

Ear Biometrics: A Small Look at the Process of Ear Recognition

Pedro Luis Galdámez and María Angélica González Arrieta

University of Salamanca, Plaza de los Caídos, 37008 Salamanca, Spain
{peter.galdamez,angelica}@usal.es

Abstract. This document provides an approach to biometrics analysis which consists in the location and identification of ears in real time. Ear features, which is a stable biometric approach that does not vary with age, have been used for many years in the forensic science of recognition. The ear has all the properties that a biometric trait should have, i.e. uniqueness, permanence, universality and collectability. Because it is a field of study with potential growth, in this paper, we summarize some of the approaches to the detection and recognition in existing 2D images in order to provide a perspective on the possible future research and the develop of a practical application of some of these methodologies to create finally a functional application for identification and recognition of individuals from an image of the ear, the above in the context of intelligent surveillance and criminal identification, one of the most important areas in the processes of identification.

Keywords: Neural Network, System Identification, Ear Recognition.

1 Introduction

The Ear does not have a completely random structure. It has standard parts as other biometric traits like face. Unlike human face, ear has no expression changes, make-up effects and more over the color is constant throughout the ear. Although the use of information from ear identification of individuals has been studied, is still an open question by specifying and determining whether or not the ear can be considered unique or unique enough to be used as a biometric. Accordingly, any physical or behavioural trait can be used as biometric identification mechanism provided which is universal, that every human being possesses the identifier, being distinctive and unique to each individual, invariant in time, finally measurable automatically or manually, the ear accomplish all these characteristics. From this reading we can conclude that the detailed structure of the ear is not only unique, but also permanent, the ear does not change during the human life. Furthermore, capturing images of the ears do not necessarily requires the cooperation of a person, so it can be considered non-intrusive. Because of these qualities, the interest in recognition systems through the ear has grown significantly in recent years and generally the increasing need to automatically authenticate people.

Traditional means of automatic recognition such as passwords or ID cards can be stolen, counterfeit, or forgotten, however, the biometric characteristics such as the ear, are universal, unique, permanent, and measurable, leaving the field open for further continuously research. The figure 1 shows the major steps of a traditional recognition system.

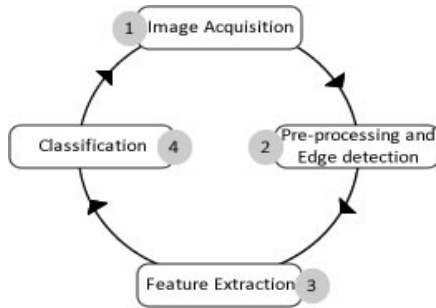


Fig. 1. Major Steps of ear recognition system

The Diagram describes task flow of a traditional recognition system, this flow is expose in the next sections, where we will see preprocessing methods and the flow catches ranging from the acquisition of the images to their identification, subsequently we will make emphasis in techniques applied in the heart system to achieve the goal of ear recognition.

2 Image Acquisition

In this section we need to differentiate from two tasks, the first one is the creation of the Ears Database, and second one is the image acquisition of the person to identify.

2.1 Ear Database

When we talk about an ear database is undisputed mentioned that exist a large group of databases for ear detection and recognition, this information is vital to test and compare the detection or recognition performance of a computer vision system, in general. The University of Science and technology in Beijing offers four collections of images that we could use to test taking any ear recognition system as they. The University of Notre Dame (UND) offers a large variety of different image databases and the Hong Kong Polytechnic University IIT Delhi has a database that consists in 421 images.

All this set of images could be used to test the performance of a ear recognition system by computer, but in our research we created our own set of images with

the help of the police school of Ávila, which is a city located in Spain. This database consist in two hundred images two for each person which does not means that we do not use the other datasets, in fact, for this research we also are using the IIT Delhi dataset. [10] The figure 2 shows a sample of the dataset created.



Fig. 2. Ear dataset. Police School of Ávila.

