

# Human Identification using Low Resolution Outer Ear Images: A Multiple Connectionist Framework

*Belén Moreno*

*Ángel Sánchez*

*José F. Vález*

*E.S.C.E.T*

*E.S.C.E.T.*

*Facultad de Informática*

Universidad Rey Juan Carlos  
28933 Móstoles (Madrid), Spain

Universidad Rey Juan Carlos  
28933 Móstoles (Madrid), Spain

Universidad Politécnica de Madrid  
28660 Boadilla (Madrid), Spain

## Abstract

In this paper we investigate the use of outer ear images for human identification. From the point of view of image processing, ears offer several advantages over complete faces: reduced spatial resolution, a more uniform distribution of colour, and less variability with expressions and orientation of the face. These advantages together with its identification richness, make ear images appropriate to be used as data for a connectionist system. A new multiple identification method, which combines the results obtained by several neural classifiers using, respectively, feature outer ear points, information obtained from ear shape and wrinkles, and macrofeatures extracted by compression network, is presented. Experimental results yields higher identification rates as well as a more robust framework using this approach as a component of a more general face identification system.

**Keywords:** neural networks, identification, outer ear images, classifier combination.

---

## 1. INTRODUCTION.

Biometrics use fingerprints, iris scanning or complete face images to recognize people. Due to its passive and non-intrusive nature for verifying personal identity, front and side (or profile) face images have been widely used. Many researchers from academia and industry have tackled this problem from fields like signal and image processing, computer science, neural networks, psychology or pattern recognition [1]. The interest for face images is justified by the extraordinarily developed capacity that

humans show to recognize and identify an individual by means of an image of his/her face. However, this does not necessarily means that the face is the part of the human body that contains the largest quantity of discriminant data; two outstanding counterexamples are fingerprints followed by outer ear images. In this paper, we propose the use of outer ear or auricle images to identify a person by means of three different classification techniques. Then, the respective outputs produced by each classifier are combined using a multiple classification scheme. The rest of the paper is organized as follows. Section 2 shows the advantages of ear images for personal identification. Section 3 describes the identification techniques. Section 4 presents the experiments and the recognition rates provided by each individual approach, and by the combination of classifiers. Section 5 concludes the paper.

## 2. USE OF OUTER EAR IMAGES FOR HUMAN IDENTIFICATION.

Face recognition and identification are tasks that can be effortlessly performed by people in a wide range of conditions: different lighting conditions, different positions of the person with respect to the observer or the presence of facial expressions. However, face identification and specially the automatic location of a face in a complex scene, are complex computer vision problems [7] [10].

Both holistic and feature information are crucial for the perception and recognition of faces [1]. An average person may not get used to employ an isolated feature to identify an

individual, unless this feature is dominant. It has been shown that outer ear images have more identification richness than the face or any other part of the human body - except for the own fingerprints- and they do not change significantly from the moment in which the subject reaches adult age. Structural countless variations on ears (with respect to the helix, lobe, cavities and wrinkles), as well as its shape and size, have helped in the absence of fingerprints to both anatomists and the police in identification tasks[4]. For the point of view of image processing, ears have some advantages over faces:

- Their surface is smaller (approximately 1/25 to 1/20 of that of the face), which allows working with reduced spatial resolution images.
- They have a more uniform distribution of colour, so that almost all information is conserved when converting the original image to grey scales (i.e. with faces, eye colour is lost).
- A face can change its appearance with the expression of the subject; the shape and appearance of an ear are fixed.

However, ears can be totally or partially occluded by hair or earrings; although occlusion can happen with faces too, due to beard, glasses or hairstyle. Good surveys to face processing problems and techniques can be found in the literature [1] [11].

### **3. APPROACHES FOR HUMAN IDENTIFICATION USING OUTER EAR IMAGES.**

The problem of recognition and/or identification a human face from a general view remains largely unsolved, because transformations such as position, orientation, scale, and illumination cause the face's appearance to vary substantially.

