Proposing a Passive Biometric System for Robotic Vision

M. Masudur Rahman, Seiji Ishikawa
Department of Mechanical and Control Engineering
Kyushu Institute of Technology, Japan
Email: rahman@ss10.cntl.kyutech.ac.jp

Abstract:
We present a passive biometric system to the detection and identification of human faces and ears and develop a multimodal biometric system using eigenfaces and eigenears. A new technique of tracking a human face and an ear from the same image of a particular person is proposed that provides us to use a single camera for developing the proposed biometric system. The proposed system uses the extracted face and ear images to develop the respective feature spaces via the PCA algorithm called eigenfaces and eigenears, respectively. A new face and ear can be characterized by calculating the Euclidean distance between the classes of eigenfaces or eigenears and the new face or ear, respectively. Eighteen persons' faces and ears are employed for developing the databases of the eigenfaces and eigenears, and new faces and ears taken from various sessions of the same persons are employed for the identification. The proposed multimodal biometric system shows promising results than individual face or ear biometrics investigated in the experiments.

1. Introduction
Many of security companies are seeking a suitable biometric system that can create its database and identify targeted person passively. Such a system is especially important for robotic vision where subject's cooperation or direct interaction of human and robot may not always be available. In some cases, a subject's involvement may cause severe threat to the biometric systems especially for identifying a suspect or a terrorist. In addition, an automatic personal identification system based solely on a particular biometric component such as fingerprint or face is often not able to meet the security satisfaction [1]. We need such a system where acquisition of biometric database and its verification could be developed without a subject's concern, and the system performance will satisfy the security requirement. A new multimodal biometric system based on eigenfaces [2] and eigenears [3] is introduced in this paper. We develop a passive multimodal biometric system where subject's concern is not necessary. The proposed system can be implemented for developing a passive identification system employing a humanoid robot and it can be of a suitable application to an airport security system.

In the past, various studies have been done in biometric systems. However, the area of passive and/or multimodal biometric systems is quite new in biometric research. We have seen some of studies where face and ear have been chosen individually as the biometric components [4, 6]. However, the present study merges the eigenfaces and eigenears for developing a passive multimodal biometric system where face and ear are the biometric component.

In this paper, we propose and develop a prototype passive biometric system using eigenears and eigenfaces. The images of the ears and faces are collected from a single image of a particular person, and the databases of eigenears and eigenfaces via the PCA algorithm are produced. New type of biometric data (ear or face) is verified using the databases, and personal verifications are done judging the match-record. Since the proposed study extracts face and ear from the same image, the system does not require the application of multiple cameras. Or, we can also employ a ceiling camera for extracting such images. To the best of our knowledge, this is a new method in the biometric systems.

We have presented a brief formulation of the proposed biometric system. In the experiment, we have presented three different results for comparing the face, ear and multimodal biometrics. A discussion and concluding remarks are also placed at the end of this study.

2. Eigenfaces and Eigenears
PCA is a very powerful technique for selecting appropriate features or to reduce the dimension in which variation in the dataset is preserved. Thus the

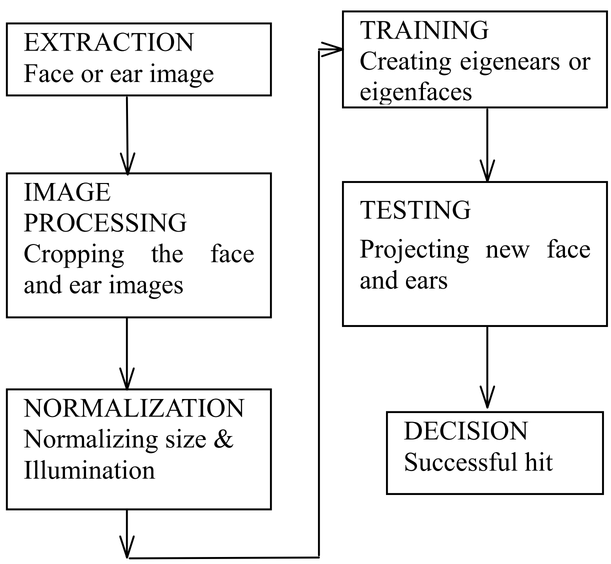


Figure 1: Flow chart of the system's algorithm.

classification is done in a lower dimensional space called the eigenspace, which is just the space defined by the principal components or the eigenvectors of the data set.

