A Novel Ear Recognition Method Using Feature Combination

Prof.Atalla I. Hashad

Dept. of Computer Engineering
AAST, Cairo
hashad@cairo.aast.edu

Dr.Gouda I. Salama

Dr_gouda80@yahoo.com

Alaa Tharwat

Dept. of Computer Science
HICIT, Cairo
engalaatharwat@hotmail.com

Abstract*: In this paper we aim to improve the accuracy of the recognition rate in the ear recognition by developing theoretical framework for combining (serial or parallel) statistical feature extraction methods (PCA, LDA, and DCT). Experimental comparisons of the combining methods demonstrate that the combination methods outperform other single feature extraction methods.*

Key words: Principle component analysis (PCA), linear discriminant analysis (LDA), discrete cosine transform (DCT), Combining features.

1. INTRODUCTION

Using ears to identify people has been interesting in the recent years. There are many methods proposed for ear recognition. The most prominent work was done by the French criminologist Alphonse Bertillon (1890) [1]. In machine vision, ear biometrics has received scant attention compared to the more popular techniques of automatic face, eye, or fingerprint recognition. However, ears have played a significant role in forensic science for many years, especially in the United States, where an ear classification system based on manual measurements has been developed by Iannarelli, and has been in use for more than 40 years [2]. He examined over 10,000 ears and found them all to be distinguishable. Iannarelli developed an anthropometric method where 12 measurements are used as features to distinguish individuals. Chang et al. [3] compared ear recognition with face recognition using a standard principal components analysis (PCA) technique known as the "eigen-face" approach on face and "eigen-ear" images. They reported that 71.6% and 70.5% are the accuracies obtained in the baseline experiment for ear and face recognition, respectively. Chang also presented results for a lighting variation experiment where the reported rates were 64.9% for face and 68.5% for ear. In a multimodal experiment for combining ear and face images, they reported

recognition results of a 90.9% rank one recognition, which is a significant gain in performance over their baseline experiment. Hurley et al. [4] considered a "force field" feature extraction approach that is based on simulated potential energy fields. They reported recognition results 99% rank one recognition rate using template matching. Then, Choras [5] presented a study on ear biometrics using geometric feature extraction. With this ideal environment for recognition, an error free recognition was reported. In other study, Moreno et al. [6] performed two experiments with neural network classifiers. For the classification technique based on feature points a 43% recognition rate was reported, and for the ear morphology approach with a neural network classifier a recognition rate of 83% was reported. Other research has been done at the University of Notre Dame by Ping Yan et al. [7], where they reported 63.8% rank one recognition rate for an "eigen-ear", PCA-based approach. The same approach was performed on range images achieving 55.3% rank one recognition rate. M. Saleh et al. [8] performed experiments using different feature extraction methods and different classification techniques. Four classifiers utilized feed-forward artificial neural networks (ANN), and three classifiers that were based on a nearest neighbor rule. The seven experiments performed in this study had rank one recognition rates ranging from 76.5% to 94.1%. Yuizono et al. treated the ear image recognition problem as an optimization problem and applied GA (genetic algorithm) to it [9]. Results showed that the highest registrant recognition rate achieved was approximately 100%, and 100% for the un-registrant rejection rate. Applied different statistical methods such as (Correlation, PCA, LDA, DCT, and ICA) to extract ear features will get a good results. And the results showed that LDA technique is the superior than other algorithms in all cases [10].

