3DLBP and SVD Fusion for 3D Face Recognition Using Range Image

El Mahdi Barrah, Rachid Ahdid, Said Safi, Abdessamad Malaoui*

Interdisciplinary Laboratory of Research in Sciences and Technologies (LIRST), Sultan Moulay Slimane University, Bni Mellal, Morocco

Copyright ©2017 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract In this paper, we present a novel approach for fusing 3D Local Binary Patterns (3DLBP) and singular Value Decomposition (SVD) for face recognition when "Kinect" is used as the 3D face scanner. Another approach is used to compare the 3DLBP method, fused with SVD, with other methods proposed in the literature for face recognition by using Kinect. Experimental results on FRGC 2.0 face dataset showed that the generated data by Kinect are discriminating enough to allow face recognition and that 3DLBP performs better than the other methods.

Keywords Face Recognition, 3D Local Binary Patterns (3DLBP), Singular Value Decomposition (SVD)

1 Introduction

Face recognition is a crucial part of many contemporary applications. Recent key applications in fields such as human-computer interface, identity verification, criminal face identification, and surveillance systems need a reliable face recognition algorithm. Unfortunately, Human face is similar in their configuration and hence offer low distinctiveness. Local Binary Patterns (LBPs) [1, 2] are a technique that has been widely applied to the problem of facial recognition [3]. Variants on LBPs have also been proposed: local gradient orientation binary patterns (LGOBPs) [4], local phase quantisation (LPQ) [5], and local Gabor binary patterns (LGBPs) [6]. LBPs, LPQs and LGBPs have also been extended to the dynamic problem in the form of LBP-TOP [7], LPQ-TOP [8] and V-LGBPs [9]. Some works have also begun to apply descriptive of this kind to the 3D problem by proposing the 3DLBP, the traditional LBP descriptor applied to the depth map of a facial mesh [10] and the Multiresolution Extended Local Binary Pattern (MELBPs) [11].

Recently, 3DLBP (3D Local Binary Pattern) method [1] was proposed for 3D face recognition by using high resolution scanners. The main goal of this work is to assess the performance of 3DLBP method, fused with SVD (Singular

Value Decomposition) face descriptor method [2], for face recognition. Another goal of our work is to analyze the performance of the 3DLBP method fused with an SVD face descriptor when compared with other methods proposed in the literature for face recognition.

The paper is organized as follows. An overview of methods used for face recognitions techniques is given in Section 2. In section 3, we present the proposed fusion scheme. The simulation results are given in Section 4, and finally the conclusion is drawn in Section 5.

2 Approaches used in this work

The global and local features demonstrate different characteristics in representing the face images. In general, in case of the global approaches a small size of feature vector is created by extracting the features from the whole image and it is able to depict the common characteristics of the faces. the proposed method includes the SVD approach in order to obtain the global features wherein some selected number of left and right singular vectors is used to create the feature vector. Likewise, the 3DLBP descriptor is able to efficiently extract the local details of face images, i.e. image edges, peaks, etc., even in the presence of noise. The 3DLBP features are invariant to changes in lighting and scale.

2.1 Local Binary Pattern

This section reports the performance of a feature as LBP extensions. Prior to matching, the face have been normalized using an affine transformation [2]. Local binary patterns are feature vectors extracted from a gray-scale image by applying a local texture operator at all pixels and then using the result of the operators to form histograms that are the feature vectors. The original LBP operator is constructed as follows: Given a 3x3 neighbourhood of pixels as shown in Figure 3, a binary operator is created for the neighbourhood by comparing the center-pixel to its neighbours in a fixed order, from the left-center pixel in counter-clockwise manner. If a neighbour has a lower intensity than the center pixel it is assigned

a zero, otherwise a one. This will yield an 8-bit binary number, whose decimal valued entry in a 256 bin histogram is increased by one. The complete LBP-histogram of an image will then depict the frequency of each individual binary pattern in the image. Due to its design, the feature vectors are robust to monotonic intensity variations since the LBP-operator is not affected by the size of the intensity difference. The feature vectors are not affected by small translations of the face either since the same patterns will be accumulated in the histogram regardless of their positions. The calculation of the LBP codes can be easily done in a single scan through the image. The value of the LBP code of a pixel (x_c, y_c) is given by :

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (1)$$

Where g_c represent the gray value of the center pixel (x_c, y_c) , g_p refers to gray values of P equally spaced pixels on a circle of radius R and s define the thresholding function as follows:

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

Figure 1. LBP encoding scheme.

