Journal of Computer Science 5 (5): 374-379, 2009
ISSN 1549-3636

Multimodal Face and Ear Images

A.A. Darwish, R. Abd Elghafar and A. Fawzi Ali
Department of Mathematics, Faculty of Science, Helwan University, Cairo, Egypt

Abstract: Problem statement: The study presented in this study to combined face and ear algorithms as an application of human identification. Biometric system to the detection and identification of human faces and ears developed a multimodal biometric system using eigenfaces and eigenears. **Approach:** The proposed system used the extracted face and ear images to develop the respective feature spaces via the PCA algorithm called eigenfaces and eigenears, respectively. The proposed system showed promising results than individual face or ear biometrics investigated in the experiments. **Results:** The final achieve was then used to affirm the person as genuine or an impostor. System was tested on several databases and gave an overall accuracy of 92.24% with FAR of 10% and FRR of 6.1%. **Conclusion:** The results display if we combined face and ear is a good technique because it offered a high accuracy and security.

Key words: Face recognition, ear recognition, PCA, algorithms, eigenfaces, eigenears, pattern recognition

INTRODUCTION

Biometrics refers to the use of physiological or biological characteristics to measure the identity of an individual. These features are unique to each individual and remain unaltered during a person's lifetime. These features make biometrics a promising solution to the society. The access to the secured area can be made by the use of ID numbers or password which amounts to knowledge based security. But such information can easily be accessed by intruders and they can breach the doors of security. The problem arises in case of monetary transactions and highly restricted to information zone. Thus to overcome the above mentioned issue biometric traits are used.

The various biometrics traits available are face, fingerprint, iris, palm print, hand geometry and ear. Among the available biometric traits some of the traits outperform others. The reliability of several biometrics traits is measured with the help of experimental results. The biometric system is basically divided into two modes i.e., unimodal biometric system and multimodal biometric system. In case of unimodal biometric system the individual trait is used for recognition or identification. The most successful applications of image analysis and understanding, face recognition has recently received significant attention, especially during the past several years, for example, recognition of face images acquired in an outdoor environment with changes in illumination and/or pose remains a largely unsolved problem. In other words, current systems are still far away from the capability of the human perception system, research in biometric systems has been increasing significantly due to international insecurity environment. Research groups around the world are developing algorithm and systems based on face, iris, fingerprint, palm print or voice and one other possible biometric source is the ear. Iannarelli performed important early research on a manual approach to using the ear for human identification[1]. Recent researchers that explore computer vision techniques for ear biometrics include those of[2,3].

1n our research laboratory, recognition with ear and face and their implementations on different databases are studying. Face recognition algorithm is mainly based on Principal Component Analysis (PCA)[4].

Background and related research: An overview on the major human face recognition techniques that apply mostly to frontal faces, advantages and disadvantages of each method are also given. The methods considered are eigenfaces and multimodal face and ear. The approaches are analyzed in terms of the facial representations they used. Eigenface is one of the most thoroughly investigated approaches to face recognition. It is also known as Karhunen-Loève expansion, eigenpicture, eigenvector and principal component. Some references used principal component analysis to efficiently represent pictures of faces. They argued that

Corresponding Author: A.A. Darwish, Department of Mathematics, Faculty of Science, Helwan University, Cairo, Egypt

any face images could be approximately reconstructed by a small collection of weights for each face and a standard face picture (eigenpicture). The weights describing each face are obtained by projecting the face image onto the eigenpicture. Another used eigenfaces, which was motivated by the technique of Kirby and Sirovich, for face detection and identification. In mathematical terms, eigenfaces are the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. The eigenvectors are ordered to represent different amounts of the variation, respectively, among the faces. Each face can be represented exactly by a linear combination of the eigenfaces. It can also be approximated using only the "best" eigenvectors with the largest eigenvalues. The best M eigenfaces construct an M dimensional space, i.e., the "face space". The authors reported 96, 85 and 64% correct classifications averaged over lighting, orientation and size variations, respectively. Their database contained 2,500 images of 16 individuals.

As the images include a large quantity of background area, the above results are influenced by background. The authors explained the robust performance of the system under different lighting conditions by significant correlation between images with changes in illumination. Recently, in[5] experiments with ear and face recognition, using the standard principal component analysis approach, showed that the recognition performance is essentially identical using ear images or face images and combining the two for multimodal recognition results in a statistically significant performance improvement. For example, the difference in the rank-one recognition rate for the day variation experiment using the 197 image training sets is 90.9% for the multimodal biometric versus 71.6% for the ear and 70.5% for the face. There is substantial related research in multimodal biometrics. For example[6] used face and fingerprint in multimodal biometric identification.