To improve the performance and robustness achievable by individual recognizers, the use of multiple classifier systems (MCSs) has been recently proposed. MCSs are currently a very active research field. Multiple classifiers systems cover a wide spectrum of applications. The effectiveness of this approach is documented by many experimental results [11-14]. Approaches for improving the performance and the robustness of a single face recognition system based on MCSs have also been proposed: The outcome of each expert, represented by a score, i.e. a level of confidence about the decision, is combined with simple fusion rules (majority voting, rank summation, Bayes's combination rule); Lucas [15] uses a n-tuple classifier for combining the decisions of experts based on sub-sampled images; Tolba [16] presents a simple combination rule for fusing the decisions of a RBF network and the decisions of a LVQ network. The works by Lucas and Tolba can be considered as the state of the art about multiple classifiers applied to face recognition. Marcialis and Roli [17] exploited the fusion of PCA and LDA for face verification KNN-based combination rule and the NM-based combination rule. In general, the performance of the KNN rule is much better than that of the NM rule: this should mean that the average template (that can be viewed

as a low-pass filtering in the domain of the PCA and LDA spaces) reduces the available information.

2. METHODOLOGY

2.1. SINGLE FEATURE EXTRACTION METHOD:

There are many different statistical approaches to extract features from ear images as described in.

- **Principal Component Analysis (PCA)**

Principal Component Analysis (PCA) [18, 19]. PCA is a linear subspace projection technique. It finds a set of representative projection vectors such that the projected samples have most information about original samples. The most representative vectors are the eigenvectors corresponding to the largest "eigen-values" of the covariance matrix. And reducing the dimensionality of the description by projecting the points onto the principal axes. In mathematical terms, we wish to find principal components of the distribution of ear images (eigenvectors of the covariance matrix of the set of ear images). Each ear image in the training set can be represented in terms of a linear combination of the "eigen-ears". The number of the "eigen-ears" is equal to the number of the ear images in the training set. The PCA algorithm is shown in list 1.

Step 1: Obtain ear images $I_1, I_2, I_3, I_4 \ldots \ldots I_M$ (Training ear images).Each image is NxN

Step 2: Represent every image as a vector Γ_i (N^2x1) .So, Γ is an N^2xM vector.

Step 3: Compute the average ψ (N^2x1)

$$\psi \;=\; \frac{1}{M} \sum_{i=1}^{M} \Gamma_i \qquad \text{(Eq. 1)}$$

Step 4: Subtract the mean (center data) :

$$\phi_i \;=\; \Gamma \;-\; \psi \qquad \text{(Eq. 2)}$$

Step 5: Form the data matrix A=[$\Phi_1\ \Phi_2\ \Phi_3 \ldots \ldots \Phi_M$] ($N^2$ x M matrix)

Step 6: Compute the covariance matrix (N^2 x N^2):

$$C \;=\; \frac{1}{M} \sum_{n=1}^{M} \phi_n \phi_n^{T} \;=\; \frac{1}{M} AA^{T} \qquad \text{(Eq. 3)}$$

Step 7: Compute eigenvalues ($\lambda_1\ \lambda_2\ \lambda_3 \ldots . \lambda_N{}^2$) and eigenvectors ($u_1\ u_2\ u_3 \ldots \ldots u_N{}^2$) of C (each $\quad N^2$xN^2). And sort eigenvectors according to

---

eigenvalues. And neglect the eigenvectors when its associate eigenvalue less then 0.01 (for example the number of eigenvalues that have values greater than 0.01 is P)

Step 8: Compute the feature matrix (PxM):

$$\Omega = U^T A \qquad \text{(Eq. 4)}$$

To identify an unknown ear image Γ (centered and in the same size) we will apply the following steps.

Step 9: normalize the testing image Γ (N^2x1):

$$\phi_i = \Gamma - \psi \qquad \text{(Eq. 5)}$$

Step 10: Project on the eigenspace $\hat{\Omega}$ (Px1) :

$$\hat{\Omega} = U^T \phi_i \qquad \text{(Eq. 6)}$$

Step 11: Compute the distance between the test image and the training image and determine the class of the unknown image

$$E_r = \min \|\Omega - \hat{\Omega}\|. \qquad \text{(Eq. 7)}$$

List 1: The PCA algorithm

▪ **Linear Discriminant analysis (LDA)**

PCA do not use ear class (category) information. The training data is taken as a whole. LDA finds an efficient way to represent the ear vector space by exploiting the class information [20], [21]. It differentiates individual ears but recognizes ears of the same individual. LDA is often referred to as a Fisher's Linear Discriminant (FLD). The images in the training set are divided into the corresponding classes. The LDA algorithm is shown in list 2.

