processed by LBP operator (right) [10].*

LBP can also be adopted as a preprocessing method on face images mainly for removing or decreasing the effects caused by the illumination changes. For instance, Heusch *et al.* [10] introduced LBP operator for pretreatment, and other methods for feature extraction act on the obtained LBP face (see Fig. 3 for an example). This claim is also supported by [13, 14], proving LBP achieves promising performance for illumination compensation and normalization.

3. CCA LEARNING METHOD

Canonical Correlation Analysis (CCA), a powerful analysis algorithm [15], is especially fully qualified for relating two sets of variables, by maximizing the correlation in the CCA subspace. Similar to Principal Components Analysis (PCA), CCA also reduces the dimensionality of original variables, while unlike PCA; CCA considers the relationship between two variable spaces in the correlation sense, which makes them better suited for regression tasks than PCA.

Given N pairs of samples (x_i, y_i) of (X, Y), i=1, 2,…, N, where $X \in \mathbb{R}^p$, $Y \in \mathbb{R}^q$. The means of both X and Y are zero. The goal of CCA is to learn a pair of directions w_x and w_y to maximize the correlation between two projections $x=w_x^T X$ and $y=w_y^T Y$, where T denotes the transpose. In the context of CCA, the two projections: x and y are also referred to as canonical variants. Formally, the directions can be found as maxima of the function:

$$\rho = \frac{E[w_x^T X Y^T w_y]}{\sqrt{E[w_x^T X X^T w_x] E[w_y^T Y Y^T w_y]}} \tag{3}$$

Where $E[f(x, y)]$ denotes empirical expectation of $f(x, y)$.

The covariance matrix of (X, Y) is

$$C(X,Y) = E\left[\begin{pmatrix} X \\ Y \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}^T\right] = E\left[\begin{pmatrix} C_{XX} \\ C_{XY} \end{pmatrix}\begin{pmatrix} C_{YX} \\ C_{YY} \end{pmatrix}^T\right] \tag{4}$$

Where C_{XX} and C_{YY} are within-set covariance matrices; C_{XY} and C_{YX} are between-sets covariance matrices.

Hence, ρ can be rewritten as

$$\rho = \frac{w_x^T C_{xy} w_y}{\sqrt{w_x^T C_{xx} w_x w_y^T C_{yy} w_y}} \tag{5}$$

Let

$$A = \begin{pmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{pmatrix}, \quad B = \begin{pmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{pmatrix} \tag{6}$$

It can be shown that the solution $W=(w_x^T, w_y^T)^T$ amounts to the extremum points of the Rayleigh quotient [8]:

$$r = \frac{W^T A W}{W^T B W} \tag{7}$$

The solution w_x and w_y can be obtained as solutions of the generalized eigen-problem:

$$AW = BW\lambda \tag{8}$$

4. ASYMMETRIC FUSION SCHEME

The proposed asymmetric face recognition fusion scheme is presented in this section. Recall that our asymmetric scheme uses two categories of face images, one range image I_{GR} and one texture image I_{GT}, extracted from each 3D face model in the gallery while 2D face texture images are used for probe. Therefore, our solution includes two independent matching steps for probe face images I_P: (1) 2D/2D face recognition based on LBP facial representation, and also (2) 3D/2D face recognition based on CCA learned LBP faces. The matching scores of both steps are then fused for final decision. Fig. 4 presents the framework of the asymmetric fusion scheme.

4.1. 2D/2D Face Matching

As follows our discussion on LBP in Section 2, LBP is used as an effective facial representation for 2D/2D step of face recognition. LBP face image I_{GT} and I_P first are divided into m regions, from each of which a histogram h_i (i=1, 2,…, m) is extracted, and the final histogram feature H_{GT} and H_P can be achieved by concatenating the separate histograms h_1, h_2,…, h_m. The weighted Chi square distance (see Eqn. (9)) is applied to compute the matching score $d_{2D/2D}$,

$$\chi^2_\omega(H_{GT}, H_P) = \sum_{j,i} \omega_j \frac{\left(H_{GTi,j} - H_{Pi,j}\right)^2}{H_{GTi,j} + H_{Pi,j}} \tag{9}$$

where i and j refer to the i^{th} bin in histogram of the j^{th} local region and ω_j is the weight for region j.

4.2. 3D/2D Face Matching

Also as follows our discussion in section 2, LBP is used to preprocess 3D range images and 2D face texture images. To each range image I_{GR} and texture image I_P, LBP faces F_{GR} and F_P are extracted respectively before the matching.

