Article

New Fusion of SVD and DCT-LBP for Face Recognition

El Mahdi Barrah, Said Safi and Abdssamad Malaoui *

Interdisciplinary Laboratory of Research in Sciences and Technologies (LIRST),
Sultan Moulay Slimane University, Beni-Mellal 23000, Morocco;
barrah.inf@gmail.com (E.M.B.); safi.said@gmail.com (S.S.)
* Correspondence: a.malaoui@usms.ma

Abstract: In this paper, we proposed the fusion of two projection based face recognition algorithms: local binary Patterns in DCT domain and singular value decomposition (SVD) characterized by its simplicity and efficiently. Standard databases ORL are used to test the experimental results which prove that proposed system achieves more accurate face recognition as compared to individual method.

Keywords: Face Recognition; Discrete Cosine Transform (DCT); singular value decomposition (SVD)

---

1. Introduction

Face recognition has become a very active research area in recent years mainly due to increasing security demands and its potential commercial and law enforcement applications. Although, face recognition systems have reached a significant level of maturity with some practical success, face recognition still remains a challenging problem due to large variation in face images.
Generally, feature extraction and classification are two fundamental operations in any face recognition system. In order to improve the recognition performance it is necessary to enhance these operations. Feature extraction is used for reducing the dimensionality of the images using some linear or non-linear transformations of face images with successive feature selection, so that exacted feature representation is possible. However, there are some problems such as lightning condition, illumination, various backgrounds, aging and individual variation with feature extraction of human face.
Then, how to solve variation problems in face recognition? It will be a highly challenging task if we want to solve those problems using visual images only. This paper presents the score level fusion of SVD [8, 9]and DCT-LBP [1]. SVD-based method used in our approach considers the left and right singular vectors as a feature matrix because its recognition rate is better than SVD-based method when using singular values as the feature vectors. The DCT-LBP is used instead of LBP in order to reduce the execution time.
The paper is organized as follows. An overview of SVD and local DCT-LBP face recognitions techniques are 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. Overview of SVD and LBP textur methods

2.1. Discrete Cosine Transform

The DCT is a popular technique in imaging and video compression, which transforms signals in the spatial representation into a frequency representation.
The forward 2D-DCT[10] of a M x N block image is defined as:

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) * Cos[\frac{\pi(2x+1)u}{2M}]Cos[\frac{\pi(2Y+1)v}{2N}] \quad (1)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2016 doi:10.20944/preprints201608.0088.v1

The inverse transform is defined as:

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v) * Cos[\frac{\pi(2x+1)u}{2M}]Cos[\frac{\pi(2Y+1)v}{2N}] \tag{2}$$

where

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{M}} & if \quad u = 0 \\ \sqrt{\frac{2}{M}}, & if \quad 1 \leq u \leq M-1 \end{cases}$$

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & if \quad v = 0 \\ \sqrt{\frac{2}{N}}, & if \quad 1 \leq v \leq N-1 \end{cases}$$

and, x and y are spatial coordinates in the image block, and u and v are coordinates in the DCT coefficients block. Fig.1 shows the properties of the DCT coefficients in M x N blocks with the zigzag pattern used by JPEG compression to process the DCT coefficients. Although the total energy remains the same in the M x N blocks, the energy distribution changes with most energy being compacted to the low-frequency coefficients. The DC coefficient is represented by C(0,0) in the forward 2D.

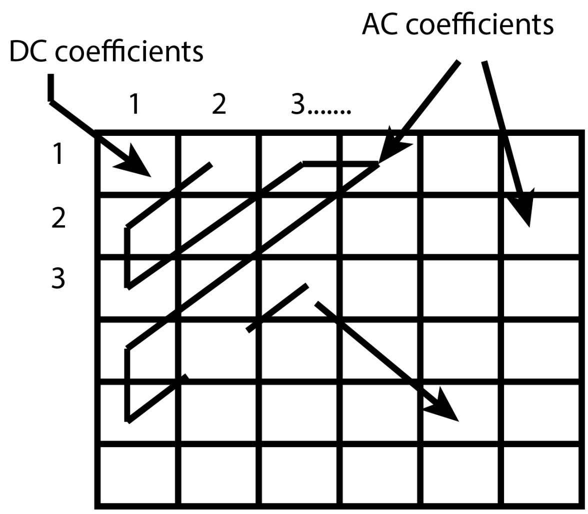

Figure 1. Block feature of DCT coefficients and their selection in zig-zag pattern.

