Turk J Elec Eng & Comp Sci, Vol.19, No.1, 2011, © TÜBİTAK
doi:10.3906/elk-0906-27

Co-occurrence matrix and its statistical features as a new approach for face recognition

Alaa ELEYAN[1], **Hasan DEMİREL**[2]
[1]*Department of Electrical and Electronic Engineering, European University of Lefke, Gemikonagi, Northern-CYPRUS, Mersin-10-TURKEY*
[2]*Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Famagusta, Northern-CYPRUS, Mersin-10-TURKEY*
e-mail: aeleyan@eul.edu.tr, hasan.demirel@emu.edu.tr

Abstract

In this paper, a new face recognition technique is introduced based on the gray-level co-occurrence matrix (GLCM). GLCM represents the distributions of the intensities and the information about relative positions of neighboring pixels of an image. We proposed two methods to extract feature vectors using GLCM for face classification. The first method extracts the well-known Haralick features from the GLCM, and the second method directly uses GLCM by converting the matrix into a vector that can be used in the classification process. The results demonstrate that the second method, which uses GLCM directly, is superior to the first method that uses the feature vector containing the statistical Haralick features in both nearest neighbor and neural networks classifiers. The proposed GLCM based face recognition system not only outperforms well-known techniques such as principal component analysis and linear discriminant analysis, but also has comparable performance with local binary patterns and Gabor wavelets.

Key Words: *Face recognition, gray-level co-occurrence matrix, Haralick features*

1. Introduction

Face recognition has been attracting the attention of the researchers as one of the most important techniques for human identification. One of the limitations of real-time recognition systems is the computational complexity of existing approaches. In the last couple decades, many systems and algorithms with high recognition rates had been introduced. The general problem of these systems is their computational cost in data pre-preparation and transformation to other spaces such as eigenspace [1, 2], fisherspace [3, 4], wavelet transform [5, 6] and cosine transform [7].

Many researchers have used the gray-level co-occurrence matrix [8] for the extraction of features to be used in texture classification. Gelzinis et al. [9] presented a new approach to exploiting information available in the co-occurrence matrices computed for different distance parameter values. In [10], an extension is introduced,

where a new matrix called motif co-occurrence matrix was proposed. Walker et al. [11] have proposed to form co-occurrence matrix-based features by weighted summation of co-occurrence matrix elements from localized areas of high discrimination. Color co-occurrence histograms were used in [12] for object recognition, where they used false alarm rate to adjust some parameters of the algorithm for optimizing the performance.

The idea behind the present paper is to use the co-occurrence matrix and its extracted features in face recognition. To the best of our knowledge, no one has attempted to implement this method before. The idea is simple and straight forward. For each face image, a feature vector is formed by converting the generated gray-level co-occurrence matrix (GLCM) to a vector and then it is used for classification. Additionally, Haralick features [8] containing 14 statistical features can be extracted from the GLCM to form a new feature vector with 14 features.

The proposed GLCM method is compared with four well-known face recognition techniques: Principal component analysis (PCA) [13, 14] which is standard technique used in statistical pattern recognition and signal processing for dimensionality reduction and feature extraction. The proposed method is also compared with the other well-known technique, linear discriminant analysis (LDA), which overcomes the limitations of PCA method by applying the Fisher's linear discriminant criterion. This criterion tries to maximize the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples. Fisher discriminant analysis method groups images of the same class and separates images of different classes. The third technique is local binary patterns [15], which is a non-parametric operator which describes the local spatial structure of an image and high discriminative power for classification. Finally, GLCM was also compared with Gabor wavelets transform. The transform involves convolving an image with a group of Gabor filters with parameterized scale and directionally.

Remainder of this paper is organized as follows. Section 2 discusses gray level co-occurrence matrix. Section 3 explains 14 well-known GLCM Haralick features. The databases are given in section four, while the experimental results, comparison with other algorithms and discussions are given in Section 5, followed by the conclusions in the final section.

2. Gray-level co-occurrence matrix

One of the simplest approaches for describing texture is to use statistical moments of the intensity histogram of an image or region [16]. Using only histograms in calculation will result in measures of texture that carry only information about distribution of intensities, but not about the relative position of pixels with respect to each other in that texture. Using a statistical approach such as co-occurrence matrix will help to provide valuable information about the relative position of the neighbouring pixels in an image.