Eigenfaces technique:

Description: Principal Component Analysis (PCA, also known as "Eigenfaces"), is one of the most known global face recognition algorithm. The main idea is to decorrelate data in order to highlight differences and similarities by finding the principal directions (i.e., the eigenvectors) of the covariance matrix of a multidimensional data. For our experiments, we use several datasets the first dataset is provided by the Massachusetts Institute of Technology (MIT), second is ORL face database, third Yale face database. Each Gallery Set contains train subjects and test subjects. For testing our system, we use some face images from test subjects (same persons of the train Set but with changes in facial expressions).

Training the PCA: From a theoretical point of our view, a face image Γ_i can be seen as a vector is a huge dimensional space, concatenating the columns. We research with normalized face images that we preprocessed. For example of a MIT normalized face and ear image which used in our system show Fig. 1. We write new code with MATLAB to combined face and ear recognition in one algorithm using PCA and GUI for facilitate used database for training and test image of face and ear which used.

The first step is to train the PCA using the Training Set, in order to generalize the ability of our system and generate eigenvectors. We compute the mean image of the training data:

$$\Psi_{train} = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$$

Then each training image is mean-subtracted:

$$\Phi_i = \Gamma_i - \Psi_{train} \quad i = 1, 2, \ldots, M$$

This set of very large vectors is then subject to principal component analysis, which seeks a set of M orthonormal vectors, U_n, which best describes the distribution of the data. The k_{th} vector, U_k, is chosen such that:

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(U_k^T\Phi_n)^2$$

The vectors U_k and scalars λ_k are the eigenvectors and eigenvalues, respectively of the Covariance Matrix (CM):

$$C = \frac{1}{M}\sum_{n=1}^{M}(\Phi_n\Phi_n^T) = AA^T$$

Fig. 1: An example of a MIT normalized face and ear image used in our system

The mean image Ψ of the gallery set is computed. Each mean-subtracted gallery image, $\Phi_i = \Gamma_i - \Psi$, i = 1….M is Then projected onto the "Face Space" spanned by the M' eigenvectors deriving from the training set[4]. This step leads to:

$$\omega_k = U_k^T \Phi_i \qquad k = 1....M'$$

This describes a set of point-by-point image multiplication and summations. The weight from the vectors:

$$\Omega = [\omega_1, \omega_2, ..., \omega_k]$$

That describes the contribution of each eigenface or eigenear in representing the input face or ear image treating the eigenfaces or eigenears as a basis set of face or ear images[4]. Calculating a Euclidian distance is the simplest way to classify the new face or ear class as follows:

$$d_k = \|\Omega - \Omega_k\|$$

where, Ω_k is a vector describing the kth face or ear class. A face is classified as belonging to class k when the minimum d_k is in the defined threshold limit of ε_k. Otherwise, the new face or ear is defined as 'unknown'. The unknown face or ear can be used for developing further database. PCA computes the basis of a space which is represented by its training vectors. These basis vectors, actually eigenvectors, computed by PCA are in the direction of the largest variance of the training vectors. As it has been said earlier, we call them eigenfaces. Each eigenface can be viewed a feature. When a particular face is projected onto the face space, its vector into the face space describes the importance of each of those features in the face. The face is expressed in the face space by its eigenface coefficients (or weights). We can handle a large input vector, facial image, only by taking its small weight vector in the face space. This means that we can reconstruct the original face with some error, since the dimensionality of the image space is much larger than that of face space. In this study, let's consider face identification only. Each face in the training set is transformed into the face space and its components are stored in memory. The face space has to be populated with these known faces. An input face is given to the system and then it is projected onto the face space. The system computes its Euclidian distance from all the stored faces. However, two issues should be carefully considered:

- What if the image presented to the system is not a face?
- What if the face presented to the system has not already learned, i.e., not stored as a known face?

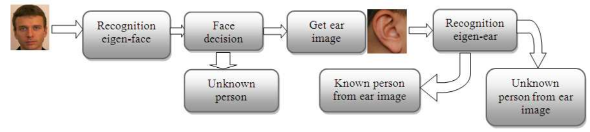

Fig. 2: Pictorial representation of the operations followed in the proposed method

The first defect is easily avoided since the first eigenface is a good face filter which can test whether each image is highly correlated with itself. The images with a low correlation can be rejected. Or these two issues are altogether addressed by categorizing following four different regions:

- Near eigenface and near stored face $\rightarrow$ known faces
- Near eigenface but not near a known face $\rightarrow$ unknown faces
- Near eigenear and near stored ear $\rightarrow$ known-faces from ear
- Near eigenear but not near a known ear $\rightarrow$ unknown-faces from ear

This is clear in Fig. 2 shows a representation of the operations followed in the proposed method which written in our project with MATLAB.