Step 1: Obtain ear images $I_1, I_2, I_3, \ldots \ldots I_M$ (Training ear images) . I(NxN).
Step 2: Represent every image as a vector Γ_i N^2x1 .So, Γ is an N^2xM vector.

Step 3: Compute the mean of each class m_i (N^2x1)and the mean of all data m (N^2x1) .

Step 4: Compute class-dependent scatter matrix (N^2x N^2):

$$S_i = \frac{1}{P} \sum_{i=1}^{K} (\Gamma_i - m_i)(\Gamma_i - m_i)^T \qquad \text{(Eq. 8)}$$

Where

P: Number of images of each class

Γ_k: Class data matrix , and K is the number of classes

S_i: Scatter matrix

Step 5: Compute within-class scatter matrix S_w ($(N^2 x N^2)$):

$$S_w = \sum_{i=1}^{K} S_i \qquad \text{(Eq. 9)}$$

Step 6: Compute between-class scatter matrix S_b ($N^2 x N^{2)}$:

$$S_b = \sum_{i=1}^{K} \frac{1}{K}(m_i - m)(m_i - m)^T \qquad \text{(Eq. 10)}$$

We are seeking the matrix W maximizing

$$W = \max \left| \frac{W^T S_b W}{W^T S_w W} \right| = \max \left| \frac{S_b}{S_w} \right| \qquad \text{(Eq. 11)}$$

Step 7: Calculate the eigen values (λ) and eigen vectors (w) to solve Eq. 11. Then sort the eigen vectors according to its eigen values.

$$S_b w = S_w w \lambda \qquad \text{(Eq. 12)}$$

Step 8: Project images onto Fisher basis vectors: Project all the original (i.e. not centered) images onto the Fisher basis vectors

$$Y = W\Gamma \qquad \text{(Eq. 13)}$$

To identify an unknown (test) ear image Γ_i we will apply the following steps.

Step 9: Project image onto Fisher basis vectors.

$$r = W \Gamma_i \qquad \text{(Eq. 14)}$$

Step 10: Compute the distance between the test image and the training images and determine the class of the unknown image:

$$E_r = \min \|Y\text{-}r\| \qquad \text{(Eq. 15)}$$

List 2 :The LDA algorithm

▪ **Discrete Cosine Transform (DCT)**

DCT is a well-known signal analysis transform used in compression standards due to its compact representation power. A detected and normalized ear image is divided into blocks of 8x8 pixels size. Each block is then represented by its DCT coefficients. DCT coefficients the ones containing the highest information are extracted via zigzag scan. To fuse the local information, the extracted features from 8x8 pixels blocks can be combined at the feature level or at the decision level [22]. The DCT algorithm is shown in list 3.

Step 1: Obtain ear images $I_1, I_2, I_3, \ldots \ldots I_M$ (Training ear images). Each image is NxN.

Step 2: A detected and normalized ear image is divided into blocks of 8x8 pixels size.

Step 3: Each block is then represented by its DCT coefficients.

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\Pi(2m+1)p}{2M} \cos \frac{\Pi(2n+1)q}{2N} \quad \begin{array}{l} 0 \leq p \leq M-1 \\ 0 \leq q \leq N-1 \end{array} \text{ (Eq. 16)}$$

Where

A: Input image.

M: Column size of A.