The operator can be extended from its nearest neighbours by instead defining a radius R where a chosen number of P points are sampled. The intensity values of the points are then calculated using bilinear interpolation. The number of points will then determine the number of possible binary patterns and also the length of the feature vector. To reduce the length of the feature vectors, Ojala et al. [11] found that patterns with at most 2 binary transitions (0 to 1 or 1 to 0) provides over 90 of all spatial texture patterns.

2.2 3D Local Binary Patterns

Local Binary Patterns (LBP) operator is first proposed by Ojala et al. [11] for texture description. It has been successfully applied to 2D face recognition by Timo et al. [12]. Also we have tested it for 3D face recognition. Motivated by the original LBP, we propose 3D Local Binary Patterns (3DLBP) operator to obtain local correlative features of facial surfaces. In 3DLBP, not only the original LBP are included, but also the information of depth differences is encoded into binary patterns. The main idea of the original LBP operator is described in Figure 2.

As shown in Figure 2, every pixel in the image is subtracted by their neighbors at the first step. Then the differences are converted to binary units: 0 or 1 according to their signs. At the third step, binary units are arranged clockwise.

Figure 2. An example of Basic LBP operator.

By doing this, we can obtain a set of binary units as the local binary pattern of the pixel. The binary pattern is further transformed to decimal number or uniform pattern (see [14] for more details). Two parameters (P, R) are used to control the selection of the number of neighbors (P) and their locations radius (R). They can be varied as (8, 2), (16, 2), (24, 3) etc.

From the process above, we can see that LBP operator actually encodes pixels relationships with their neighbors. We name these relationships as correlative features in this paper. So LBP can be seen as a kind of local correlative features. In our opinion, the structure information of facial surface should exist in the correlative features of points on the surface. According to this intrinsic property of LBP analyzed above, we suppose LBP operator should own the potential power to encode structure information of 3D face surface.

Furthermore, in the process of LBP operator, only encoding signs of depth differences is not adequate for describing 3D faces, because different depth differences on the same point of the facial surfaces distinguish different faces. See a specific example in Figure 3. Though A and B are two different persons, LBP of their nose tips are the same, because all the points around the nose tips are lower than them. So we can see that, if two facial regions of different persons on the same place have the same trend of depth variation, LBP is inadequate to distinguish them. However, though the signs of differences between the two nose tips and their neighbors are the same, the exact values of the differences are different. We consider this clue should be crucial for 3D face recognition.

So we further encode the exact values of differences into binary patterns as what is shown in Figure 3. According to our statistical analysis, more than 93 % percent of the depth differences between points in $R = 2$ are smaller than 7, so we add three binary units to encode each depth difference between a pixel and its neighbor. Three binary units ($i_2 i_3 i_4$) can correspond to the absolute value of depth difference (DD) : 0-7. All the $|DD| \geq 7$ are assigned to 7. Combining with signs of the differences denoted by 0, 1 as the head binary unit (i_1) like what the original LBP does, we finally obtain a set of 4 binary units $i_1 i_2 i_3 i_4$ to denote DD between two points as follows.

$$i_1 = \begin{cases} 1 & \text{if } DD \geq 0 \\ 0 & \text{if } DD < 0 \end{cases} \quad (3)$$

$$|DD| = i_2 * 2^2 + i_3 * 2^1 + i_4 * 2^0 \quad (4)$$

A specific example is shown in the blue window of Figure 3. Four binary units are divided into four layers. The binary units of each layer are arranged clockwise. Finally, we will get four decimal numbers: P1, P2, P3, P4 at each pixel point as its representative (see Figure 2 for an example). We call these as 3D Local Binary Patterns (3DLBP). For matching, 3DLBP are first transformed into four maps according to P1, P2, P3, P4 respectively: 3DLBPMAP1 (equal to original LBPMAP), 3DLBPMAP2, 3DLBPMAP3, 3DLBPMAP4. Then histograms of local regions of the four maps are concatenated as a local statistics of correlative features (LSCF) for matching.

