IADIS International Journal on Computer Science and Information Systems
Vol. 8, No. 1, pp. 31-46
ISSN: 1646-3692

ROBUST EAR DETECTION FOR BIOMETRIC VERIFICATION

José F. Vélez. *Depto. Ciencias de la Computación, Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), SPAIN.*

Ángel Sánchez. *Depto. Ciencias de la Computación, Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), SPAIN.*

Belén Moreno. *Depto. Ciencias de la Computación, Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), SPAIN.*

Shamik Sural. *School of Information Technology, Indian Institute of Technology, Kharagpur 721302, India.*

ABSTRACT

Ear biometric recognition has received increasing attention in recent years. However, not so much work has been done on the ear verification problem. Automatic ear detection (or segmentation) from facial profile images becomes an essential preprocessing stage with high impact on the subsequent recognition/verification tasks. This paper presents a new ear detection method based on the use of circular Hough transform and some anthropometric proportions to detect the ear region accurately. After detection, the extracted contours of the segmented ear region are used to verify the identity of an individual by adjusting a fuzzy snake model on it. The proposed ear detection and verification methods were successfully tested with images from three different databases presenting different variations to evaluate the robustness of this approach.

KEYWORDS

Biometrics; verification; ear detection; active contours; fuzzy model; control access system.

1. INTRODUCTION

Biometrics is considered as one of the most promising solutions for the development of secure systems. Due to the many practical applications of this technology, there is currently an increasing demand of biometric applications in the industry. According to the experts of the International Biometrics Group (IBG) in their "Biometrics Market and Industry Report 2009-2014" [BMIR, 2009], there is a foreseeable increase in the annual revenues of the world biometrics industry by more than 270%. Among the different physical and behavioral human characteristics that can be used as biometric traits, some of the most common ones are fingerprints, facial images, voice patterns, iris and handwritten signatures, among others.

Computer biometrics is an evolving research area related to automatic security systems and relying on specific modes of uniquely recognizing or authenticating individuals. It is based on one or more intrinsic physical or behavioral characteristics [Bolle et al., 2004] [Li and Jain, 2009]. Two main types of methods for authentication people are considered in biometrics: identification and verification. Identification is based on comparing biometric templates of an individual to the corresponding ones of enrolled individuals in a database (i.e. it is an 1:N matching problem). On the other side, verification only performs just one comparison to determine the degree of similarity between one test template and one reference biometric template to determine if both correspond to the same person (i.e. it is an 1:1 matching problem).

Ear biometrics is an emerging modality which has not been exploited in terms of applications. Only those ones related to forensics (i.e. latent ear prints) have been reported [Bolle et al., 2004]. However, the advantages of using outer ear images for human identification/verification are multiple [Goudelis et al., 2008]. Ears are one of the most stable anatomical facial features and they present a rich set of distinctive points and geometrical characteristics [Burge and Burger, 2000]. The location of these elements, their direction, angles, size and relations within the ear have significant variations among humans [Abaza et al., 2011]. Moreover, these patterns have a much smaller size than the face (i.e. their area is only around 5% of the facial profile size), and are not subject to the variations produced by facial expressions or pose. The use of ears in conjunction with facial images as multimodal biometrics has been successfully presented in [Chang et al., 2003]. Another interesting advantage of ear images is that they are more secure than face images; this is because it is difficult to associate ear image with the identity of a given person. In consequence, ear image databases present a lower risk of attacks than facial databases even when the images are not encrypted [Abaza et al., 2011]. Specific difficulties in handling ear images are their partial or complete occlusions by hair, the use of earrings or the presence of hearing aids.

A potential application of ear biometrics is its usage in access control systems as an alternative to other modalities like fingerprints or hand geometry. In this application, the ear image becomes less intrusive since there is no need to touch any sensing device and the pattern can be captured from a facial profile image using a conventional camera. To analyze and compare the performance of different ear biometrics systems, public image databases are necessary. In recent years, different databases containing facial profile images (with changing pose and/or illumination conditions) or directly 2D or 3D ear images presenting some variabilities, are available. Some examples of image databases for ear biometrics application are: University of Notre Dame [UND], University of Science and Technology in Beijing [USTB] or West Pommeranian University of Technology [Frejlichowski and Tyszkiewicz, 2010] databases.