Given an image I, of size $N \times N$, the co-occurrence, matrix P can be defined as

$$P(i,j) = \sum_{x=1}^{N} \sum_{y=1}^{N} \begin{cases} 1, & \text{if } I(x,y) = i \text{ and } I(x+\Delta_x, y+\Delta_y) = j \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Here, the offset $(\Delta_x\,,\ \Delta_y)$, is specifying the distance between the pixel-of-interest and its neighbour. Note that the offset $(\Delta_x\,,\ \Delta_y)$ parameterization makes the co-occurrence matrix sensitive to rotation. Choosing an offset vector, such that the rotation of the image is not equal to 180 degrees, will result in a different co-occurrence matrix for the same (rotated) image. This can be avoided by forming the co-occurrence matrix using a set

of offsets sweeping through 180 degrees at the same distance parameter Δ to achieve a degree of rotational invariance (i.e., [0 Δ] for 0°: **P** horizontal, [-Δ, Δ] for 45°: **P** right diagonal, [-Δ 0] for 90°: **P** vertical, and [-Δ $-\Delta$] for 135°: **P** left diagonal).

Figure 1 illustrates the details of the process to generate the four co-occurrence matrices using N_g=5 levels for the offsets {[0 1], [-1 1], [-1 0], [-1 -1]} that are defined as one neighboring pixel in the possible four directions. We can see that two neighboring pixels (2, 1) of the input image is reflected in $\mathbf{P}_H$ concurrence matrix as 3, because there are 3 occurrences of the pixel intensity value 2 and pixel intensity value 1 adjacent to each other in the input image. The neighboring pixels (1, 2) will occur again 3 times in $\mathbf{P}_H$, which makes these matrices symmetric. In the same manner, the other three matrices $\mathbf{P}_V$, $\mathbf{P}_{LD}$, $\mathbf{P}_{RD}$ are calculated.

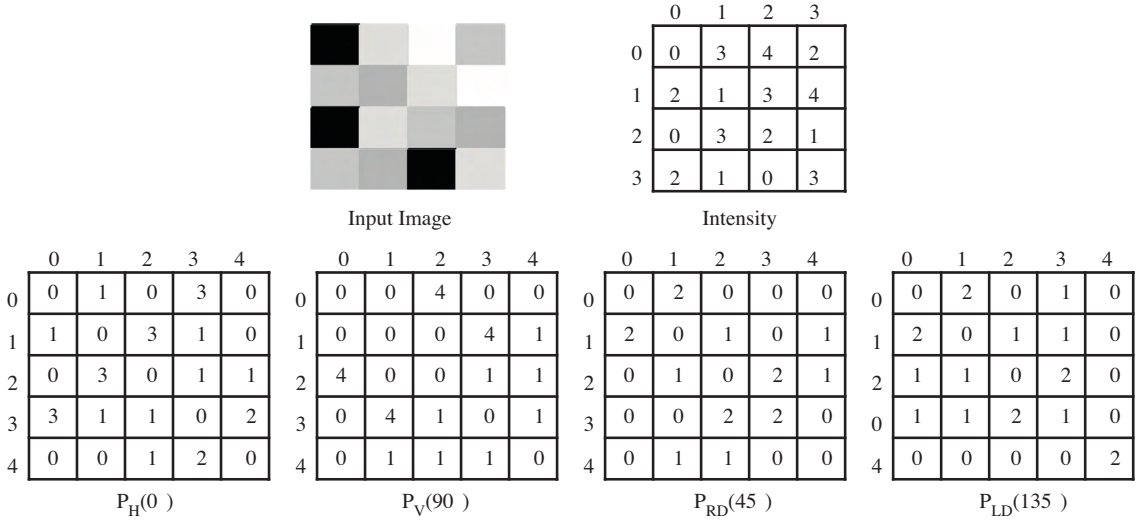
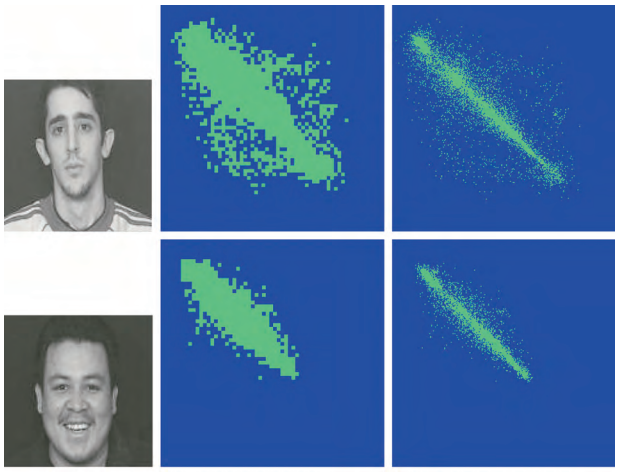