Figure 3. An example of 3DLBP and its comparison to LBP.

2.3 Process of Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) represents a significant topic in linear algebra. SVD has many practical and theoretical values? special feature of SVD is that it can be performed on any real (m, n) matrix. Let's say we have a matrix A with m rows and n columns, with rank r and $r \leq n \leq m$. Then the A can be factorized into three matrices:

$$A = USV^T \quad (5)$$

Figure 4. Illustration of Factoring A to USV^T .

Where Matrix U is an $m * m$ orthogonal matrix

$$U = [U_1, U_2, \dots, U_r, U_{r+1}, \dots, U_m]$$

The column vectors U_i , for $i = 1, 2, \dots, m$, form an orthonormal set:

$$U_i^T U_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (6)$$

And matrix V is an $n * n$ orthogonal matrix.

$$V = [V_1, V_2, \dots, V_r, V_{r+1}, \dots, V_m]$$

Column vectors V_i for $i = 1, 2, \dots, n$ form an orthonormal set:

$$V_i^T V_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (7)$$

Here, S is a $m * n$ diagonal matrix with singular values (SV) on the diagonal. The matrix S can be showed in the following matrix

$$S = \begin{bmatrix} \sigma_1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \sigma_{r+1} & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & \sigma_n \end{bmatrix} \quad (8)$$

For $i = 1, 2, \dots, n$, σ_i are called Singular Values (SV) of matrix A. It can be proved that

$$\begin{aligned} \sigma_1 &\geq \sigma_2 \geq \dots \geq \sigma_r \geq 0 \\ \text{and} \\ \sigma_{r+1} &= \sigma_{r+2} = \dots = \sigma_N = 0 \end{aligned} \quad (9)$$

The V_i 's and U_i 's are called right and left singular vectors of A [1].

3 Proposed method

In this work we propose the fusion of 3DLBP method with Singular Value Decomposition (SVD) face descriptor method for face recognition, three main steps are applied to the dataset: segmentation, feature extraction, and classification.

The segmentation is centered at the subjects nosetip, and then a square region with the width and height equal to the Euclidean distance of the subjects eyes is cropped, This cropped region is downsized to a 32x32 square region, which is utilized in the next steps for the feature extraction.

The feature extraction step is carried out by applying the 3DLBP on all 32x32 face region. Each region that goes through the feature extraction has a histogram describing it, in case of having multiple regions in the same face image the face descriptor is created by concatenating all 3DLBP histograms.

The face descriptor obtained in the previous step are used to train a SVM (Support Vector Machine) classifier. The whole process is summarized in the Fig. 5.

Figure 5. Diagram of the proposed method.

4 Experimental results

In order to assess the proposed method, two experiments were carried out on the FRGC 2.0 Face Dataset [14]. These experiments, as well as, the database are described in this section.

4.1 FRGC 2.0 database

In order to demonstrate the effectiveness of the proposed method, we use the public database for experiments. The FRGC 2.0 [14] is a benchmark database released last year. To our best knowledge, the 3D face dataset of FRGC 2.0 is the largest 3D face database till now. It consists of six experiments. Our experiments belong to Experiment 3 which measures the performance of 3D face recognition. The 3D data set is divided into training and validation partitions. The validation dataset contains 4007 shape data of 466 different persons. All the 3D data were acquired by a Minolta Vivid 900/910 series sensor in three semesters in two years. The 3D shape data are given as raw point clouds, which contain spikes and holes. And manually labeled coordinates of the eyes, noses, and chin are given.

Figure 6. Persons from FRGC 2.0 database .

4.2 Experiment

The experiment was carried out to compare the results obtained by 3DLBP, SVD, and by their fusion against. In this experiment only the open mouth, smile, neutral and light on images from both sessions of the FRGC 2.0 Dataset were utilized. Those images are divided into two groups: gallery and probe. The gallery was composed of 75% of subject images available in the sets: open mouth, smile and light on. The neutral set of faces was utilized as probe. The Figure 8 shows an example of the two sets, probe and gallery, of a subject in the FRGC 2.0 database.

Figure 7. Gallery and probe for a subject in the FRGC 2.0 database .