In the images obtained we tried to prevent the occlusion that hair and ear rings can make. The goal of create this dataset is to test our system with images that we already know to whom it belongs, in short words, this is our test set and sometimes we could use these images to train the system, in fact, we have the profile face of each person which we use to test algorithms to identify ears. This brings us to the second point to be analyzed in this section.

2.2 Image to Identify

Now, we need to capture the image to be analyzed, for that we use a simple web camera. The figure 3 shows a screen capture from the application that we are developing in order to achieve an application able to recognize a person in real time from a snapshot of its ear. For now it is a small system that can capture the ear from a web cam video using EmguCV [4] which is a wrapper from OpenCV that allows to develop in Visual Studio .Net and Java.

The application allows the user to create and manipulate images applying image filter like color filters and edge detectors. It is an image lab processing that also permits to introduce video documents as shown in figure 3, also we can edit the configuration to select what object we want to detect, this can vary, for example we could detect frontal and profile faces, right and left ear and eyes. But in this project we are focussing in ears, how we use the emgucv framework, we have access to the viola-jones classifier included in that library, to detect the ear in video, this application is using the haar-cascade classifier developed by

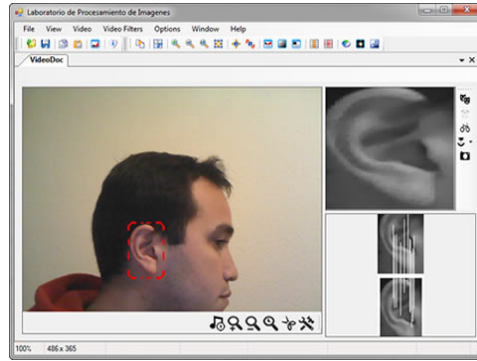


Fig. 3. Ear capture in the application

Modesto Castrillón-Santana [2] this allow us to capture ears on a easy way in real-time video. With the acquisition of the images defined, we must move into the second stage of the recognition system which is the pre-processing of the images.

3 Image Pre-processing and Feature Extraction

With each image in the database and the captured image, we begin the normalization process, first we perform the segmentation of the image applying a mask to extract only the ear, then the image is converted to an edge map using the canny edge filter. The Figure 4 is a summary of the pre-processing activities.

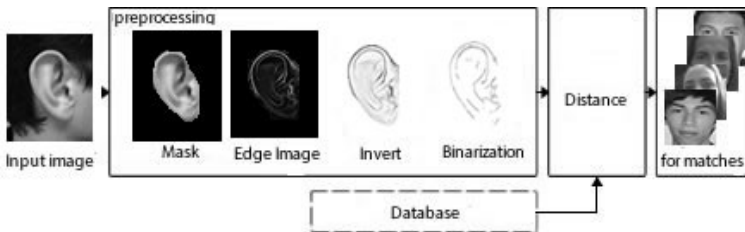


Fig. 4. Image pre-processing

If w is the width of the image in pixel and h is the height of the image in pixel, the canny edge detector takes as input an array $w \times h$ of gray values and sigma. The output is a binary image with a value 1 for edge pixels, i.e., the pixel which constitute an edge and a value 0 for all other pixels. We calculate a line between major and minor y value in the edge image to rotate and normalize each image, trying to put the lobule of the ear in the centre. This process is to try to get all the images whose shape is similar to the image to identify. We identify some

points on the external shape of the ear and the angle created by the center of the line drawn before and the section in the ear’s tragus with the major x value. All these values will be used like complement of an input in a neural network. Once we have all data processed we proceed to extract the features for the recognition.

4 Classification

4.1 Principal Component Analysis (PCA)

The ear recognition algorithm with eigenears is described basically in the figure 6. First, the original images of the training set are transformed into a set of eigenears E , Then, weights are calculated for each image on the (E) set, and then are stored in the (W) set. Observing an image X unknown, weights are calculated for that particular image, and stored in the vector W_X . Subsequently, W_X compared to the weights of images, which is known for sure that they are ears (the weights of the training set W) [8].