Two main approaches have been mainly used for this problem in the literature: one is based in the detection and analysis of facial features (i.e. eye distance, chin's angle,...) [10] ; the other approach avoids feature extraction and processes faces as general images using appropriate tools (i.e. neural networks, principal component analysis, histogram analysis,...) [3].

In this work, the above dichotomy has been considered. Two classifiers corresponding to the first approach have been implemented: the first one is based in the extraction of ear feature points, and the second one considers the morphology of the outer ear. Another classifier, corresponding to the second approach, which uses a compression network [11] is also developed. All classifiers produce their feature vectors that are the inputs of a respective perceptron. Finally, the predictions of the three constituent classifiers are combined by defining a multiple classification technique [6]. The following subsections give a brief description of the mentioned classifiers.

#### **3.1. Identification using feature points of ear.**

Fig. 1 shows the location of considered feature points defining the *biometric vector*. This vector, which is previously normalized to be invariant with respect to translations and scales, is the input for a perceptron. The outputs of the classifier correspond to each one of the individuals to be identified.

We use seven known feature points of outer ear, which are clearly described in anatomy books and police reports [4]. These points are automatically extracted from the profile of an ear image using an heuristic procedure. This profile is obtained as follows. First, a pair Sobel filters are applied over the original image in the orthogonal directions (horizontal and vertical), then the resulting images are respectively thresholded, and finally added to get the ear profile.

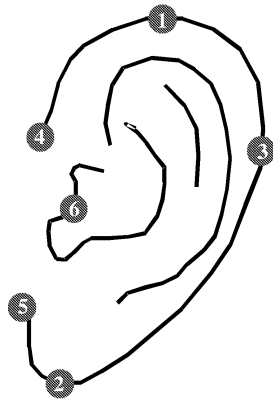


Figure 1.- Points.

### 3.2. Identification using the morphology of ear.

As in the previous identification technique, a vector is automatically obtained from the profile of an ear image. This *morphology vector* captures the shape and wrinkles structure that is specific for each individual ear. The vector determines the input for a perceptron, which can classify ear images after a training process [2].

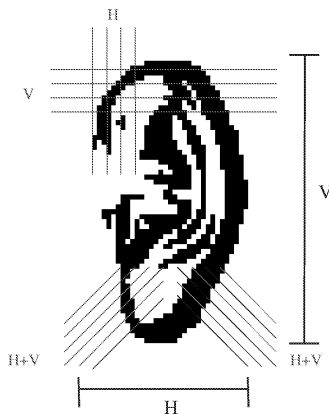


Figure 2.- Cuts.

The construction of morphology vector is as follows. A profile ear image is obtained as in the previous technique. Then a number of  $h$  horizontal cuts,  $v$  vertical cuts, and  $2(h+v)$  diagonal cuts are performed over the profile for an  $h \times v$  size image, as showed in Fig. 2. The number of intersections among the ear profile and the different defined cuts determines a vector that is then normalized to obtain the

morphology vector for each ear image.

### 3.3. Identification using compression network.

The technique uses neural networks in two stages. A first network, called compression network, is trained autoassociatively on the original ear image to extract its statistically salient properties of the image data or macrofeatures. This vector which is an intermediate codified representation of the original image is the *compression vector*. It constitutes the input to a single perceptron that performs the identification task itself (each one of its outputs corresponds to each one of the individuals to be identified). Fleming and Cottrell [3] have initially proposed a similar approach for face recognition.

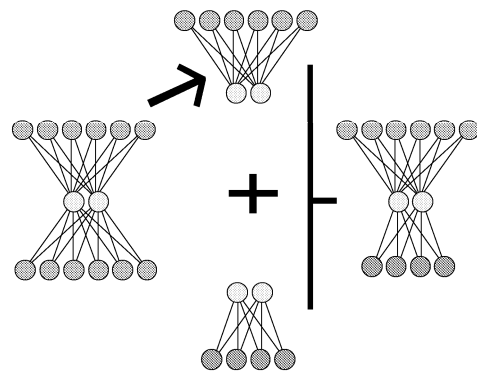


Figure 3.- Building a compression Network.