Typical stages of an automatic ear-based recognition system [Abaza et al., 2011] are the following ones: detection (or segmentation), normalization and enhancement, feature extraction and matching (recognition or verification). Ear detection consists of first extracting the position of the ear in a facial profile image. The ear region is commonly returned as a rectangular boundary that contains and adjusts to the region as best as possible. Since the other processing stages of the ear images (e.g. recognition or verification) depend on its correct detection, this stage is considered to be critical. Once segmented, the ear region can be normalized (in orientation or in size) and enhanced to make easy the further feature extraction and matching processes. Different automatic ear detection methods have been published in recent years. According to Abaza et al [Abaza et al, 2011], these methods can be classified into the following categories: computer-assisted ear segmentation that are semi-automated methods needing from user-defined landmarks indicated on the image (see as example [Alvarez et al, 2005]); template matching techniques use a template model to represent the ear in order to detect it in the image (see [Ansari and Gupta, 2007]); morphological techniques use combination of these non-linear operators to segment the ear region (e.g. [Hajsaid et al, 2008]); hybrid methods that combine two or more approaches to detect the ear region (e.g. the skin color and the use of templates like in [Prakash et al, 2009]; and Haar/Adaboost-based where a cascaded Adaboost technique based on Haar features was employed (see for example [Islam et al, 2008] or [Castrillon-Santana et al, 2011]); among other detection techniques.

Different types of features [Kumar and Rao, 2009] have been extracted from ear images like: intensity and shape features, Fourier descriptors, wavelet-based (i.e. Gabor) features or SIFT points. Some works also apply dimensionality-reduction techniques [Hurley et al., 2005] on these extracted features. Several types of recognition (or identification) techniques have also been used for the problem considered. Reported accuracy results for ear biometric identification vary depending on the methods and the databases used [Pflug and Busch, 2012] but these results are in most cases higher than 80% of success. The use of three dimensional (3D) ear patterns provides more robustness for images captured at a distance [Chen and Bhanu, 2007]. Recent survey papers on ear biometrics like [Abaza et al., 2011] and [Pflug and Busch, 2012] point out an increasing interest in this modality for person verification (or authentication).

Snakes are a type of active contour models [Blake and Isard, 1998] that have been mainly applied to contour segmentation of complex structures in medical images [Colliot et al., 2006]. In these applications, a high deformation capacity in the snake during the adjustment process is needed. However, when these models are applied to the shape matching problem (as it is the case for ear image-based biometric verification), an excessive deformation of the snake during its adjustment to the ear-contour image does not result appropriate. To avoid this undesired effect inherent in the original snake formulation, a modified model called shape-memory snake was introduced [Velez et al., 2009a]. This model "remembers" its original geometry (in particular, the relative proportions of the adjacent snake segments and the angles between these segments) during the iterative adjustment to the object boundary for shape verification. A detailed description of the model and its application to the off-line signature verification problem can be found in [Velez et al., 2009a].

This paper presents both automatic ear detection and verification methods using 2D ear images as patterns. Our detection approach combines the use of circular Hough transform with anthropometric ear proportions to robustly detect the ear region. The verification method uses fuzzy snakes to match one ear contour image with a compared snake model. Due to the shape and size variability in the ear contours of the different subjects, a system with tolerance to

imprecision would be very useful for this complex verification problem. Consequently, we have adapted and applied the fuzzy shape-memory snake framework described in [Velez et al., 2009a] that was used for handwritten signature verification. In Velez et al [Velez et al., 2009b], we demonstrated that this framework outperforms the equivalent crisp solution. People recognition based on ear biometrics is more commonly studied than the verification one [Abaza et al., 2011]. However, the verification modality presents interesting applications like those ones related with access control systems.

The rest of the paper is organized as follows. Section 2 describes the proposed ear detection method. Section 3 summarizes the ear verification approach based on fuzzy snakes. Detection and verification experiments carried out on three image datasets are presented in Section 4. Finally, Section 5 outlines the conclusions of this work.

2. EAR DETECTION

Object detection is an important task in computer vision that aims to find instances of semantic objects belonging to searched classes (such as humans, buildings, or cars) in digital images and videos. This task involves extracting the regions of pixels in the image where the searched objects are present. A widely studied detection domain in Biometrics is the detection of faces that is commonly performed using the Viola-Jones method [Viola and Jones, 2001]. Object detection has many applications in areas of computer vision (e.g. image retrieval) and it is a crucial stage for further recognition processes.