Figure 1. Co-occurrence matrix generation for N_g=5 levels and four different offsets: P_H (0°), P_V (90°), P_{RD} (45°), and P_{LD} (135°).

The four matrices are used separately for classification, then the final decision is formed by fusing the four decisions. As these matrices are symmetric, it is more convenient to use the upper or lower diagonal matrix coefficients in forming the vectors. So, instead of having a vector length of $N_g \times N_g$, the vector size is reduced to $(N_g \times N_g + N_g)/2$ which helps to speed up the process without affecting the recognition performance.

3. Haralick features

In 1973, Haralick [8] introduced 14 statistical features. These features are generated by calculating the features for each one of the co-occurrence matrices obtained by using the directions 0°, 45°, 90°, and 135°, then averaging these four values. The symbol Δ, representing the distance parameter, can be selected as one or higher. In general, Δ value is set to 1 as the distance parameter. A vector of these 14 statistical features is used for characterizing the co-occurrence matrix contents. These features can be calculated by using the following equations.



(a) (b) (c)

Figure 2. (a) Example from FRAV2D Database (b) corresponding CLGM at (64×64) and (c) corresponding CLGM at (256×256).

3.1. Notation

$p(i,\ j)$ is the $(i,\ j)$-th entry in normalized co-occurrence matrix; N_g denotes the dimension of co-occurrence matrix (number of gray levels), and $p_x(i)$ and $p_y(j)$ are the marginal probabilities:

$$p_x(i) = \sum_{j=1}^{N_g} p(i,j), \quad p_y(j) = \sum_{i=1}^{N_g} p(i,j).$$

μ is the mean of μ_x and μ_y;

$$\mu_x = \sum_{i=1}^{N_g} i p_x(i), \quad \mu_y = \sum_{i=1}^{N_g} i p_y(i);$$

and σ_x and σ_y are the standard deviations of p_x and p_y, respectively. All together,

$$\sigma_x = \left(\sum_{i=1}^{N_g} p_x(i)(i-\mu_x)^2 \right)^{1/2}, \quad \sigma_y = \left(\sum_{i=1}^{N_g} p_y(i)(i-\mu_y)^2 \right)^{1/2}$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ i+j=k}}^{N_g} p(i,j); \quad k = 2, 3, \ldots, 2N$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=k}}^{N_g} p(i,j); \quad k = 0, 1, \ldots, N-1$$

$$HXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log\{p_x(i) p_y(j)\}$$

$$HXY2 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_x(i)p_y(j)\log\{p_x(i)p_y(j)\}$$

$$Q(i,j) = \sum_{k=1}^{N_g} \frac{p(i,k)p(j,k)}{p_x(i)p_y(j)},$$

where HX and HY are the entropies of p_x and p_y, respectively:

$$HX = -\sum_{i=1}^{N_g} p_x(i)\log\{p_x(i)\}, \quad HY = -\sum_{j=1}^{N_g} p_y(j)\log\{p_y(j)\}.$$

3.2. Haralick Features

Haralick Features are defined by the following fourteen functions:

$$f_1 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)^2. \tag{2}$$

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j) \right\}. \tag{3}$$

$$f_3 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{(i-\mu_x)(i-\mu_y)p(i,j)}{\sigma_x\sigma_y}. \tag{4}$$

$$f_4 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} (i-\mu)^2 p(i,j) \tag{5}$$

$$f_5 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{1}{1+(i-j)^2} p(i,j) \tag{6}$$

$$f_6 = \sum_{k=2}^{2N_g} k p_{x+y}(k) \tag{7}$$

$$f_7 = \sum_{k=2}^{2N_g} (k-f_6)^2 p_{x+y}(k) \tag{8}$$

$$f_8 = -\sum_{k=2}^{2N_g} p_{x+y}(k)\log\{p_{x+y}(k)\} \tag{9}$$

$$f_9 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log\{p(i,j)\} \tag{10}$$