N: Row size of A.

$$\alpha p = \begin{cases} 1/\sqrt{M}, & p = 0 \\ \sqrt{2/M}, & 1 \leq p \leq M-1 \end{cases} \qquad \text{(Eq. 17)}$$

$$\alpha q = \begin{cases} 1/\sqrt{N}, & q = 0 \\ \sqrt{2/N}, & 1 \leq q \leq N-1 \end{cases} \qquad \text{(Eq. 18)}$$

Step 4: From the remaining DCT coefficients, the ones containing the highest information are extracted via zigzag scan.

Step 5: The DCT coefficients obtained from each block are concatenated to construct the feature vector which is used by the classifier. These steps (1, 2, 3, and 4) are performed on all the training images.

To identify unknown ear image (centered and in the same size) we will apply the following steps

Step 6: The previous steps are performed on the unknown (test) image to construct the feature vector.

Step 7: Compute the minimum distance from the feature vectors of the training images to identify the class of the unknown image.

List 3: The DCT algorithm

2.2. COMBINED FEATURE EXTRACTION METHODS

Combining feature extraction methods used to improve the accuracy of the recognition rate of the ear images. The combination can be implemented in many levels (similarity measurements, ranking scores, and final decisions). In this proposed technique we use combination in the decision level. There are two methods of combinations (serial and parallel).

Correlation coefficient (ρ) should be calculated when choosing feature extraction methods to prevent using two dependent feature extraction methods. When ρ tends to zero so, the feature extraction methods that had been chosen are independent and the combination between them will give us good results. As shown in table.1 the correlation coefficient between feature extraction methods shows that the combination methods illustrated into the table will give us a good results because the correlation coefficient tends to zero.

Combined methods	Correlation Coefficient
PCA & LDA	-0.000796
PCA & DCT	0.000020816
DCT & LDA	-0.0011

Table.1 Correlation coefficient between feature extraction methods

▪ **Parallel Combination:**

Fig.1 shows the steps of combining two methods of feature extraction. The figure is divided into three stages. In the first step (training stage) we collect ear images and perform images preprocessing to all images and then apply two different feature extraction methods, and enroll the result (represented as vectors) into two databases. In the second stage (test stage) we apply the same preprocessing steps and then apply the same two feature extraction methods on the samples of ear images. In the third stage (matching and combining step) match every feature extraction method with its database to get the decision, then apply the combination technique to combine the two decisions.

The combination step is started from computing the distance vectors d^{f1} and d^{f2} (f1 and f2 denotes the first and second feature extraction methods) that are computed from the matching of the sample images with the two databases. Then append the final decisions from two methods into one vector d and then apply the KNN on this vector to get final decisions.