The usage of some contextual information about the searched object (e.g. its shape or color) is very useful to automatically detect it. The ear pattern is composed of several circles (some of them are concentric) that form the outer and inner borders of the helix and antihelix ear regions. This morphological feature is a characteristic from ears. The external ear contour can be represented by an ovoid [Alvarez05]. The ovoid model resembles the egg-shape and it can be obtained by the deformation of an ellipse. In practice, an algorithm is needed to estimate the ovoid parameters given an ear contour.

The main idea in our approach for ear detection is based on finding in a facial profile image the concentric circles present in the ear region. More specifically, we apply the general Hough transform to detect the center of the higher number of concentric circles in the edge image of a facial profile. This center pixel corresponds to the upper ear region and the larger circle given by the Hough transform using this center (see Figure 1.a) is used in combination with some anthropometric measures to extract in a precise form a rectangle in the image adjusted to the ear region (see Figure 1.b). It is interesting to note that although other circular regions in the facial profile can be detected using the Hough transform, the upper ear region is the one that contains most concentric circles.

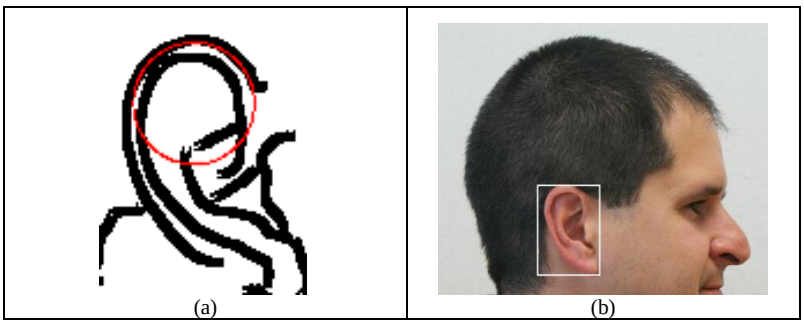

Figure 1. Ear detection method: (a) upper external ear circle found (in red) to detect the ear helix by applying the generalized Hough transform to the facial contour (for clarity we zoomed the image and only show the ear region in it) and (b) final ear region segmentation produced (white rectangle).

The stages involved in the ear detection task are shown with an example in Figure 2. Next, we explain the method used to segment the ear region from a facial profile image. It can be decomposed into three main stages: preprocessing, contour extraction and ear region extraction, respectively. Ear image preprocessing consists of converting the original color image into gray scale (since our method is based on contour detection and only the luminance information is needed), and applying a median filter for noise removal. Contour extraction consists of first applying a Canny edge detector to obtain a binary edge facial profile image which is next inverted. This resulting binary edge image is properly dilated with a 4×3 disk-shaped structuring element for further detection stages. After that, the small-sized connected components are removed to complete the contour extraction. Ear region extraction consists of finding a rectangular region in the image which better adjusts to the external ear contour. This stage is carried out by searching circles by applying the circular Hough transform to the facial edge image obtained previously. We first search the upper helix region (i.e. ear upper rim) and once this region is detected, the remaining part of the ear is found by considering some anthropometric ratios of this organ. Since the helix contour resembles the three-quarters circle show in Figure 3, then this circular arc is searched instead the complete circle. Pei and Horng [Pei and Horng, 1995] have proposed a circular arc detection method based in the Hough transform that we adapted for this task. The information of centers in the Hough accumulator matrix of circular arcs is used to detect the point that produces the highest accumulation value (see Figure 2.h) that will be considered the center of helix rim. The largest radius value associated to this center will determine the upper outer contour of the ear. Finally, we have used some parameter values explained in anthropometric studies [Farkas and Munro, 1987][Alexander and Laubach, 1968] to detect the rectangle region adjusted to the external contour of the ear. These parameters are shown in Figure 4 and their estimated values are presented in Table 1.