MATERIALS AND METHODS

The datasets: The research in this study was done using several datasets. The first dataset is provided by the Massachusetts Institute of Technology (MIT) containing a collection of facial images and side images used to construct the ear images from them (10 individuals with 4 face images and 4 ear images per individual). Figure 3 shows an example of images used in our research from the MIT dataset which is composed of 40 individuals with 10 face images per individual. Figure 4 shows an example of images used in our research from the ORL dataset which is composed of 15 individual with 11 face images per individual. Figure 5 shows an example of images used in our research from the SEARCH ear dataset. Images for individuals that we considered for our research were taken from different datasets to insure that they are taken on different sessions, different days and at different times of day. Some of the images were excluded from the datasets that we used due to poor quality or movement distortions.

Main recognition process: In this study, the recognition process is divided into two main steps.

Each step is treated as a separate recognition problem. This means that if we decide to identify an individual, we will have two images for him, one for his face and other for his ear, representing each image to be recognized separately, Fig. 6 shows a general view of the recognition process for individual images. To recognize that individual correctly each image will have to be classified correctly to be belonging to that individual. In the following of this study we will present the different datasets used with their corresponding recognition rates. We will also be presenting a method for combining the results of classification of individual images to come to a unified decision about the classification of the individual in question.

Fig. 3: Example of images from the MIT database used in our research

Fig. 4: Example of images from the Yale database used in our research.

Fig. 5: Example of images from the SEARCH ear database used in our research

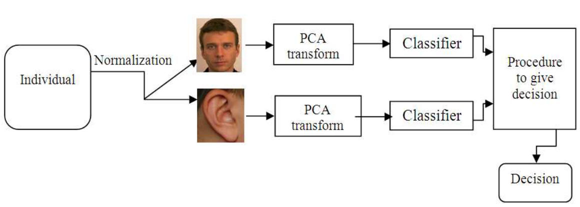

Fig. 6: A general view of the recognition process for individual images

RESULTS

Experimental results and analysis: experimental results that were obtained from the proposed face and ear recognition system are given, how we write the code using MATLAB for combined face and ear.

At first level face and ear algorithms are tested individually. At this level the individual results are computed. At this level the individual accuracy for face is found to be 68.16% as shown in Table 1.

However in order to increase the accuracy of the biometric system as a whole the individual results are combined at matching score level. At second level of experiment the matching scores from the individual traits are combined and final accuracy graph is plotted as shown in Fig. 7. Table 1 shows the accuracy and error rates obtained from the individual and combined system. The overall performance of the system has increased showing accuracy for face and ear of 92.24% with FAR of 10% and FRR of 6.1% respectively. FAR graph is plotted as shown in Fig. 8 and FRR graph is plotted as shown in Fig. 9.

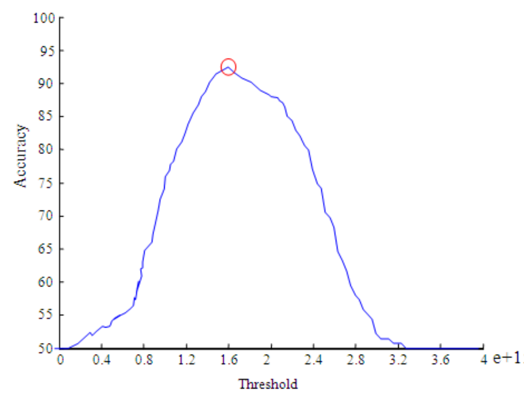

Fig. 7: FRR curve for combined face and ear

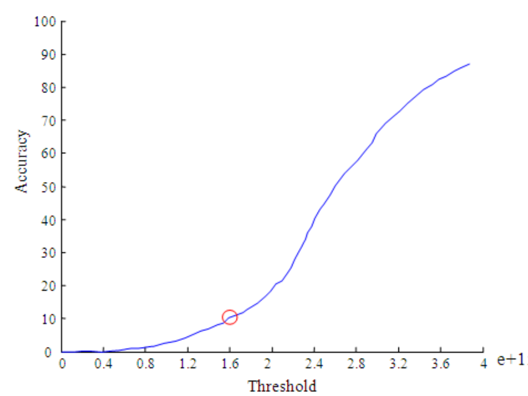

Fig. 8: FAR curve for combined face and ear

Table 1: Showing individual and combined accuracy

Trait	Algorithm	Accuracy (%)	FAR (%)	FRR (%)
Face	PCA	68.16	31.2	14.1
Face and ear	PCA	92.24	10.0	6.1