$$f_{10} = \sum_{k=0}^{N_g-1} [k - \sum_{l=0}^{N_g-1} l p_{x-y}(l)]^2 p_{x-y}(k) \tag{11}$$

$$f_{11} = -\sum_{k=0}^{N_g-1} p_{x-y}(k) \log\{p_{x-y}(k)\} \tag{12}$$

$$f_{12} = \frac{f_9 - HXY1}{\max(HX, HY)} \tag{13}$$

$$f_{13} = (1 - \exp[-2(HXY2 - f_9)])^{1/2} \tag{14}$$

$$f_{14} = (\text{second largest eignvalue of } Q)^{1/2}. \tag{15}$$

4. Proposed GLCM technique

We propose two scenarios for the implementation of GLCM based face recognition. In the first scenario, we used the co-occurrence matrix directly by converting it into a column vector for each image in the recognition process. In the second scenario, we used the vector of Haralick features extracted from the co-occurrence matrix for recognition. Figure 3 shows these scenarios.

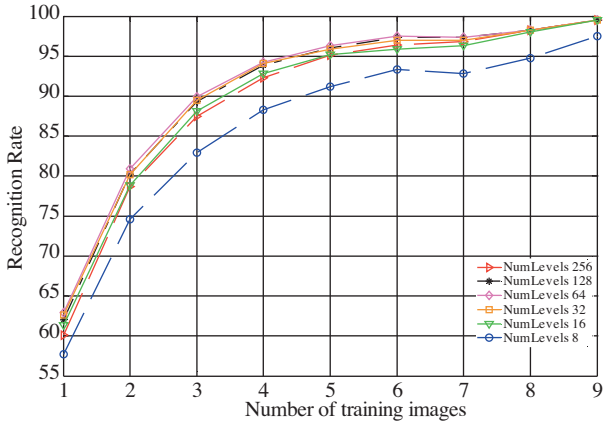


Figure 3. Proposed scenarios for using co-occurrence matrix and its features for face recognition.

In the first scenario, we use the GLCM directly after converting it to a column vector. However, for 8-bit gray level representation the size of this vector is 256×256, which corresponds to a 65536 dimensional vector. The dimension of this vector can be reduced by reducing the number of quantization/gray levels N_g of the image. For example, if N_g is reduced to 16, the GLCM will be (16×16) 256 dimensional vector. The effects of using different number of gray levels (bins) on the recognition performance are studied in the following section.

In the second scenario, a vector of 14 Haralick features is used to represent the given image. This size of this vector is independent of the size of the GLCM.

Figure 4 shows the application of data fusion after the classification of the resulting 4 matrices separately. The decisions coming from the 4 matrices were fused using majority voting. Decision fusion was also applied for the Haralick features resulting from each matrix. As the system has even number of decisions to fuse, ties were break by selecting the first choice.

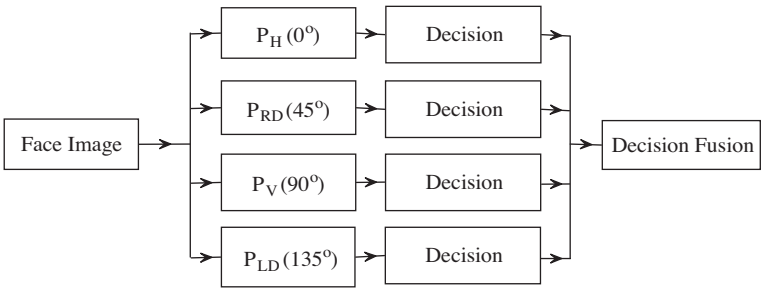


Figure 4. Decision Fusion for the four GLCM matrices (P_H , P_V , P_{RD} , and P_{LD}).

5. Simulation results and discussions

The simulations have been conducted on the four data sets: FERET, FRAV2D, Yale B and ORL face databases. All experiments (except for FERET) were run randomly 10 times, after which results were averaged. The standard test bed adopted in similar studies for the FERET database [17] has been used to test our algorithm for identification. The results are reported in terms of the standard recognition rate. FERET database scenario includes 200 people with 3 upright frontal face images each. In order to test the algorithms, two images of each subject are randomly chosen for training, while the remaining one is used for testing.

ORL face database [18] is the second database we experimented with. It consists of 40 people with 10 images acquired for each one with different facial expressions and illumination variations.