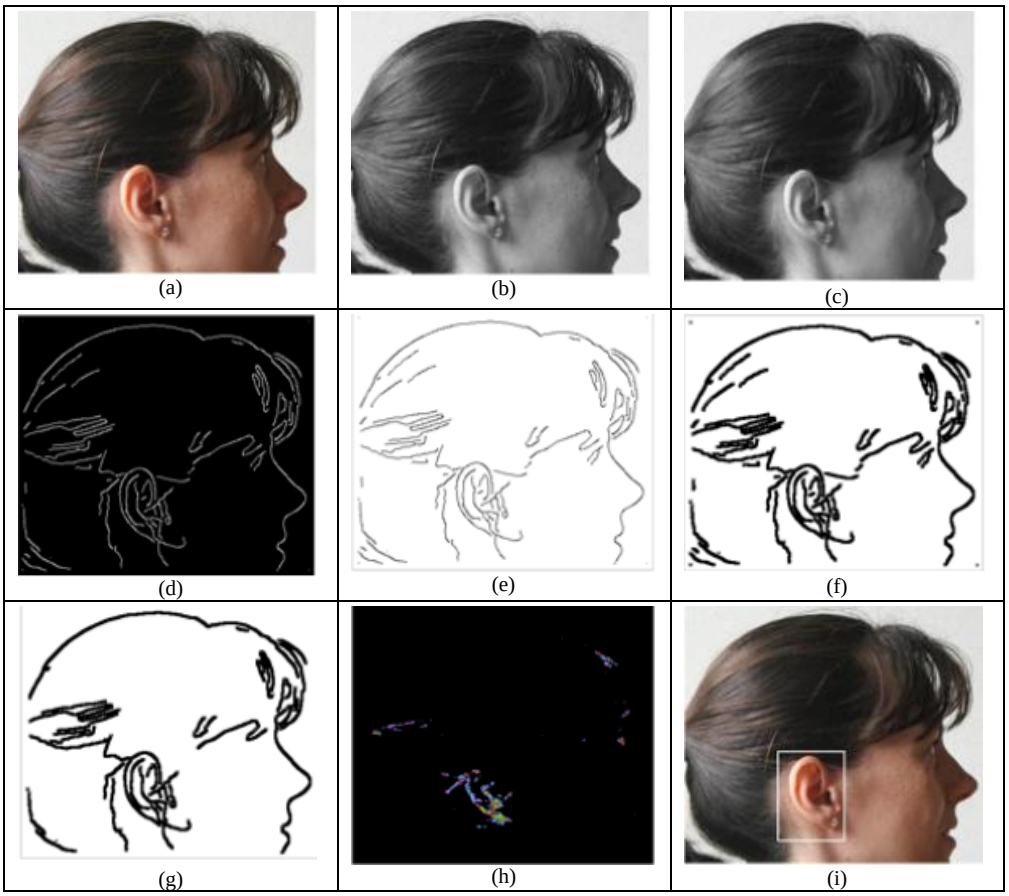

Figure 2. Stages followed to detect the ear region: (a) initial color facial profile image, (b) conversion to grayscale, (c) application of a median filter, (d) edge extraction (Canny algorithm), (e) inversion of the edge image, (f) mathematical morphology processing, (g) remove small regions, (h) application of circular Hough transform and (i) final detection of the ear region in the original image.

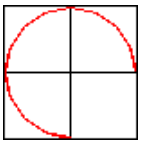

Figure 3. Circular arc (in red color) searched in the image using Hough transform.

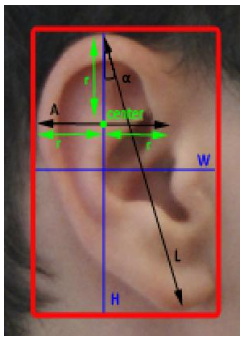

Figure 4. Anthropometric measures used to detect the ear rectangle region.

Table 1. Parameters and their corresponding values used to detect the ear region (see also Figure 4).