DCT equation. As the cosine of zero is one, the equation is simplified to:

$$C(0,0) = \frac{1}{M*N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \tag{3}$$

The DC coefficient, which is located at the upper left corner, holds most of the image energy and represents the proportional average of the M x N blocks. The remaining ((M x N)) coefficients denote the intensity changes among the block images and are referred to as AC coefficients. The DCT is performed on the entire image obtained after processing the input face images by histogram equalization.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2016 doi:10.20944/preprints201608.0088.v1

For that reason the researchers are investigated to add a step which could reduce the computing time. For example the local 2D DCT [10] transform has been used as a feature extraction step in face recognition (figure 2).

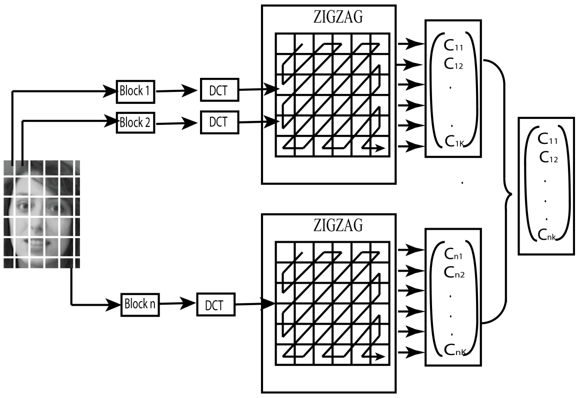

Figure 2. Feature Extraction in DCT-LBP

2.2. Local binary patterns

This section reports the performance of a feature as LBP extensions. Prior to matching, the face have been normalized using an affine transformation [4]. 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 :

$$LBP_{P,R} = \sum_{P=0}^{P-1} s(g_P - g_c)2^P \tag{4}$$

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:

$$s(x) = \begin{cases} 1 & if \quad x \geq 0 \\ 0 & if \quad otherwise \end{cases} \tag{5}$$

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. [7] found that patterns with at most 2 binary transitions (0 to 1 or 1 to 0) provides over 90 of all spatial texture patterns.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2016 doi:10.20944/preprints201608.0088.v1

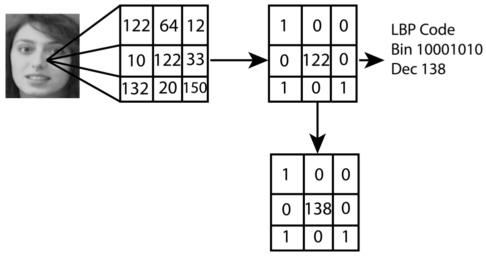

Figure 3. the local binary pattern is determined by comparing the center pixel to its nearest neighbours in the fixed order assigned by the arrow. The binary pattern yields a binary number which is accumulated by one in the LBP histogram.

2.3. *face recognition algorithm based on PCA*

The main idea of the PCA[2] algorithm is to express our M starting images as a basic individual orthogonal vectors (eigenvectors). This is done as follows:

- Transform each learning image column vector, and concatenate all training vectors to form a matrix X.
- Adjust learning data with respect to the average μ.

$$X = X - \mu \tag{6}$$

- Calculate the covariance matrix.

$$G = X^T X \tag{7}$$

- Determine the matrix of eigenvectors W ordered according to the matrix of eigenvalues. $\wedge$, itself sorted in descending order, by solving the following equation:

$$GW = \wedge W \tag{8}$$

- Projecting all training images by the matrix W as follows:

$$Y = W^T X \tag{9}$$

- Transform each image test column vector T and then used to obtain the T_p model and a similarity measure can be used to laclassification.
  To measure the similarity between two vectors, we used the Minkowski distance of order p:

$$L_p = \left(\sum_{i=1}^{n} |X_i - Y_i|^p\right)^{\frac{1}{p}} \tag{10}$$

such as $X = (x_1, ..., x_n)$ et $Y = (y_1, ..., y_n)$.

In our experiments, we used two derived distances from the Minkowski distance, the first obtained by p = 1 called the Manhattan distance (L1) and the second obtained by p = 2 and called the Euclidean distance (L2).