Compression networks [11], which are trained as autoassociative memories, allow the coding of neural patterns in a small dimensional subspace by extracting salient features (see Fig. 3). It has been demonstrated that a compression network with  $h$  hidden units can span the space of the  $h$  first eigenvectors (with bigger eigenvalues) of the covariance matrix corresponding to the input image.

### 3.4. Combining the predictions of previous identification techniques.

One of the key ideas for obtaining a more robust identification system with good

performances is the use of several complementary techniques to perform a given identification task. The objective of multiple classifiers is to benefit from the results provided by each classifier in the hope that a combination can be found which will yield a final recognition/identification rate superior to those of each constituent [2] [6]. Multiple classifiers require a necessary uniformization in their inputs (outputs of the component classifiers) in order to make possible the combination; this could mean a decrease in the execution time of the algorithm.

Many combination techniques have been proposed recently. Most of them are based on ideas like voting, Bayesian combination, linear weighting, or neural network techniques. A simple categorization of multiple classifiers can be established according to the kind of output provided by the constituent classifiers [2]:

- *Deterministic classifiers* are most simple ones. They provide an output, which corresponds to the class being identified, for each given input. In this work, we have used the *Majority Vote*.
- *Rank classifiers* provide for each input pattern an output, which is an ordered list of classes, such that a beginning position of a class into the list

determines that this class is more probable. As a representative technique in this group we used the *Borda Count*.

- *Probabilistic classifiers* provide for each input pattern an output vector with a number of components equal to the possible classes, and for each position (or class) they assign the probability corresponding to that class. In this work, we used the *Weighting Bayesian (WB)* method as representative of this category.

Many other combination methods of interest can be found in the literature [2][6]. Some of the most popular are the *Dempster-Shafer* or the *Behavioural Knowledge Space* methods. These methods require a bigger sample of patterns belonging to each class, which is not available in our actual ear's database. The considered multiple classifiers do not have this restriction.

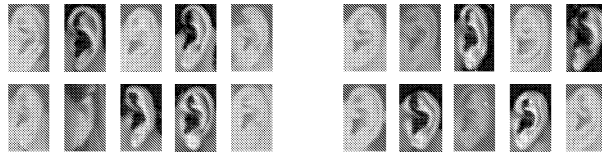
## 4. EXPERIMENTAL RESULTS.

### 4.1. Acquisition and preprocessing of outer ear images.

For the experiments 2 set was acquired (see Fig. 4 and 5). The first set, composed of 168 images, correspond to 6 photos for each one of 28 distinct individuals. Images were taken in several different sessions, and individuals were invited to change his/her expression and face orientation. Ears were manually extracted



**Figure 4.-** Data set, composed of 168 images, correspond to 6 photos for each one of 28 distinct individuals.



**Figure 5.-** Set of strange individuals.

from the profile images, stored as PGM files and slightly normalized to get an approximate vertical image. So all images have similar brightness conditions and height/width ratio (in our case, this ratio is 1.6). The second set, composed of 20 images, correspond to 20 different individuals, distinct to the individuals of the first set.

The first set of images is organized in three subsets [8]: a training set (TRS) (three images per individual), a validation set (VS) (one image), and the test set (TS) (two images). The second set was used entirely for test (SS). It was expected that all the patterns of this second set were regrests, because it corresponds to individuals distinct to the individuals that the system is trained to identify.