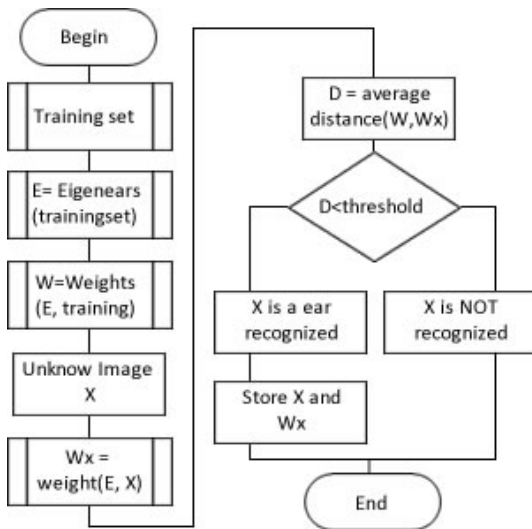


Fig. 5. Ear recognition algorithm based on eigen vectors

Classification of a New Ear. The process of classifying a new ear in the Γ_{new} to another category (known ears) is the result of two steps. First of all, the new image is transformed into its eigenear components. The resulting weights forms the weight vector Ω_{new}^T .

$$\omega_k = u_k^T(\Gamma_{new} - \Psi) \quad k = 1, \dots, M'$$

$$\Omega_{new}^T = [\omega_1 \ \omega_2 \ \dots \ \omega_{M'}]$$
(1)

The Euclidean distance between two vectors of weights $d(\Omega_i, \Omega_j)$ provides a measure of similarity between the corresponding images i and j . If the Euclidean distance between Γ_{new} and the rest of images on average exceeds a certain threshold value, through this can be assumed that Γ_{new} is not a recognizable ear [8].

4.2 Fisher Algorithm

The PCA algorithm is a linear combination of functions that maximizes the variance of the information. This can result in poor performance, especially when we are working with image noise such as changes in the background, light and perspective. So the PCA algorithm can find faulty components for classifying. To prevent this problems, we implement the Fisher algorithm to compare results in the ear recognition process. The Fisher algorithm that we implement basically goes like this [1,13]:

Construct the Image matrix x with each column representing an image. Each image is assigned to a class in the corresponding class vector c . Project x into the $(N - c)$ dimensional subspace as P with the rotation matrix $WPca$ identified by a Principal Component Analysis, where N is the number of samples in x .

c is unique number of classes ($length(unique(C))$) Calculate the between-classes scatter of the projection P as:

$$Sb = \sum_{i=1}^c N_i * (mean_i - mean) * (mean_i - mean)^T \quad (2)$$

Where $mean$ is the total mean of P $mean_i$ is the mean of class i in P , N_i is the number of samples for class i . Calculate the within-classes scatter of P as:

$$Sw = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - mean_i) * (x_k - mean_i)^T \quad (3)$$

Where x_i are the samples of class i x_k is a sample of x_i $mean_i$ is the mean of class i in P . Apply a standard Linear Discriminant Analysis and maximize the ratio of the determinant of between-class scatter and within-class scatter. The solution is given by the set of generalized eigenvectors $Wfld$ of Sb and Sw corresponding to their eigenvalue. The rank of Sb is atmost $(c - 1)$, so there are only $(c - 1)$ non-zero eigenvalues, cut off the rest. Finally obtain the Fisher ears by $W = WPca * Wfld$ [13].

4.3 Ear Classification

With all the pictures processed, an array is created with labels indicating to whom belongs each image, the vector values are calculated with the PCA and Fisher algorithm. Subsequently, it obtains the Euclidean distance of the comparison of the weight vectors obtained. Defining a threshold value of 1000 representing a similarity between two images of 90%, under the assumption that a smaller distance means a greater similarity between the sets.