The linear CCA algorithm is introduced for learning the mapping between the LBP faces of range image and texture image. In the training process, N pairs of 3D/2D LBP faces are given as $(F_{GR}, F_P) = \{(f_{GRk}, f_{Pk})\}$, ($k$=1, 2,…, N), where (f_{GRk}, f_{Pk}) is a corresponding pair of 3D and 2D faces. Two directions w_{GR} and w_P are learned, and $w_{GR}^T F_{GR}$ and $w_P^T F_P$ are best correlated. In the matching process, LBP faces, F_{GR} and F_P, are first projected into CCA subspace:

$$F'_{GR} = w_{GR}^T F_{GR}, \quad F'_P = w_P^T F_P \tag{10}$$

The matching score, $d_{3D/2D}$, is calculated according to the following normalized correlation function:

$$S(F'_{GR}, F'_P) = \frac{F'_{GR} \bullet F'_P}{\|F'_{GR}\| \|F'_P\|} \tag{11}$$

In the end, by Min-Max normalization, matching scores $d_{2D/2D}$ and $d_{3D/2D}$ are normalized to the interval of [0, 1], and the sum of two corresponding matching scores is considered as the discriminant criterion for the final decision.

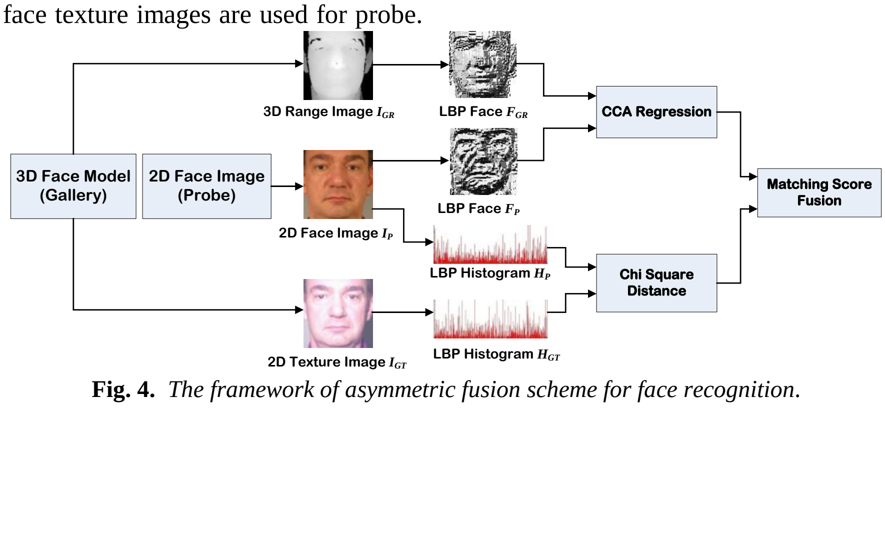

Fig. 4. *The framework of asymmetric fusion scheme for face recognition.*

5. EXPERIMENTS

The dataset for evaluating the proposed scheme contains 200 subjects selected from FRGC 1.0 and 2.0 datasets so that for each of subjects, there are enough 2D face images with the illumination changes and slight variations of head pose and facial expression. In our case, five 2D images are for probe. Fig. 5 shows several samples of the database: the first two images are range and texture image extracted from 3D face model for gallery; the last two are 2D face images for probe. All the images are cropped to 80x80 pixels.

Fig. 5. *Some samples of the database.*

For experiments, the database is divided into two parts; one for training CCA consists of 170 subjects, and the other for testing contains the remaining 30 subjects. To evaluate the performance of proposed asymmetric fusion scheme, for each subject, one 3D range image and one 2D texture image are used for gallery while probe contains 2D texture images only. In our experiments, LBP operator used in 2D/2D face recognition step is $LBP^{U2}_{(8,2)}$, and the images are divided into 8x8 square blocks of the same size 10x10 pixels; the weight strategy comes from [11].

Table 1. *Recognition rates of different schemes.*

Scheme	Recognition Rate
(1) 2D/2D $LBP^{U2}_{(8,2)}$	0.7388
(2) 3D/2D CCA Original Images	0.4000
(3) 3D/2D CCA LBP Faces	0.5420
(1)+(2)	0.7667
(1)+(3) Proposed Asymmetric Scheme	0.8236

The recognition rates are given in Table 1. We can see that the performance of proposed asymmetric fusion scheme by combining both matching steps is better than that of any single matching; for 3D/2D matching, CCA learning based on LBP faces outperforms that based on the original images.