2.4. *Process of Singular Value Decomposition (SVD)*

Singular Value Decomposition (SVD) represent 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2016 doi:10.20944/preprints201608.0088.v1

(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:

$$A = USV^T \tag{11}$$

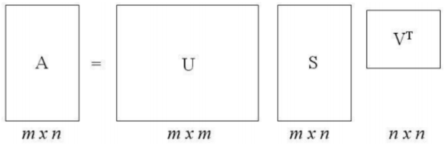

Figure 4. Illustration of Factoring A to USV^T.

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

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

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

$$U_i^T U_j = \delta_{ij} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases} \tag{12}$$

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

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

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

$$V_i^T V_j = \delta_{ij} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases} \tag{13}$$

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

$$S = \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 & 0\cdots & 0 \\ 0 & \sigma_1 & 0 & \cdots & 0 & 0\cdots & 0 \\ \vdots & 0 & \cdots & 0 & 0 & 0\cdots & 0 \\ \vdots & 0 & \cdots & \sigma_r & 0 & 0\cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_r & 0\cdots & 0 \\ \vdots & 0 & \cdots & 0 & 0 & 0\cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0\cdots & \sigma_n \end{bmatrix} \tag{14}$$

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

$\sigma_1 \geq \sigma_2 \geq .... \geq \sigma_r \geq 0$ and $\sigma_{r+1} = \sigma_{r+2} = ...\sigma_N = 0$

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

3. THE PROPOSED APPROACH

In this section we present our methodology for fusing two appearance-based approaches for face recognition: the SVD and the DCT-LBP. Figure 5 shows the block diagram of the proposed method. It is composed of the following steps:

- Representation of the face according to the SVD and the DCT-LBP approaches;
- Computation of the distance vectors D_{SVD} and $D_{DCT-LBP}$ from all the M faces in the database;
- Normalization of these dissimilarity distance vectors;
- Combination of the two vectors according to a given fusion rule.

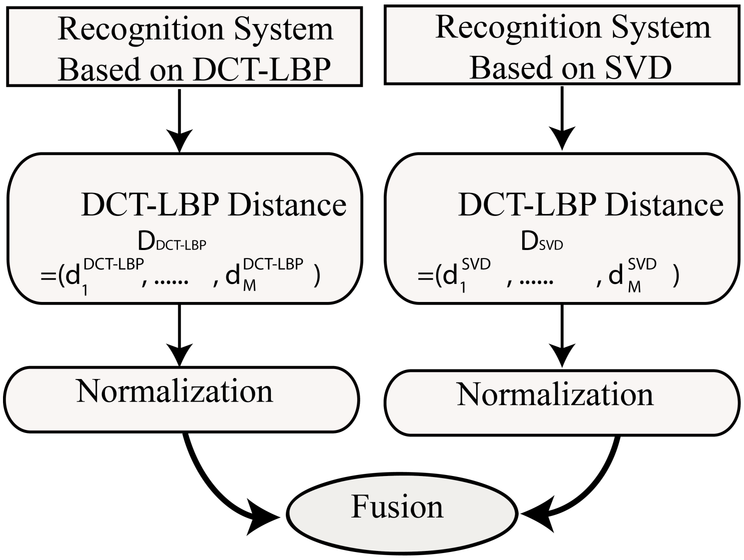

Figure 5. Diagram of the proposed fusion.

3.1. Normalization

Since the dissimilarity distance scores output of different recognition techniques are heterogeneous, score normalization is needed to transform these scores into a common domain, prior to combining them.
In this work we have used MIN-MAX and Z-score normalization techniques.
We denote a distance vector as $D = (d_1, ...d_M)$ and the coresponding normalized score as $D_n = (d_{1n}, ...d_{Mn})$

3.1.1. MIN-MAX method

This method maps the distance vector to the [0, 1] range. The quantities D_max and D_min specify the end points of the distance range [7] and D_n is given by:

$$D_n = \frac{D - D_{min}}{D_{max} - D_{min}} \tag{15}$$

where $D_{min} = min(d_1, ...d_M)$ and $D_{max} = max(d_1, ...d_M)$

 doi:10.20944/preprints201608.0088.v1

3.1.2. Z-score

This method transforms the scores to a distribution with mean of 0 and standard deviation of 1. The operators mean() and std() denote the arithmetic mean and standard deviation operators, respectively [7]:

$$D_n = \frac{D - mean(D)}{std(D)} \tag{16}$$

4. RESULTS AND DISCUSSIONS

In order to demonstrate the effectiveness of the proposed method, we use the public database for experiments The ORL (Olivetti Research Laboratory) [10] for evaluating performance with large number of subject and complicated conditions and to analyze the accuracy on 2D face simples with extreme pose change. It contains 400 images for 40 individuals as shown in Figure 6, for each person we have 10 different images of size 112x92 pixels. For some subjects, the images are captured at different times.
The two sub-tables in table 1 represent the obtained results based on fusion of DCT-LBP layer with different parameters. Recall that P is the number of sampling points and R is the radius value, and SVD.