$$d = \left\{ d^{f1}_{1}, \ldots\ldots, d^{f1}_{N}, d^{f2}_{1}, \ldots\ldots., d^{f2}_{N} \right\}$$
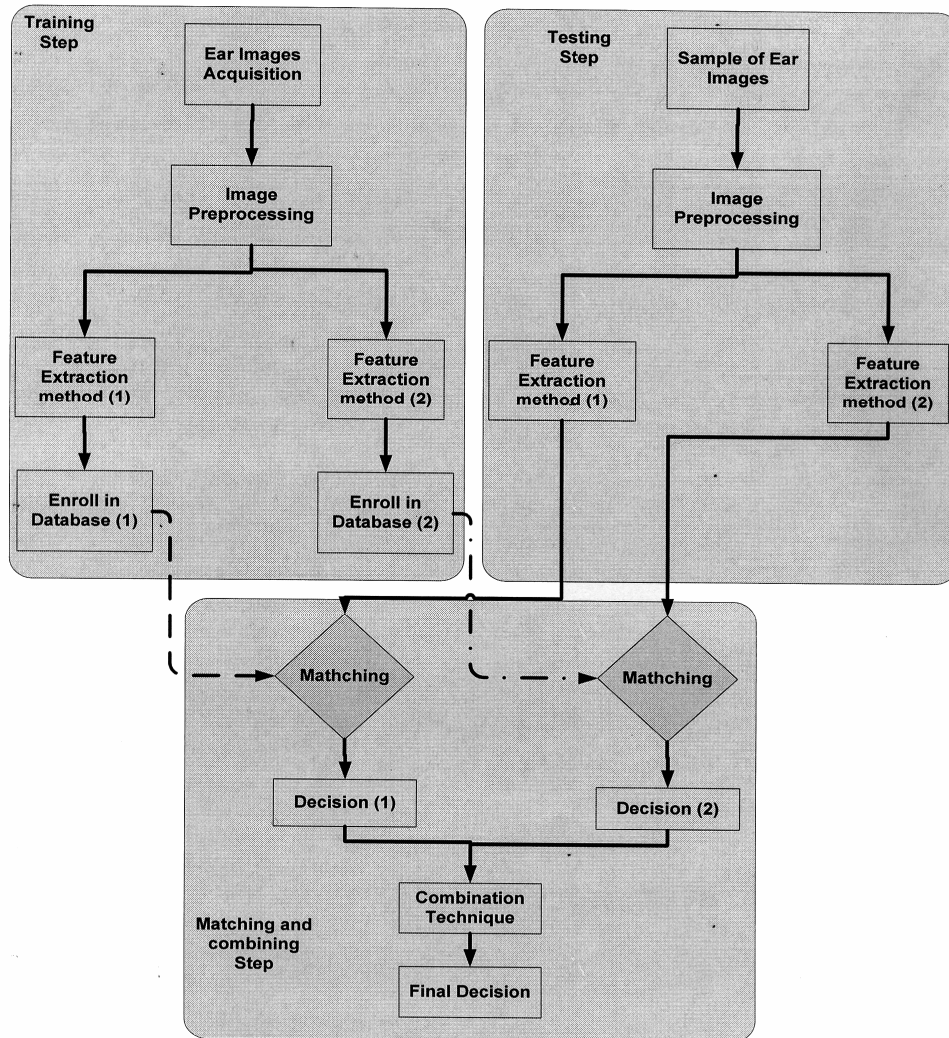
Fig.1 Schematic block diagram of parallel combining feature extraction methods

▪ Serial Combination

Figure2 shows the steps of combining two methods of feature extraction methods. After the images acquisition and preprocessing steps apply the first feature extraction method (LDA, PCA, or DCT) and then use the transformation matrix and apply the second feature extraction method. After applying all the

٨

previous steps on the training and testing images, match the testing images with training images using minimum distance or K-nearest neighbor classifier (for example when use the PCA transformation matrix as in equation (4) to apply DCT feature extraction method on it and use the output of DCT as an transformation matrix (apply this two feature extraction methods in this order onto the training and test images). The two feature extraction methods can be done as in the table.2 order, but we can't reverse the sequence of these methods because the dimensions are not matched. Also not all methods can be used in this combination method due to the dimensions conflict.
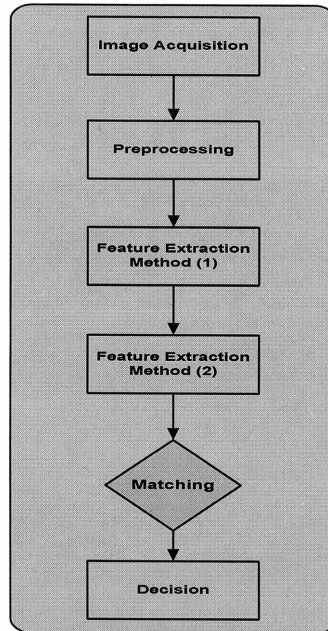
```
┌─────────────────────────┐
│   Image Acquisition     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│      Preprocessing      │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Feature Extraction    │
│       Method (1)        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Feature Extraction    │
│       Method (2)        │
└─────────────────────────┘
            │
            ▼
        ◇ Matching ◇
            │
            ▼
┌─────────────────────────┐
│        Decision         │
└─────────────────────────┘
```