6. CONCLUSIONS

In this paper, an asymmetric fusion scheme is proposed for face recognition, which utilizes the range image and texture image extracted from 3D face model for enrollment and 2D face image for test. As a result, the proposed fusion scheme contains two steps: 2D/2D face recognition and 3D/2D face recognition. LBP is applied as facial description for both 2D texture image and 3D range image in order to highlight the local structure variations on face images and also to reduce the influences by illumination changes. CCA is used to learn the mapping between two types of LBP faces. The matching scores of both steps are further fused to make final decision.

The Experimental results demonstrate that the proposed fusion scheme achieves better performance than only using single matching, showing that 2D/2D and 3D/2D provide the complementary information for the face recognition task. Moreover, for 3D/2D face matching, CCA learning based on LBP faces outperforms that on the original face images, proving the fact that LBP is efficient for the preprocessing again.

To sum up, compared with traditional 2D/2D algorithm, the proposed solution achieves better recognition accuracy with additional information provided by 3D/2D matching; and compared with 3D/3D methods, it decreases the cost for both data acquisition and computation, for only 2D images are required for probe.

In our future work, we will test our approach on other datasets, in particular IV2 dataset [16, 17].

7. REFERENCES

[1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, “Face recognition: a literature survey,” *ACM Computing Surveys*, 2003.

[2] A. Scheenstra, A. Ruifrok, and R.C. Veltkamp, “A survey of 3D face recognition methods,” *AVBPA*, 2005.

[3] K.W. Bowyer, K. Chang, and P. Flynn, “A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition,” *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1-15, 2006.

[4] M. Reiter, R. Dormer, G. Langs, and H. Bischof, “3D and infrared face reconstruction from RGB data using canonical correlation analysis,” *ICPR*, 2006.

[5] D. Yi, R. Liu, R. Chu, Z. Lei, S. Li, “Face matching between near infrared and visible light images,” *ICB*, 2007.

[6] A. Rama, F. Tarres, D. Onofrio, and S. Tubaro, “Mixed 2D-3D information for pose estimation and face recognition,” *ICASSP*, 2006.

[7] D. Riccio and J.-L. Dugelay, “Geometric invariants for 2D/3D face recognition,” *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1907-1914, 2007.

[8] W. Yang, D. Yi, Z. Lei, J. Sang, and S. Li, “2D-3D face matching using CCA,” *FG*, 2008.

[9] X. Tan, S. Chen, Z. Zhou, and F. Zhang, “Face recognition from a single image per person: a survey,” *Pattern Recognition*, vol. 39, no. 9, pp. 1725-1745, 2006.

[10] G. Heusch, Y. Rodriguez, and S. Marcel, “Local binary patterns as an image preprocessing for face authentication,” *FG*, 2006.

[11] T. Ahonen, A. Hadid, and M. Pietikäinen, “Face recognition with local binary patterns,” *ECCV*, 2004.

[12] T. Ojala, M. Pietikäinen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. PAMI*, vol. 24, no. 7, pp. 971–987, 2002.

[13] R. Gross and V. Brajovic, “An image preprocessing algorithm for illumination invariant face recognition,” *AVBPA*, 2003.

[14] J. Ruiz-del-Solar, and J. Quinteros, “Illumination compensation and normalization in eigenspace-based face recognition: a comparative study of different pre-processing approaches,” *Pattern Recognition Letters*, vol. 29, no. 14, pp. 1966–1979, 2008.

[15] D.R. Hardoon, S.R. Szedmak, and J.R. Shawe-taylor, “Canonical correlation analysis: an overview with application to learning methods,” *Neural Comput.*, vol. 16, no. 12, pp.2639–2664, 2004.

[16] D. Petrovska-Delacretaz, S. Lelandais, J. Colineau, L. Chen, B. Dorizzi, M. Ardabilian, E. Krichen, M.-A. Mellakh, A. Chaari, S. Guerfi, J. D’Hose and B. BenAmor, “IV2 multimodal (2D, 3D, stereoscopic face, talking face and iris) biometric database, and the IV2 2007 evaluation,” *BTAS*, 2008.

[17] J. Colineau, J. D'Hose, B. BenAmor, M. Ardabilian, L. Chen, and B. Dorizzi, “3D face recognition evaluation on expressive faces using the IV2 database,”*ACIVS*, 2008.