FRAV2D Database [19] is the third database that is used, which comprises 109 people, each with 32 images. It includes frontal images with different head orientations, facial expressions and occlusions. In our experiments, we used a subset of the database consisting 60 people with 32 images each with no overlap between the training and test face images.

The fourth database used in our experiments was the Yale B Database [20], which comprises 10 people, each with 5760 images (9 poses $\times$ 64 illuminations). For conducted simulations, a subset from the database was used with 90 images (3 pose $\times$ 30 illuminations) for each person.

Table 1 shows the results obtained for 4 matrices at different angles (i.e. 0°, 45°, 90°, and 135°). The results are for the FERET database with gray level N_g=32. The recognition rates for the different matrices ranged from 90.5% to 92.33%. Fusing the four decisions helped to improve the recognition rate to 93.08%. The rest of the results in the following tables show the performance of the fused decisions directly.

Table 1. Face recognition performance (%) for FERET database with 4 different angles of Direct GLCM with N_g=32.

N_g	Angle	Performance (%)
32	0^o	92.17
	45^o	91.87
	90^o	92.33
	135^o	90.51

Table 2 shows the result of using GLCM column vector directly (direct GLCM) and the Haralick features on four different databases; ORL, FERET, FRAV2D, and Yale B databases using nearest neighbor classifier namely L_1 norm distance, δ_{L1}. Reduced number of gray levels has been used for GLCM which in turns

reduced the resulting vector length. For ORL database, the training set has been formed by using 5 samples of each individual, where the test set contains the remaining 5 samples. The results clearly show that the direct GLCM approach outperforms the Haralick feature based approach. It is also, important to note that, the dimensionality reduction based on the reduction of the number of gray levels does not significantly change the recognition performance as we reduce the number of gray levels from 256 down to 16 levels. The highest performance is achieved by using 64 gray levels.

Table 2. Face fecognition performance (%) with different number of levels for ORL, FERET, FRAV2D and Yale B databases using Haralick features and Direct GLCM.

		Number of Gray Levels						
Approach		256	128	64	32	16	8	4
Haralick features	ORL	72.70	71.00	71.30	75.70	80.30	79.10	70.40
	FERET	30.23	30.57	27.07	27.57	35.40	37.73	14.06
	FRAV2D	50.56	51.99	56.36	58.87	60.77	55.10	47.08
	Yale B	26.18	26.42	26.82	26.25	27.22	21.48	15.81
Direct GLCM	ORL	95.50	96.40	96.70	96.30	95.60	91.60	70.80
	FERET	89.42	91.42	92.42	93.08	89.42	75.58	15.08
	FRAV2D	94.31	94.23	94.37	96.02	87.55	80.24	42.56
	Yale B	78.83	79.67	79.33	81.67	76.17	59.80	18.43

Table 2 also shows the results from the FERET database. The training set is formed by using 2 samples of each individual, where the test set contains the remaining 1 sample per subject. Similarly, the results clearly show that the direct GLCM approach outperforms the Haralick features based approach. Similar results are achieved for FRAV2D and Yale B databases. For both FERET and FRAV2D, the highest performance of direct GLCM is achieved by using 32 gray levels.

The results in Table 2 indicate that, the proposed direct GLCM method is superior to the method using well-known Haralick features. Furthermore, it is important to note that changing the number of gray levels N_g provides important flexibility, which results scalable matrix size for the GLCM. An important notice which can be observed in table 2 that in some cases when the co-occurrence matrix become smaller due to the use of smaller number of gray levels, the recognition performance of direct GLCM approach slightly increases. This can be attributed to the quantization process that helps to suppress the noise in the input image, which as a result, helps to increase the recognition performance.

Figure 5 shows the performance of the direct GLCM method for changing number of training samples in the training set for different number of gray levels. The results show that the recognition performance of the proposed system changes slightly as we reduce the number of gray levels from 256 to 16 gray levels. Although, this observation is specific to ORL dataset and cannot be generalized, the results confirm the scalability of the proposed direct GLCM approach.