Parameter	Meaning	Value
r	Radius of circular arc detected	17.25 mm
L	Largest distance between ear contour points	3.89·r = 67.1 mm
A	Largest distance in helix contour	2·r = 34.5 mm
α	Inclination angle of ear (defined by *H* and *L*)	15°
H	Vertical height of the ear region	$L \cdot \cos \alpha = 3.76 \cdot r = 67.1$ mm
W	Horizontal width of the ear region	$A + H/2 \cdot \sin \alpha = 2.71 \cdot r =$ 46.74 mm

3. EAR VERIFICATION

Previous to the verification task, a feature extraction step is used that determines the ear contour which is a normalized ear edge image. The process of ear contour extraction is illustrated in Figure 5. First, the original image is converted into a gray scale (if the initial image is in colors) and this is smoothed by the application of a Gaussian filter. Next, the edges are extracted using a Canny filter, and small non-connected edges are removed. After that, the edge contour is obtained by applying an appropriate morphological dilation. Ear contour normalization in orientation and scale is performed by automatically detecting the super-auricular point (i.e. the higher-position point at the ear contour) and the sub-auricular point (i.e. the lower-position point at the ear contour). Using the segment defined by the two previous points, we normalize the ear patterns that will be used for verification purposes.

In the verification stage, we need one snake model for each class (i.e. subject in the database to be authenticated). This model is built by using only one training ear contour image per person. We produce an open polygonal line (snake) *P* composed by segments that join a variable number of equally-spaced neighbor control points, as explained in [Velez et al., 2009a]. This snake is placed over the test ear contour image such that the centers-of-gravity (COG) of both structures are made coincident. Later, the snake *P* is iteratively adjusted to the ear contour image to be verified using an energy minimization formulation similar to those presented by the authors in [Velez et al., 2009a] and [Velez et al., 2009b] that was applied to the off-line signature verification problem.

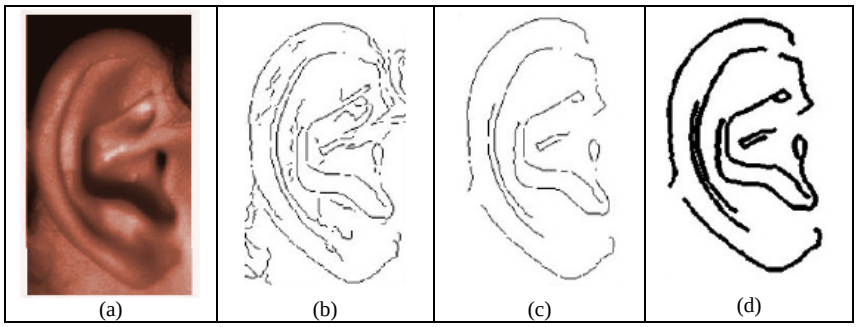

Figure 5. Ear contour image pre-processing: (a) original segmented image, (b) edge extraction, (c) removal of edge noise components and (d) template edge contour (obtained through morphological dilation from previous image).

This process is illustrated in Figure 6 with two examples. The first two images shows a successful adjustment (i.e. where the compared snake model and test ear contour are both from the same person) that will produce an acceptable result during verification. The last two images show a failed adjustment (i.e. where the compared snake model and test ear contour are from different persons) that will produce a rejection result in the verification task.

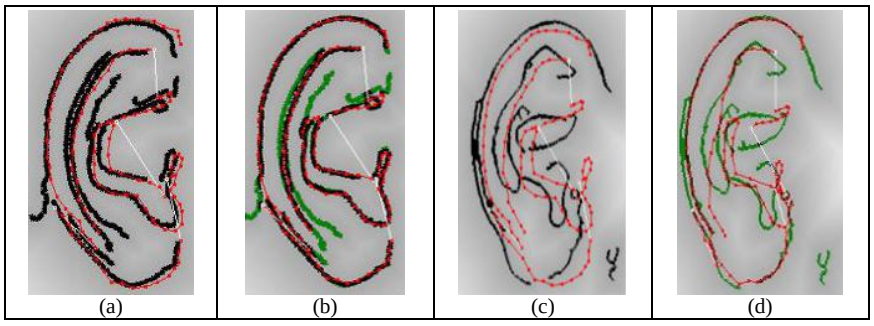

Figure 6. Two examples of ear-snake adjustments: (a)-(b) and (c)-(d), respectively: (a) initial snake placement (red dotted line) on an ear contour (solid black line) of the same individual and (b) successful adjustment result (green lines are the non-adjusted parts), (c) initial snake placement (red dotted line) on an ear contour (solid black line) of a different individual and (d)) failed adjustment result (green lines are the non-adjusted parts).