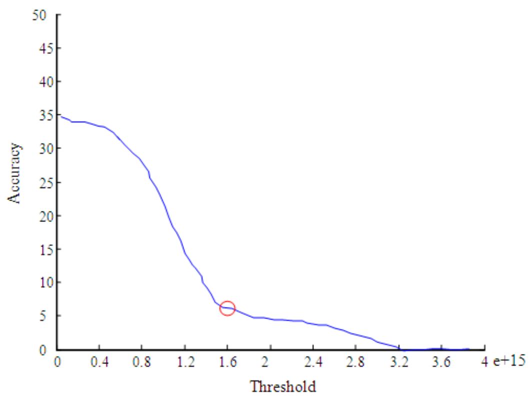

Fig. 9: FRR curve for combined face and ear

Applying recognition procedure on used datasets: Each image database used in our research is divided into two training set images, face and ear images and their two corresponding test set images. Images for 3 individuals were put in the test set of each dataset to construct the set of imposters and the images for a 10 chosen individuals from every dataset where divided into training and test face and ear sets as follows: The first two images per individual will construct the training set and the last two images per individual will be part of the test set. Principal components will be calculated for each individual image separately and the images will be transformed to the PCA space using their corresponding transformation matrix as discussed previously in Eigenfaces technique.

The individual images are normalized and preprocessing operations performed and then clear face and ear images are constructed. The individual images are transformed to the PCA space. Every image is recognized with its corresponding classifier. A Procedure is applied for reaching a unified decision.

DISCUSSION

We discuss the pros and cons of using ears as a biometric for identifying humans. We also explain the logic behind combining the images post PCA calculations and the weakness of facial recognition and the benefit of using a multimodal system for human identification.

What's wrong with using face recognition? Face recognition has been researched a lot in the past years and a lot of algorithms, feature extraction techniques and classification techniques have been developed for that purpose, but it all comes down to the efficiency of the feature extraction. Facial features are susceptible to many factors such as mood, health, facial hair and facial expressions. This is a natural barrier in using face as a reliable means for human identification. The feature extraction technique used will have to deal with the material at hand, so no matter how good the feature extraction process used is, the condition of the face presented will determine the outcome.

Why use ears? The use of ears as a biometric for human identification has not been researched as intensively as other biometrics has been researched. Although research in this area is relatively small, the research that has been done showed a lot of promise in using the ear as a biometric for human identification.

The ear much like the face is a visible part of the human body that can be used for a non invasive biometric technique. Humans most likely will have to keep their ears uncovered to be able to hear. The ears unlike the face are unaffected by ageing, in fact the ear undergoes very slight changes from infancy to adulthood, in fact the only change that happens is elongation due to gravity. The ears also do not suffer the change in appearance by hair growth like the face does.

Although these are all pros for using the ears as a biometric, but using the ears for human identification has some disadvantages. These disadvantages are embodied in occlusions. Sources of occlusion may be long hair, earrings and multiple piercings.

CONCLUSION

In this study, the present study has aimed to develop a multimodal biometric system for personal identification. Experimental results have shown that combined face and ear recognition system. This system, after studying is implemented on different databases. The results display if we combined face and ear is a good technique because it offers a high accuracy and security. In a near future, we plan to use other algorithms and compare it to do best and increase the accuracy for ear and face. We plan to study the implementation of some preprocessing steps such as face detection, eyes detection and face normalization. We are also researching on fusion of iris and face and we are trying to develop a bimodal biometric system for recognition.

ACKNOWLEDGMENT

Portions of the research in this study use the MIT, ORL and Yale database of facial images collected under the FERET program.

REFERENCES

1. Iannarelli, A., 1989. Ear Identification, Forensic Identification Series. Fremont, Paramont Publishing, Calif, ISBN: 10: 0962317802, pp: 213.
2. Burge, M. and W. Burger, 2000. Ear biometrics in computer vision. Proceeding of the 15th International Confereence on Pattern Recognition, IEEE Xplore Press, USA., pp: 822-826. DOI: 10.1016/j.cviu.2004.11.001
3. Hurley, D.J., M.S. Nixon and J.N. Carter, 2002. Force field energy functional for image feature extraction. Image Vis. Comput. J., 20: 311-317. DOI: 10.1016/S0262-8856(02)00003-3
4 Turk, M. and A. Pentland, 1991. Eigenfaces for recognition. J. Cognit. Neurosci., 3 : 71-86. http://portal.acm.org/citation.cfm?id=1326894
5. Chang, K., K.W. Bowyer and S. Sarkar, 2003. Comparison and combination of ear and face images in appearance-based biometrics. IEEE. Trans. Patt. Anal. Mach. Intell., 25: 1160-1165. DOI: 10.1109/TPAMI.2003.1227990
6. Hong, L. and A. Jain, 1998. Integrating faces and fingerprints for personal identification. IEEE. Trans. Patt. Anal. Mach. Intell., 20: 1295-1307. DOI: 10.1109/34.735803