In other words, the eigen and fisher vectors are obtained from the set of images stored in the database, the vectors values and the unknown image are compared both with the distance measurement, resulting in a distance vector from the input image regarding the collection of the database. The lowest value is obtained and compared with the threshold, being the label with the lower value the identified user.

5 Neural Network for Ear Recognition

Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, speech, vision and control systems. In this project, there is a neural network that identifies each person in the database. After calculating the eigen and fisher ears, the feature vectors of the ears are stored in the database. These vectors are used as inputs to train the network. In the training algorithm, the vectors of values belonging to a person, are used as positive for returning said individual neuron 1 as the neuron output assigned to that user and 0 in other neurons.

When the new image has been captured, the feature vectors are calculated from the eigenears obtained before, we compute new descriptors of the unknown image. These descriptors are entered into the neural network, the outputs of individual neurons are compared, and if the maximum output level exceeds the predefined threshold, then it is determined that the user belongs to the ear assigned to the neuron with the index activated. The algorithm implementation in the approach of ear recognition using eigen, fisher and geometric image pre-processing with canny edge filter basically is summarized building a library of ears, choose a training set M that includes if it is possible more than one image per person, these images can have variation in lighting and perspective, use the eigen and fisher values as input in a neural network, in our case we add at the input two values previously calculated.

These values are the average of the points detected in the geometric normalization, and the angle that forms the center and the major x value on the edge map in the tragus area of the ear. Create one output neuron per person in the database and finally after the training select the neuron with the maximum value. If the output of the selected neuron passes a predefined threshold, is presented as the recognized person.

6 Experimental Results

In this section we present each of the mentioned ear recognition techniques applied over the new database created, our first hypothesis says that variable illumination could affect the performance of the algorithms. In the process, all images were cropped within the ear therefore the contour around the ear was excluded. classification was performed using a nearest neighbour classifier. All training images of an individual were projected into the feature space. Each image in the database was taken while the subject ear was being illuminated by enough light. The methods reveal a number of interesting points:

1. Both algorithms perform perfectly with frontal light, however, as we are using a real-time video, the change of perspective, and when the lighting became darkness made a significant performance difference between the two methods.
2. The algorithm with less errors classifying the ear when we change the illumination and perspective was Fisher method.
3. Neural network using as input the eigenvalues computed have accomplished better performance in ear recognition process with change on illumination.

Table 1. Normal Conditions

	<i>PCA</i>		<i>Fisher</i>		<i>NeuralNetwork</i>	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	101	19	177	17	163	18
<i>Negative</i>	11	88	18	21	8	33

Under normal conditions we obtain the previous confusion matrix, assuming that true positive values are people classified correctly. True negatives are the users that the system should not recognize because they do not exist in the database, and indeed does not recognize, a false negative is when the system predicts that the user does not exist but it does exist, finally a false positive is when the system make a mistake to identify a person, but really it is not in the database.

Table 2. Changing Illumination and Perspective Conditions

	<i>PCA</i>		<i>Fisher</i>		<i>NeuralNetwork</i>	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	64	72	93	32	84	44
<i>Negative</i>	35	67	23	46	23	47

With perspective and illumination in normal conditions, we get 87% of succeed in recognition with PCA, 91% with fisher algorithm, using the neural network, the percentage increased to 93%, over more than 200 attempts of different individuals. When we change illumination and perspective conditions the error rate increase, leaving the PCA algorithm with a success rate of only 54%, neural network with 71% and Fisher Algorithm in 77%. This percentages are calculated using the F_1Score (over tables 1 and 2) which associate recall and precision measures to give us a value that represent, how well the system makes the recognition.

The method that has being used in this research is to try to put together some of the most common approaches in the recognition process, the project is

not presented as unique and exceptional, but upon the approaches that other researchers have proposed, combining and comparing them, and trying to select a combination of these approaches to successfully implement a fully functional system capable of recognizing a person across its ear and use this system to identify criminals, using the database created in the police academy of Ávila. As the title of this paper indicates this is a small look at the process of recognition of people from captured images of their ears in real-time video. Throughout the article presents the first techniques studied, therefore the quantitative results of the project can be considered preliminary, but they provide a clearer picture to where should point this research in the future, observing some of the strengths and weaknesses of the algorithms studied in order to strengthen pre-processing tasks and / or implementation of more robust algorithms.