Fig.2 Schematic block diagram of Serial combining feature extraction methods


3. EXPERIMENTAL RESULTS

In this section we will compare between single (PCA, LDA, and DCT) and (parallel and serial) combination (PCA+LDA, PCA+DCT, and LDA+DCT) feature extraction methods. We will use small database (Carreira-Perpinan, 1995) consists of 102 grayscale images (6 images for each of 17 subjects) in PGM format. Six views of the left profile from each subject were taken under uniform, diffuse

lighting. Slight changes in the head position were encouraged from image to image. In our experiment we will use the database two times. First, we will use three images as training images and three for test. Second, one images as a training images and five for test. We have been used minimum distance classifier based on two distance measurements Cosine and Chebychev to measure similarities.

• **Parallel Combination** :

Table 2 and Fig.3 presents a comparison between parallel combination of features and single feature extraction methods based on single classifier. We note that when combining two feature extraction methods, the result will be better than or equal when using single feature extraction methods for one training images as in Figure 4.2. While parallel feature combination gives us results approximately the same or lower than single feature extraction methods for three training images.
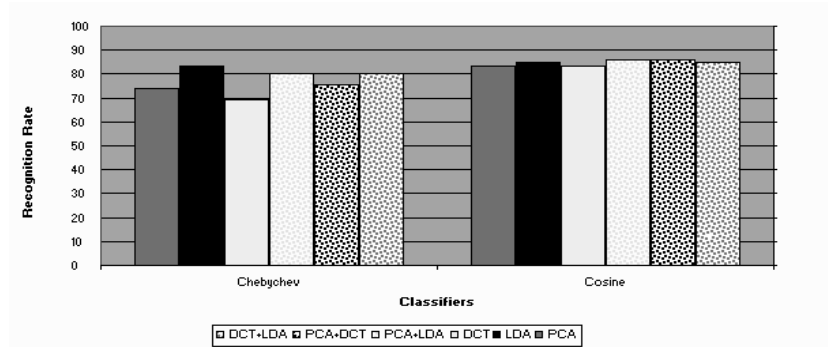
(a)
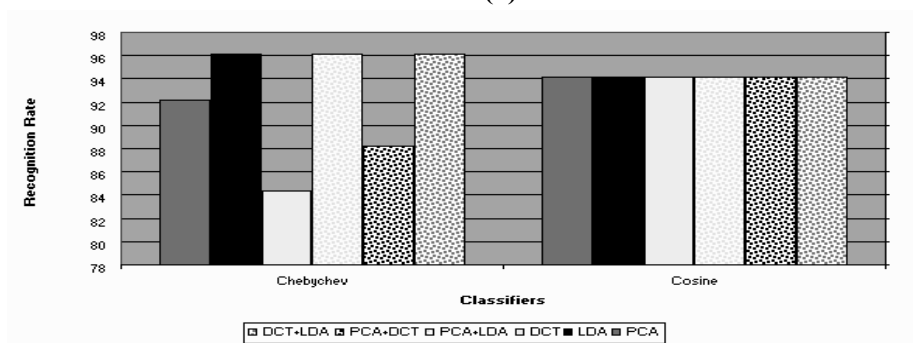
(b)

Figure 3: Comparative performance between single and parallel combination of feature extraction methods (a) For one training image   (b) For three training image

**Table 2: Comparative performance between single and parallel combination of feature extraction methods**

|	PCA		LDA		DCT		PCA+LDA		PCA+DCT		DCT+LDA	
	1	3	1	3	1	3	1	3	1	3	1	3
Cosine	83.53	94.12	84.71	94.12	83.53	94.12	85.9	94.12	85.9	94.12	84.71	94.12
Chebychev	74.12	92.16	83.53	96.08	69.41	84.31	80	96.08	75.3	88.24	80	96.08

- **Serial Combination** :

A comparison between serial combination of feature extraction methods and single feature extraction methods is introduced in table 3 and Figure 4. We note that, serial combination (PCA+LDA, PCA+DCT, and LDA+DCT) of two feature extraction methods will give us good results compared with single (PCA, LDA, and DCT) feature extraction methods for one training images. While when using three training images serial combination is approximately the same as single feature extraction methods.