Figure 6. Persons of the ORL database .

Table 1. RECOGNITION RATES BASED ON DCT-LBP WITH DIFFERENT PARAMETERS with SVD in %.

P=4	**R=2**	**R=4**	**R=6**	**R=8**
DCT-LBP and SVD	94.62	95.20	95.80	98.2
DCT-ACP and SVD	94.62	95.20	95.80	98.2
P=8	**R=2**	**R=4**	**R=6**	**R=8**
DCT-LBP and SVD	94.41	95.43	95.93	98.7
DCT-ACP and SVD	94.62	95.20	95.80	98.2
P=16	**R=2**	**R=4**	**R=6**	**R=8**
DCT-LBP and SVD	94.75	95.82	95.98	98.83
DCT-acp and SVD	94.62	95.20	95.80	98.2

In this simulation, we have randomly selected five persons for learning, and the rest of image for testing. Thus, the total number of training image and testing is 200 for both. We calculate the learning recognition rates of SVD and DCT-LBP.
In Table II, we represent the times of learning and Identification (number of image of learning and identification is 200), times are calculated with the developed software in Matlab. We note that SVD significantly reduces the learning time and identification.

Table 2. Training and testing time in second.

Time	Training	testing
SVD	24.284	28.334
DCT-LBP	28.863	29.645
DCT-PCA	28.863	29.645
DCT-LBP with SVD	18.762	19.398
DCT-PCA with SVD	18.789	20.354

5. Conclusion

Since fusion of information is a relatively new research area. In this paper, we have represented a new rapid method which is the combination of DCT and PCA.PCA is considered as a very fast algorithm with a more or less high robustness and DCT is used for time reduction of recognized output images. So finally we can conclude that combination of PCA and DCT it will offers higher rates of recognition .This face recognition method verifies improvement in parameters in comparison to the existing method. The main advantage of the DCT transform is that it discards redundant information and we have used the left and right singular vectors of SVD instead of the singular values in order to obtain a better performance. In the future, we test our fusion for more complex image such as vary large size 2D images or 3D images.

References

1. 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; IEEE Transactions
2. G.L. Marcialis and F. Roli. *Fusion of LDA and PCA for Face Recognition*. Proceedings of the Workshop on Machine Vision and Perception, 8 th Workshop of the Italian Association for Artificial Intelligence (AIIA'02).
3. ORL, 1992. *The ORL face database at the AT and T (Olivetti) Research Laboratory*. Available from: < *http : //www.uk.research.att.com/facedatabase.html* >
4. C. Shan, al. *Facial expression recognition based on Local Binary Patterns: A comprehensive study IVC*. vol. 27, no. 6, pp. 803-816, 2009. IEEE Transactions
5. 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
6. D. Huang, al. *A novel geometric facial representation based on multiscale extended local binary patterns in FG 2011*. IEEE, 2011, pp. 1-7D. Huang, al., A novel geometric facial representation based on multiscale extended local binary patterns in FG 2011. IEEE, 2011, pp. 1-7
7. T. Ojala M. Pietikainen and D. Harwood. *A comparative study of texture measures with classification based on feature distributions* . Pattern Recognition, Volume 29, 51-59, 1996. Springer-Verlag.
8. Y. Pang; N. Yu; R. Zhang; J. Rong; Z. Liu. *Fusion of SVD and LDA for face recognition*, Image Processing, 2004. ICIP apos; 04. 2004 International Conference on Volume 2, Issue, 24-27 Oct. 2004 Page(s):1417 - 1420 Vol.2.
9. C. H. Hsu and C. C. Chen. *SVD-Based Projection for Face Recognition*, Proceedings of IEEE International Conference on Electro/Information Technology, 703-707, May 17-20, 2007
10. W. Chen, M. J. Er and S. Wu, *PCA and LDA in DCT domain*; Pattern Recognition Letters, Volume 26, Issue 15, 2005, pp. 2474 – 2482.