Table 3 compares the performance of the proposed GLCM based recognition systems with the state of the art techniques PCA, LDA, Gabor wavelets (GW) and LBP by using the ORL face database and nearest neighbor classifier. The results demonstrate the superiority of the proposed direct GLCM based face recognition system over the well know face recognition techniques PCA and LDA. In same manner, Table 4 shows the performance results of same algorithms with different number of training images using subset of FRAV2D database and nearest neighbor classifier. In both tables, the GLCM approach outperformed the PCA and LDA approaches. While GLCM performed nearly same as LBP and GW in some cases, it is clear that LBP and GW have the

overall highest similar performance among them. LBP in both tables used 256 bins and 16 regions. For Gabor wavelets we used the same scenario as in [21] with 5 scales and 8 directions.

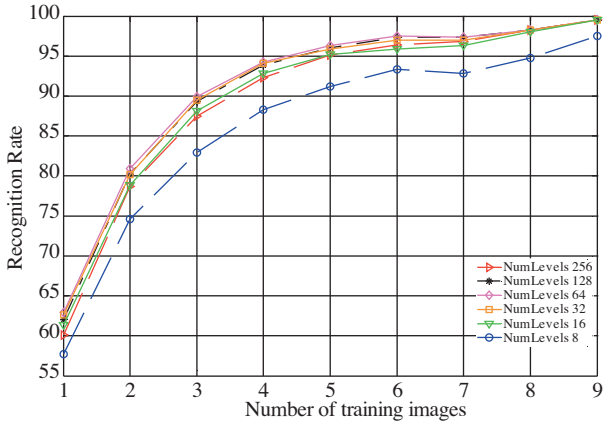


Figure 5. Face recognition performance of the Direct GLCM for ORL database, using different number of training images and different number of grey levels using L_1 norm.

Table 3. Performance of PCA, LDA, GW, LBP, Direct GLCM and Haralick features using ORL Database, Distance Measure δ_{L1} and Gray Level N_g=256.

# Train	Haralick Features	Direct GLCM	PCA	LDA	GW	LBP
1	42.76	60.56	61.44	-	67.47	70.11
2	56.95	79.03	74.63	76.33	78.50	80.33
3	63.20	87.80	82.86	86.67	87.56	89.10
4	68.76	92.76	88.42	92.86	93.10	93.43
5	72.34	95.50	90.00	95.47	95.33	95.17
6	74.95	96.78	91.13	96.67	97.88	97.33
7	75.17	97.25	94.00	97.10	98.15	98.67
8	76.44	98.67	94.50	97.33	99.78	100
9	76.44	99.60	97.50	98.56	100	100

Table 4. Performance of PCA, LDA, GW, LBP, Direct GLCM and Haralick Features using FRAV2D Database, Distance Measure δ_{L1} and Gray Level N_g=256.

# Train Images	Haralick Features	Direct GLCM	PCA	LDA	GW	LBP
1	38.13	55.91	57.24	-	63.55	65.91
2	53.40	75.59	71.53	73.35	77.67	77.23
3	59.96	84.53	79.96	83.56	85.20	86.20
4	65.53	89.53	85.62	90.10	90.33	90.63
5	69.40	93.46	87.50	93.33	93.56	93.50
6	72.30	94.18	88.93	94.56	95.33	95.13
7	72.62	94.78	91.95	94.88	96.76	96.62
8	74.34	96.59	92.84	96.10	98.47	98.34
9	75.50	98.00	95.00	97.67	98.67	98.50

In Table 5, instead of nearest neighbor classifier we used neural networks classifier. The feedforward neural network was applied with back-propagation learning algorithm. The neural networks inputs depends on the length of the vector which determined by the used algorithm. The output layer contains 40 neurons corresponding to the number of people in the ORL database. For training, hyperbolic tangent sigmoid was used as the activation function with MSE 10^{-3} and 1000 iterations. The results in table 5 show that GLCM approach is outperforming Haralick features approach and has competitive results with LBP approach.

Table 5. Performance of PCA, LDA, LBP, Direct GLCM and Haralick Features using ORL Database, NN classifier and Gray Level N_g=256.

# Train Images	Haralick Features	Direct GLCM	LBP
1	44.50	62.10	71.56
2	57.25	81.25	82.47
3	64.25	88.80	89.67
4	69.33	93.76	94.50
5	72.88	96.50	96.77
6	75.50	97.78	98.10
7	76.10	98.47	99.50
8	77.47	99.50	100
9	78.33	100	100

6. Conclusions

In this paper, we propose a new method, direct GLCM, which performs face recognition by using Gray-Level Co-occurrence Matrix. The direct GLCM method is very competitive with state of the art face recognition techniques such as PCA, LDA, Gabor wavelets, and LBP. Using smaller number of gray levels (bins) shrinks the size of GLCM which reduces the computational cost of the algorithm and at the same time preserves the high recognition rates. This can be due to the process of quantization which helps in suppressing the noise of the images at higher grey levels. It is obvious from the results that the GLCM is a robust method for face recognition with competitive performance.