In case of face or ear images, the principal components are called eigenfaces or eigenears, respectively. The proposed system performs various procedures that are shown in **Figure 1**. In the extraction stage, acquired training face and ear images are cropped to a specified size from the same image of a particular subject. Image processing step includes geometric normalization and masking. Different masks can be used to avoid the unnecessary parts of the images. In the verification stage, it consists of two stages; training stage and testing stage. In the training phase, the eigenears and eigenfaces are developed choosing the appropriate eigenvectors. In the testing phase, the decision of the identification is taken based on the Euclidean distance between the classes of eigenfaces or eigenears and new faces or ears, respectively.

2.1 Calculating Eigenfaces or Eigenears

Let a face or ear image *I*(x,y) be a two dimensional *N* by *N* array of intensity or a vector dimension N^2. Let the training set of face or ear images be

$$\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots \boldsymbol{x}_M \qquad (1)$$

The average face or ear of the set is defined by

$$\boldsymbol{\psi} = \frac{1}{M}\sum_{n=1}^{M} \boldsymbol{x}_n \qquad (2)$$

Each face or ear differs from the average by the vector $\boldsymbol{\varphi}_i = \boldsymbol{x}_\mathrm{i} - \psi$. A covariance matrix can be obtained by

$$\begin{aligned} C &= \frac{1}{M}\sum_{i=1}^{n} \boldsymbol{\varphi}_i \boldsymbol{\varphi}_i^\mathrm{T} \\ &= \mathbf{X}_n \mathbf{X}_n^T \end{aligned} \qquad (3)$$

where, X is the new image set obtained by the subtraction result between each image and φ_i. This set of very large vectors is then subject to principal component analysis that seeks a set of *k* orthogonal vectors or the principal components $\boldsymbol{e}_i$ and their associated eigenvalues λ_i which best describes the distribution of data. The vectors $\boldsymbol{e}_k$ and scalars λ_k are the eigenvectors and eigenvalues, respectively, of the covariance matrix. The *N* dimensional space defined by all the eigenvectors of *C* is reduced via the principal component analysis [8]. One may find a simple method to select the eigenvectors proposed by Turk and Pentland [2]. As a result, the chosen *K* ($1 \le k << N^2$) eigenvalues $\boldsymbol{\lambda}_k$ (*k*=1,2,…,*K*) and corresponding eigenvectors $\boldsymbol{e}_k$ are obtained. The *k*-dimensional vector defined by the eigenvectors $\boldsymbol{e}_k$ is called an eigenface or eigenear. Because these vectors are the eigenvectors of the covariance matrix corresponding to the original face or ear images and they are faces or ears like their appearances.

2.2 Classifying new Faces or Ears using Eigenfaces or Eigenears

After creating the eigenfaces or eigenears, classification of new images becomes a pattern recognition matter. A new face or ear image (**x**) is transformed into its eigenface or eigenear component projected onto the eigenfaces or eigenears by the following way,

$$w_k = \boldsymbol{e}_k^T(\mathbf{x} - \psi) \qquad (4)$$

for $k = 1, 2, ..., k'$. This describes a set of point-by-point image multiplication and summations. The weight from the vectors $\boldsymbol{\Omega}^\mathrm{T} = [w_1, w_1, ..., w_{k'}]$ 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 an Euclidian distance is the simplest way to classify the new face or ear class as follows,

$$d_k = \left\| (\Omega - \Omega_k) \right\| \qquad (5)$$

where Ω_k is a vector describing the *k*th 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.

3. Experimental Results

Three different issues are considered for performing this study. We have initially developed two individual biometrics using face and ear separately. Finally, a multimodal biometric system is developed merging the face and ear biometrics.

As we have described in the Section 2 and Figure 1, there are three main steps involved in developing the proposed system: (a) extraction of images, pre-processing and normalization, (b) creating eigen databases for ears and faces, and (c) verification. We have employed the databases of university of Notre Dame [9] which is composed of five different sessions and imaging conditions. The database contains five different variations based on the sessions collected in the database. We employ one set for creating the databases of eigenears and eigenfaces, and four other variations are used for the verification. In the preprocessing step, we have used only the side faced-images for obtaining the (right) ears and faces together. **Fig. 2** shows 5 person's images out of total 18 employed in this study. Some of image samples used in the experiments are shown in **Fig. 3** along with extraction procedures of the ear and face

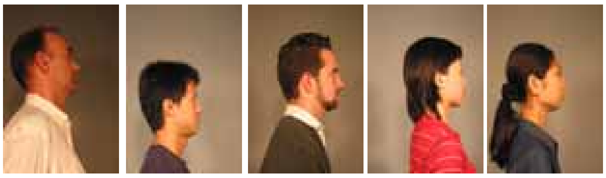


Figure 2: Some of models employed in the experiment out of total 18.