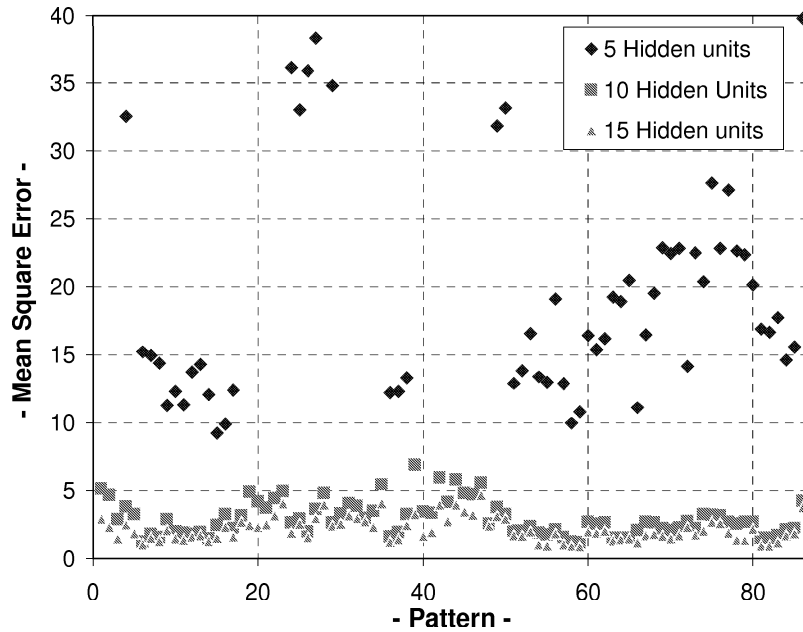
#### 4.2. Experiments.

A number of experiments were performed to evaluate the proposed framework. Results corresponding to considered identification approaches are resumed following the same order of Section 3.

- In relation with the classification technique based on *feature points*, 90% of ear points were correctly found. Different neural classifiers were trained, and a 9-9-28 (9 input, 9 hidden and 28 output units) MLP network was considered appropriate. We used QuipProp learning algorithm for the adjustment of weights, with a learning rate parameter  $\mu=0.05$ , and have trained the network during 7000 iterations to

achieve the optimal generalization with the TRS and the VS sets.

- For the *ear morphology approach*, a neural classifier was defined with a number of input units equal to the size of the morphology vector, a variable number of hidden units, and a number of outputs equal to the possible classes (number of individuals). Experiments showed that optimal generalization point was achieved using 35 hidden units, for a network trained with QuickProp with  $\mu=0.1$ , after 140000 iterations with the TRS and the VS sets.
- For the *compression network* approach, different number of hidden units were tested this network. As it is shown in Fig. 6, the *Mean Square Error* (MSE) decreased as the number of units increased. We found that a number of hidden units superior to 15 was unnecessary due to MSE did not decrease. QuickProp learning algorithm was used with  $\mu=0.001$ , and initial weight set had small values (0.001 in magnitude) to avoid overflow errors. Finally a 15-28 single perceptron was trained during 10000 iterations to achieve optimal generalization with the TRS and the VS sets.
- The combination of predictions corresponding to previous identification techniques produced the best results using the WB method. This is because this combination technique considers the confidence in each constituent classifier (in relation with its identification rate).



**Figure 6.** MSE corresponding to the compression network using a variable number of hidden units for the different training patterns.

Table I shows experimental identification rates produced by each implemented combination scheme (rows 4-7) together with the results produced by the individual classifiers (rows 1-3), with and without rejection thresholds [2]. These results correspond to the TS set.

It is important to classify correctly as many patterns as possible. But also is important, in order to avoid wrong classifications, to regret a pattern when the evidence were insufficient. From this

viewpoint, when a pattern is presented to the system, there are 3 possibilities: to accept the pattern getting it rights, to accept the pattern making a mistake, and to regret the classification. Dimauro [2] presented 2 methods to regret a classification:

- The most probable class is not classified with sufficient evidence. A threshold in the output units in the neural network can control this.
- The most probable class is not sufficient

|                               | Results without rejection thresholds |       | Results with optimal rejection thresholds |           |
|-------------------------------|--------------------------------------|-------|---|-----------|
|                               | Identification                       | Error | Identification                            | Rejection |
| Feature points of ear         | 43%                                  | 57%   | 3%  | 97%       |
| Morphology of ear             | 83%                                  | 17%   | 48%                                       | 52%       |
| Compression network           | 93%                                  | 7%    | 79%                                       | 21%       |
| Majority Vote Combination     | -                                    | -     | 34%                                       | 66%       |
| Borda Combination             | 84%                                  | 16%   | 34%                                       | 66%       |
| Bayesian Combination          | 84%                                  | 16%   | 78%                                       | 22%       |
| Weighted Bayesian Combination | 93%                                  | 7%    | 84%                                       | 16%       |

**Table I.** Summary of experimental results provided by the considered classification techniques. The first column does not consider rejection thresholds, and in the second column rejection thresholds have been chosen to eliminate the error during the identification process. Note that in the Majority Vote Combination approach makes no sense to have results without thresholds since at least two of the classifiers must coincide in its vote.

far from the second most probable class.

Using the WB method, and applying a threshold of 0,4 of distance between the first and the second more probable classes, we get a 81% of identify rate and a 0% of error over the TS set. Also we get a 100% of regret over the SS set.

## 5. CONCLUSIONS AND FUTURE RESEARCH.

Our experimental results have demonstrated that best identification results, without considering rejection thresholds, have been obtained using a compression network (93%). The combination of classifiers do not increase in general the identification rates since the classifiers are not independent at all. But taking into account rejection thresholds, these combination techniques have increased the identification rates without errors (84% achieved by the Weighting Bayesian method versus 79% of compression network).

There a number of directions for future work. The main limitation of the current system is that the size of the learning and test corpuses is small. Actually, we are extending our ear database since the difficulty of finding multiple profile images per individual in existing facial databases [5]. We are also working in the integration of this framework into another more general identification system using complete profile images. In this system, different

resolution images (i.e. higher-resolution ear and/or nose images combined with lower-resolution complete profile images) are constructed to be inputs of classification schemes.

## 6. REFERENCES.

- [1] R. Chellappa et al. "Human and Machine Recognition". Proc. of the IEEE (1995), vol. 83, pp. 705-740.
- [2] G. Dimauro, S. Impedovo. "Automatic bankcheck processing: A new engineered system". Intl Journal on Pattern Recognition and Artificial Intelligence (1997), vol. 11, pp. 467-504.
- [3] M.K. Fleming, G.W. Cottrell. "Categorization of faces using unsupervised feature extraction". Proc. Intl. J. Conf. On Neural Networks (1990), vol. II, pp. 65-70.
- [4] M. García-Pindado et al. "Estudios de Policía Científica: Identificación". División de Formación y Perfeccionamiento de la Dirección General de la Policía, 1992. (In Spanish)
- [5] S. Gutta et al. "Benchmark studies on face recognition". Proc. Int. Workshop on Automatic Face and Gesture Recognition (1995), pp. 227-231.
- [6] D.J. Hand. "Construction and Assessment of Classification Rules". John Wiley & Sons, 1997.
- [7] S. Lawrence, C.L. Gilles, A.C. Tsoi, A. Back. "Face Recognition: A Convolutional Approach". IEEE Trans. on Neural Networks and Pattern Recognition (1997), vol. 8, pp. 98-113.
- [8] C. G. Looney. "Pattern Recognition using Neural Networks: Theory and Algorithms for Engineers and Scientists". Oxford University Press, 1997.
- [9] O. Lutz, A. Dengel. "A comparison of neural net simulators". IEEE Expert (1993), vol. 8, pp. 43-51.
- [10] A. Samal, P.A. Iyengar. "Automatic recognition and analysis of human faces and facial expressions: A survey". Pattern Recognition (1992), vol. 25, pp. 65-77.
- [11] D. Valentin, H. Abdi, A.J. O'Toole, G.W. Cottrell. "Connectionist models of face processing: A survey". Pattern Recognition (1994), vol. 27, pp. 1209-1230.