The Figure 8 shows the Cumulative Match Characteristic (CMC) curves obtained utilizing the FRGC 2.0 Face Dataset. The fusion between 3DLBP and SVD has the higher identification accuracy at Rank 1 and 2. The 3DLBP method has the higher identification accuracy at Rank 3 and 4, at 5 it ties with the fusion. The SVD method starts with the worst performance but passes SVD at Rank 5.

Figure 8. DET curve for the 3DLBP, SVD, and the fusion between them.

5 Conclusion

In this paper, we proposed a new scheme for 3D face recognition based on range images. From the experimental results presented in section 4 we can conclude that:

- Although the SVD descriptor has the lowest performance in several cases, when fused with 3DLBP it

improves its performance, becoming a good alternative to increase face recognition performance;

- The used of (3DLBP) to encode correlative features of points on facial surface, which has been proved to perform much better than SVD.

All experiments in this work were performed on the full set of 3D face data FRGC 2.0. So finally we can conclude that combination of 3DLBP with SVD, it will offer higher rates of recognition, a superior preprocessing method and other better classification methods, such as SVM, can be adopted to further improve the final performance of our method.

Acknowledgements

I would like to thank to Dr. Abdessamad MALAOUI, Dr. Said SAFI for their assistances, discussions and suggestions.

REFERENCES

- [1] E. BARRAH, S. SAFI and A. MALAOUI. "Recognition of Tifinagh Characters with Missing Parts Using Neural Network." *International Journal of Computer, Electrical, Automation, Control and Information Engineering* Vol:10, No:2, 2016.
- [2] E. BARRAH, S. SAFI and A. MALAOUI. "Extended Set of DCT-TPLBP and DCT-FPLBP for Face Recognition", *International Journal of Computer, Electrical, Automation, Control and Information Engineering* Vol:9, Paper Code: JRNL37640, 2015; World Academy of Science, Engineering and Technology;.
- [3] E. BARRAH, S. SAFI and A. MALAOUI. "New Technique to Determination of Electronic Circuits for Semiconductor Components by Recognizing Nyquist Curve". *Journal of Computer Sciences and Applications*, 2015, Vol. 3, No. 4, 100-104
- [4] C. Shan, al., "Facial expression recognition based on Local Binary Patterns": A comprehensive study IVC, vol. 27, no. 6, pp. 803-816, 2009.
- [5] V. Ojansivu, J. Heikkila, "Blur insensitive texture classification using local phase quantization , *Image and Signal Processing*, pp. 236-243, 2008.
- [6] W. Zhang, al., Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition, *ICCV 2005. IEEE*, 2005, vol. 1, pp. 786-791.
- [7] G. Zhao, M. Pietikainen, Dynamic texture recognition using local binary patterns with an application to facial expressions, *IEEE PAMI*, pp. 915-928, 2007. *IEEE Transactions*
- [8] B. Jiang, M.F. Valstar, M. Pantic, Action unit detection using sparse appearance descriptors in spacetime video volumes in *FG 2011*, March 2011. *IEEE Transactions*.
- [9] S. Xie, al., V-LGBP: Volume based local Gabor binary patterns for face representation and recognition in *ICPR 2008. IEEE*, 2009, pp. 1-4. *IEEE Transactions*.
- [10] Y. Huang, Y. Wang, and T. Tan. Combining statistics of geometrical and correlative features for 3d face recognition. In *Proceedings of the British Machine Vision Conference*, pages 90.190.10. BMVA Press, 2006. doi:10.5244/C.20.90.
- [11] T. Huynh, R. Min, and J.-L. Dugelay. An efficient LBP-based descriptor for facial depth images applied to gender recognition using RGB-D face data. In *ACCV 2012, Workshop on Computer Vision with Local Binary Pattern Variants*, Daejeon, Korea, November 5-9, 2012 / Published also as LNCS, Vol 7728, PART 1, Daejeon, KOREA, DEMOCRATIC PEOPLES REPUBLIC OF, 11 2012.
- [12] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, 24(7):971987, 2002.
- [13] D. Huang, al., A novel geometric facial representation based on multi-scale extended local binary patterns in *FG 2011. IEEE*, 2011, pp. 1-7.
- [14] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, Overview of the face recognition grand challenge, *Proc. of CVPR*, 2005.
- [15] Fehr, J., Burkhardt, H.: 3d rotation invariant local binary patterns. In: *ICPR*. (2008) 14