Table 3: Performance comparison between single and serial combination of feature extraction methods

	PCA		LDA		DCT		PCA+LDA		PCA+DCT		DCT+LDA	
Classifiers	1	3	1	3	1	3	1	3	1	3	1	3
Cosine	83.53	94.12	84.71	94.12	83.53	94.12	92.94	96.08	83.53	94.12	84.71	96.08
Chebychv	74.12	92.16	83.53	96.08	69.41	84.31	83.53	96.08	78.4	92.16	87.06	96.08
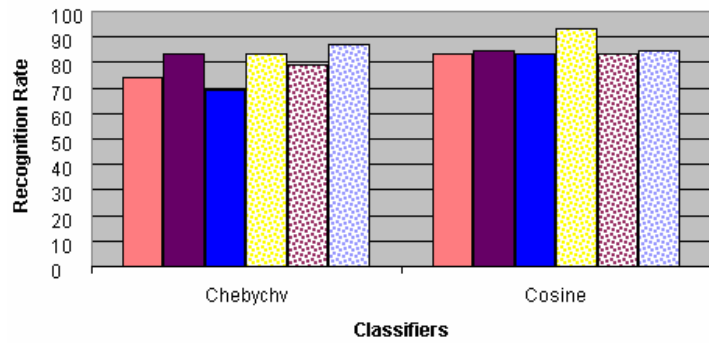
(a)
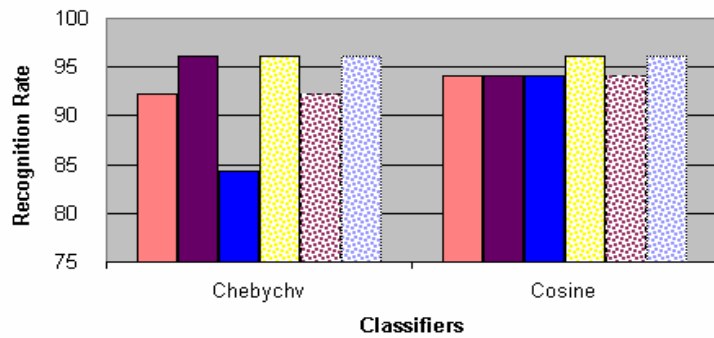
(b)

Figure 4: Performance comparison between single and serial combination of feature extraction methods (a) For one training image (b) For three training images

4. CONCLUSION

Combination of two single extraction methods will increase the recognition rate. And the effect of combining is obvious when using one training image and independent feature extraction methods because the recognition rates in such cases are low and when using the combination technique the recognition rate will increase. Also, we note that parallel and serial combination techniques give us good results. But when the serial combination will give us results better than parallel combination as in table 4.

Classifiers	Serial						Parallel					
	PCA+LDA		PCA+DCT		DCT+LDA		PCA+LDA		PCA+DCT		DCT+LDA	
	1	3	1	3	1	3	1	3	1	3	1	3
Cosine	*92.94*	*96.08*	83.53	94.12	84.71	*96.08*	85.9	94.12	*85.9*	94.12	84.71	94.12
Chebychev	*83.53*	96.08	*78.82*	*92.16*	*87.06*	96.08	80	96.08	75.3	88.24	80	96.08

Table 4: Performance comparison between parallel and serial combination of feature extraction methods

5. References

[1] A. Bertillon, La photographie judiciaire, "avec un appendice sur la classification et l_identification anthropometriques", Gauthier-Villars, Paris, 1890.

[2] A. Iannarelli, "Ear Identification", Paramont Publishing Company, 1989.