7 Conclusion and Future Work

The algorithms perform a good ear recognition process if the video captures an image very similar with one in the training set. Fisher method appears to be better over variation in lighting. Use a neural network with the eigenvalues calculated as input makes a better performance than the eigenear traditional method over changes on illumination and perspective. Changes in pre-processing process allows better results if all images have the same angle and illumination, other techniques of pre-processing images may improve the ear recognition process. If these techniques allow recognize a person through the ear, exist other methods like Ray Image Transform, Histograms of Categorized Shapes, Edge orientation pattern that can obtain better results.

Our goal is to create an application that identify one person in a real-time video for that we are interested in the study of these techniques. This paper is our first look to the ear recognition process with encouraging results in real-time video identify process. Our future work will be modify the application to use more complex algorithms like Ray Image Transform and improve the pre-processing step.

References

1. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1997)
2. Castrillón-Santana, M., Lorenzo-Navarro, J., Hernández-Sosa, D.: An Study on Ear Detection and Its Applications to Face Detection. In: 14th Conference of the Spanish Association for Artificial Intelligence, La Laguna, Spain, pp. 313–322 (2011)
3. ElBakry, H.M., Mastorakis, N.: Ear Recognition by using Neural Networks. Technical University of Sofia Faculty of Computer Science and Information, Sofia (2007)
4. Intel OpenCV wrapper EmguCV (April 08, 2013), http://www.emgu.com/wiki/index.php/Main_Page
5. Islam, S., Davies, R., Mian, A., Bennamoun, M.: A Fast and Fully Automatic Ear Recognition Approach Based on 3D Local Surface Features, The University of Western Australia, Crawley, WA 6009, pp. 1081–1092. *ACIVS, Australia* (2008)

6. Lammi, H.-K.: Ear Biometrics. Lappeenranta, Lappeenranta University of Technology, Department of Information Technology, Laboratory of Information Processing, Finland (2004)
7. Marti-Puig, P., Rodríguez, S., De Paz, J.F., Reig-Bolaño, R., Rubio, M.P., Bajo, J.: Stereo Video Surveillance Multi-Agent System: New Solutions for Human. *Journal of Mathematical Imaging and Vision* (2011), ISSN: 0924-9907, doi:10.1007/s10851-011-0290-2
8. Dimitri, P.: Eigenface-based facial recognition (2002)
9. Prakash, S., Gupta, P.: An Efficient Ear Recognition Technique Invariant to Illumination and Pose. Indian Institute of Technology Kanpur, Kanpur-208016, India.: Department of Computer Science and Engineering (2010)
10. Pug, A., Busch, C.: Ear Biometrics: A Survey of Detection, Feature Extraction and Recognition Methods. *IET Biometrics*, Darmstadt (2012)
11. Islam, S.M.S., Davies, R., Bennamoun, M., Mian, A.S.: Efficient Detection and Recognition of 3D Ears. *International Journal of Computer Vision*, 52–73 (2011)
12. Narendira Kumar, V.K., Srinivasan, B.: Ear Biometrics in Human Identification System. *Information Technology and Computer Science*, 41–47 (2012)
13. Wagner, P.: Fisherfaces (January 13, 2013), <http://www.bytefish.de/blog/fisherfaces/>
14. Xin Dong, Y.G.: 3D Ear Recognition Using SIFT Keypoint Matching. *Journal of Theoretical and Applied Information Technology* (2013)
15. Yuan, L., Mu, Z., Xu, Z.: Using Ear Biometrics for Personal Recognition. In: *Advances in Biometric Person Authentication*, pp. 221–228 (2005)