References

[1] M. Turk, A. Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, vol. 3, pp. 71–86. 0898-929X, 1991.

[2] A. Pentland, B. Moghaddam, T. Starner, View Based and Modular Eigenspaces for Face Recognition, In *Proceedings of Computer Vision and Pattern Recognition*, pp. 84–91, 0-8186-5825-8, *IEEE Computer Society*, Seattle, USA, June 1994.

[3] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 771–720, 0162-8828, July 1997.

[4] W. Zhao, R. Chellappa, N. Nandhakumarm, Empirical Performance Analysis of Linear Discriminant Classifiers. In *Proceedings of Computer Vision and Pattern Recognition,* pp. 164–169, 0-8186-5825-8, *IEEE Computer Society*, Santa Barbara, Canada, June 1998.

[5] L. Wiskott, J.-M. Fellous, N. Kruger, C. von der Malsburg, Face Recognition By Elastic Bunch Graph Matching, *IEEE Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, July 1997.

[6] C. Liu, H. Wechsler, Independent Component Analysis of Gabor Features for Face Recognition, *IEEE Transactions on Neural Networks*, vol. 14, no. 4, pp. 919–928, July 2003.

[7] Manjunath, B., Chellappa, R., von der Malsburg, C. A Feature Based Approach to Face Recognition. In: Proc. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 373–378, 1992.

[8] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 3, no. 6, pp. 610–621, 1973.

[9] A. Gelzinisa, A. Verikasa, M. Bacauskienea, Increasing the discrimination power of the co-occurrence matrix-based features, *Pattern Recognition,* vol. 40, pp. 2367–2372, 2004.

[10] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique, Content Based Image Retrieval Using Motif Co-Occurrence Matrix, *Image Vision Computing*, vol. 22, pp. 1211–1220, 2004

[11] R.F. Walker, P.T. Jackway, D. Longstaff, Genetic Algorithm Optimization of Adaptive Multi-Scale GLCM Features, *International Journal on Pattern Recognition and Artificial Intelligence,* vol. 17, no. 1, pp.17–39, 2003.

[12] P. Chang, J. Krumm. Object Recognition with Color Cooccurrence Histograms. *IEEE Conference on Computer Vision and Pattern Recognition*, Fort Collins, co, June 1999.

[13] M. Kirby, L. Sirovich, Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces, *IEEE Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 0162-8828, January 1990.

[14] L. Sirovich, M. Kirby, Low-Dimensional Procedure for the Characterization of Human Faces, *Journal of the Optical Society of America*, vol. 4, no. 3, pp. 519–524, 1084-7529, March 1987.

[15] T. Ahonen, A. Hadid, and M. Pietikainen. Face Recognition with Local Binary Patterns. *In European Conference on Computer Vision* (ECCV), 2004.

[16] R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, 3^{rd} Ed. Prentice Hall, 2008.

[17] P. J. Philipps, H. Moon, S. Rivzi, P. Ross, The Feret Evaluation Methodology Fro Face-Recognition Algorithms, *IEEE Pattern Analysis and Machine Intelligence,* vol 22, pp. 1090–1104, 2000.

[18] The Olivetti Database; http://www.cam-orl.co.uk/facedatabase.html.

[19] Á. Serrano, I. M. de Diego, C. Conde, E. Cabello, L Shen, L. Bai, Influence of Wavelet Frequency and Orientation in an SVM Based Parallel Gabor PCA Face Verification System, H. Yin et al. (Eds.): IDEAL Conference 2007, Springer-Verlag LNCS 4881, pp. 219–228, 2007.

[20] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From Few to Many: illumination Cone Models for Face Recognition under Variable Lighting and Pose, *IEEE Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[21] C. Liu and H. Wechsler, "Independent Component Analysis of Gabor Features for Face Recognition face", *IEEE Transactions on Neural Networks*, vol. 14, no. 4, pp. 919–928, July 2003.