[3] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and Combination of Ear and Face Images in Appearance-based Biometrics", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25 pp. 1160-1165, 2003.

[4] D. J. Hurley, M. S. Nixon, and J. N. Carter, "Force Field Feature Extraction for Ear Biometrics. Computer Vision and Image Understanding", vol. 98, pp. 491-512, 2005.

[5] M. Choras, "Ear Biometrics Based on Geometric Feature Extraction". Electronic Letters on Computer Vision and Image Analysis 5(3), 84-95, 2005.

[6] B. Moreno, A. Sanchez, and J. F. Velez, "On the Use of Outer Ear Images for Personal Identification. in Proc", IEEE International Carnahan Conf. on Security Technology, pp. 469-476, 1999.

[7] P. Yan and K. W. Bowyer, "Empirical Evaluation of Advanced Ear Biometrics", Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshop Empirical Evaluation Methods in Computer Vision, pp. 41-48, 2005.

[8] M. Saleh, S. Fadel, and L. Abbott, "Using Ears as a Biometric for Human Recognition", Proc. International Conf. on Computer Theory and Applications, pp. 311-314, 2006.

[9] T. Yuizono, Y. Wang, K. Satoh, and S. Nakayama, "Study on Individual Recognition for Ear Images by Using Genetic Local search", Proceedings of the 2002 Congress on Evolutionary Computation (CEC2002), p. 237-242, 2002.

[10] Atalla H., Gouda I., Alaa T., "Personal Identification Using Ears Based on Statistical Features", AEIC2008, Egypt, 2008.

[11] J. Kittler, and F. Roli Eds, "Proceedings of the First International Workshop on Multiple Classifier Systems", Springer LNCS 1857, Is Molas Golf Hotel Cagliari Italy, 2000.

[12] J. Kittler and F. Roli Eds, "Proceedings of the Second International Workshop on Multiple Classifier Systems", Springer LNCS 2096,Cambridge UK, 2001.

[13] F. Roli and J. Kittler Eds., "Proceedings of the Third International Workshop on Multiple Classifier Systems", Springer LNCS 2364, Grand Hotel Chia Laguna Cagliari Italy, 2002.

[14] B. Achermann and H. Bunke, "Combination of Classifiers on the Decision Level for Face Recognition", Technical Report IAM-96-002, Institut für Informatik und angewandte Mathematik, Universität Bern, 1996.

[15] SM. Lucas, "Continuous n-Tuple Classifier and its Application to Real-time Face Recognition", IEE Proceedings of Visual Image and Signal Processing, vol. 145 num.343-348, 1998.

[16] AS. Tolba and AN. Abu-Rezq, "Combined Classifier for Invariant Face Recognition", Pattern Analysis and Applications. vol. 3 num.289-302, 2000.

[17] G. L. Marcialis, F. Roli, "Fusion of LDA and PCA for Face Recognition". Department of Electrical and Electronic Engineering - University of Cagliari – Italy, 2002.

[18] M. Turk and A. Pentland, "Eigenfaces for recognition", J. of Cognitive Neuro-science, 3(1), 1991.

[19] M. Turk and A. Pentland. "Face recognition using eigenfaces", In Proc. IEEE Conf. on Comp. Vision and Patt. Recog., pages 586, 1991.

[20] P.Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", Proc. of the Fourth European Conference on Computer Vision, Vol. 1, 14-18 , Cambridge, UK, pp. 45-58, 1996.

[21] A. Martinez, A. Kak, "PCA versus LDA". IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp. 228-233, 2001.

[22] H. Kemal E., R. Stiefelhagen, "Local Appearance Based Face Recognition Using Discrete Cosine Transform", Interactive Systems Labs, University of Karlsruhe, 2005.

[23] Carreira-Perpinan, M. A, "Compression neural networks for feature extraction: Application to human recognition from ear images" (in Spanish). MSc thesis, Faculty of Informatics, Technical University of Madrid, Spain, 1995.