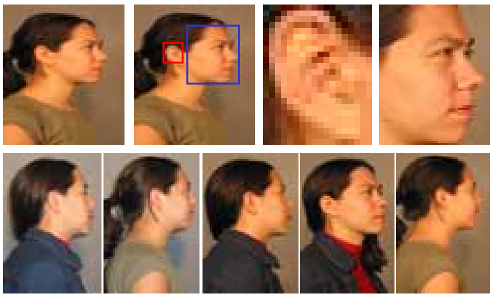


Fig. 3. Procedures of extracting ear and face images (upper), and image variations used in the experiments (lower).

area (upper row). In the Fig. 3 (lower row), 5 images of a particular subject taken from 5 different sessions are shown. The extraction of faces and ears are yet not automatic, as we need some manual interruption for selecting the exact size of the ear and face. The images are cropped to a specific size with just containing the face and ear. Finally, the original size of the cropped images is reduced to a size of 32x32 pixels for the sake of efficiency. Now, the images are normalized for minimizing the lighting effect.

We have developed the databases of eigenears and eigenfaces of 18 subjects as described in the section 2. Performance of the verifications is obtained based on successful hits by calculating the distance within the prescribed threshold. In case of multimodal biometric system, if the system recognizes any one of the ear or face of a particular person successfully, we consider the correct identification of the subject. Detail of experimental dataset is placed in **Table 1.** We have also obtained the classification results on various issues that are placed in **Fig. 4, Fig., 5, Fig. 6** and **Fig. 7**. Performance based on eigenvectors, eigenvector vs. CPU time, and performance based on day variations are shown in Fig 4, Fig. 5 and Fig. 6, respectively. Fig. 7 shows the performance of the proposed multimodal system based on required eigenvectors. It has shown that multimodal biometrics has better performance over individual face and ear biometrics. In comparison between face and ear biometrics, face biometrics has shown better performance than ear biometrics. However, ear biometrics has shown better performances in case of CPU time and day variations. Ear biometrics has absolutely no effect on day variations. We have obtained the best recognition rate of 94.44% by multimodal biometric system on 40 eigenvectors. The best performances of each system are also placed in Table 1.

4. Discussion

The present study has aimed to develop a multimodal biometric system for personal identification that can also work passively. Experimental results have shown that the concept of tracking the ear and face from the same image of a person has received better performances over the many available methods [4-7]. Although there was no intention to make any recommendation in favor of any particular biometric systems, it is observed that face biometrics has performed better than ear biometrics. The

Table 1: Experimental setup

Biometrics	Total images	Test Variations	Best Perform.
Face	5x18=90	4	88.88 %
Ear	5x18=90	4	77.77%
Multimodal	2x5x18=180	4	94.44%

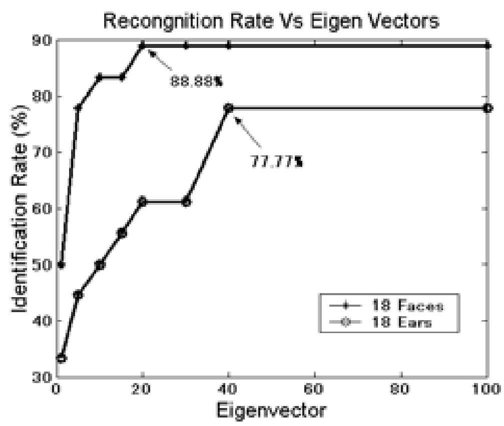


Figure 4: Experimental results on eigenvectors vs. recognition rates.

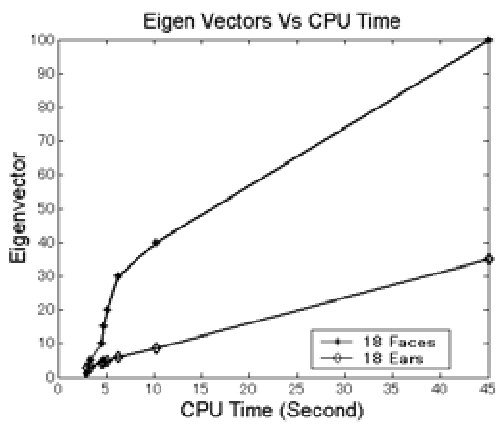


Figure 5: Experimental results on CPU time vs. eigenvectors.

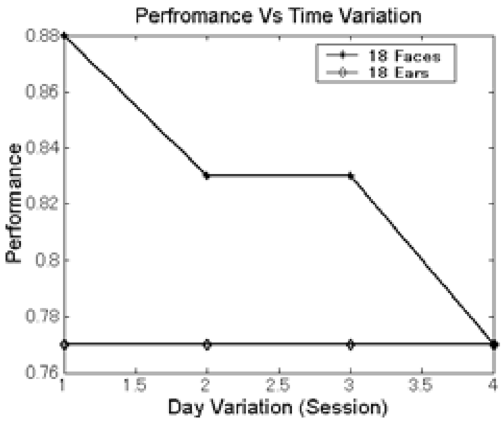


Figure 6: Experimental results on recognition rate vs. day variations.

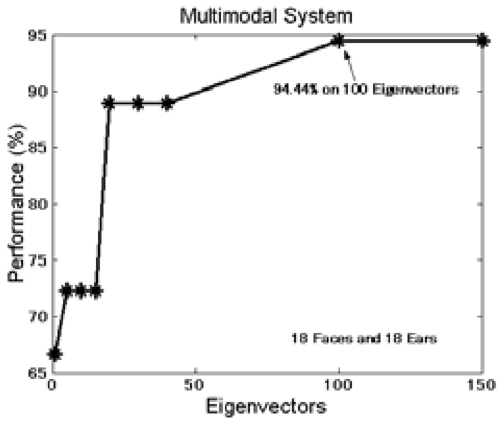
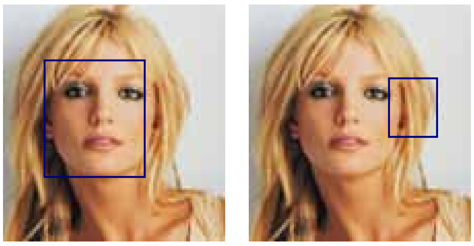


Figure 7: Multimodal biometric system. Experimental results on eigenvectors vs. recognition rate.



(a) (b)

Figure 8. Occlusion problems in extracting face and ear images. Extraction of (a) face is possible but (b) ear is not possible due to long hair. *Images of Britney Spears collected from the Internet*

results presented in this paper differ slightly with the paper by Burge and Burger [7] where they have showed ear-based biometrics better than face biometrics. Variations on the ears and faces may give us different results as described by Chang and Bower [3]. Our results, on the other hand, do support the conclusion that a multimodal biometric using eigenfaces and eigenears can perform better than either one alone. We have seen the same results in case of using face and fingerprint or face and voice [1]. However, use of face and ear has the best combination for developing a passive biometric system. This research may open up some new multimodal biometric systems with merging two or more biometric systems such as face and gait or face, ear and gait.

Since the proposed biometric system can detect and identify human face and ear passively, it can be of a suitable application in a robotic vision field. A robot can be a part of security systems in an airport for security checking.

The present study has some limitations on several factors, e.g., occlusion, use of appropriate mask, etc. The occlusion may sometimes greatly influence to the proposed system. **Fig. 8** shows a simple example where occlusion problem makes impossible to use the ear as a biometric component. Although face biometric can be implemented in this particular case, a multimodal system cannot be developed due to absence of ear image. There are some extensive works on extracting face and ear images using different masking. Although use of such masks needs some manual work, better results can be obtained in different imaging conditions [3,9]. We need some further studies with considering the mentioned factors such as occlusion by facial hair or partially covered ear by long hair, and different masking.

5. References

1. L. Hong and A. Jain: "Integrating Faces and Fingerprints for Personal Identification", *IEEE Trans. on PAMI*, Vol. 20, No. 12, pp. 1295-1307 (1998).
2. M. Turk and A. Pentland: "Eigenfaces for Recognition", *J. of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86 (1991).
3. K. Chang and K. W. Bowyer: "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics", *IEEE Trans. on PAMI*, Vol. 25, No. 9, pp. 1160-1165(2003).
4. M. Turk and A. Pentland: Face recognition Using Eigenfaces, *Proceedings of the CVPR*, pp. 586-591(1991).
5. L. Sirovich and M. Kirby: "Low Dimensional Procedure for Character of Human Faces", *J. of Opt. Society of America*, Vol. 4, No. 3, pp. 519-524(1987).
6. B. Victor, K. Bower and S. Sarkar: "An Evaluation of Face and Ear Biometrics", *Proceedings of the ICPR* CDROM Version (2002).
7. M. Burge and W. Burger: "Ear Biometrics in Computer Vision", *Proc. of the Int'l Conf. on Pattern Recognition*, CD ROM (2002).
8. R. C. Gonzalez and P. Wintz: *Digital Image Processing*, Addison-Wesley Publishing Company, Inc. (1992).
9. P. J. Flynn, K. W. Bowyer, and P. J. Phillips: "Assessment of time dependency in face recognition: An initial study," *Audio and Video-Based Biometric Person Authentication*, pp.44-51 (2003).