Next, we detail the specific fuzzy energy formulation for the problem considered. Let $\vec{v} = (d, \theta, p)$ be the 3-tuple vector of input variables defining the snake adjustment energy E_{snake}. At each iteration *t*, this energy results from the contributions of the snake control points in the adjustment process. The input variables are: *d* representing the distance from any control snake point to the closest one in the ear contour, θ defining the variation of the angle formed by any three adjacent snake control points with respect to its initial position, and *p* that represents the variation of the proportions of two consecutive snake segments with respect to its initial proportions.

The vector of membership functions $U\tilde{P}_{\vec{v}}$ of each input variable is:

$$U\tilde{P}_{\vec{v}} = (U\tilde{P}_d, U\tilde{P}_\theta, U\tilde{P}_p) \tag{1}$$

where:

$$U\tilde{P}_d = (\tilde{d}_o, \tilde{d}_1), U\tilde{P}_\theta = (\tilde{\theta}_0, \tilde{\theta}_1), U\tilde{P}_p = (\tilde{p}_0, \tilde{p}_1) \tag{2}$$

The subscript '0' in the fuzzy sets represents the linguistic label 'small' and the subscript '1' in the fuzzy sets represents the linguistic label 'large' (for distances, angle and proportion variations, respectively). In our ear contour verification problem, given the input vector, the proposed zero-order Takagi-Sugeno (TS) inference system has 2^3=8 fuzzy rules which define the snake adjustment energy E_{snake}:

$$\begin{aligned} R_0 &: IF\ (d\ is\ \tilde{d}_0)\ AND\ (\theta\ is\ \tilde{\theta}_0)\ AND\ (p\ is\ \tilde{p}_0)\ THEN\ E_{snake} = E_{Low} \\ R_1 &: IF\ (d\ is\ \tilde{d}_1)\ AND\ (\theta\ is\ \tilde{\theta}_0)\ AND\ (p\ is\ \tilde{p}_0)\ THEN\ E_{snake} = E_{High} \\ &\ldots. \\ R_7 &: IF\ (d\ is\ \tilde{d}_1)\ AND\ (\theta\ is\ \tilde{\theta}_1)\ AND\ (p\ is\ \tilde{p}_1)\ THEN\ E_{snake} = E_{High} \end{aligned} \tag{3}$$

As explained in [Velez et al., 2009a], the system output computation E_{snake} can be reduced to:

$$E_{snake} = \sum_{i=0}^{R-1} E_i \prod_{j=1}^{P-1} \tilde{n}_{i,j,k}(x_j) \tag{4}$$

where $\tilde{n}_{i,j,k}(x_j)$ represents the degree of fulfillment of the fuzzy number $\tilde{n}_{i,j,k}$ for the variable x_j. To simplify the corresponding computation, two membership functions (that form a partition of unity) are defined in each interval of every variable $x_j \in \bar{x}$:

$$\tilde{n}_{i,j,k} \in [\tilde{n}_{i,j,0}, \tilde{n}_{i,j,1}] = [\tilde{n}_{i,j,0}, 1 - \tilde{n}_{i,j,0}] \tag{5}$$

where: $0 \leq i \leq R\text{-}1$ and $0 \leq j \leq P\text{-}1$ (being *R* the number of rules and *P* the number of variables). The corresponding membership functions to these fuzzy sets are:

$$\mu_{\tilde{n}_{i,j,k}} = [\mu_{\tilde{n}_{i,j,0}}, \mu_{\tilde{n}_{i,j,1}}] \tag{6}$$

and the degree of fulfillment of the fuzzy set $\tilde{n}_{i,j,k}$ by the variable x_j is $\mu_{\tilde{n}_{i,j,k}}(x_i)$.

As all considered fuzzy numbers for each input variable belong to a partition of unity, the final snake adjustment energy E_{snake} for the zero-order TS system in (3) can be simply computed as:

$$E_{snake} = \alpha_0 E_{Low} + (1-\alpha_0)E_{High} = [\mu_{\tilde{d}_0}(d) \cdot \mu_{\tilde{\theta}_0}(\theta) \cdot \mu_{\tilde{p}_0}(p)]E_{Low} + [1 - \mu_{\tilde{d}_0}(d) \cdot \mu_{\tilde{\theta}_0}(\theta) \cdot \mu_{\tilde{p}_0}(p)]E_{High} \quad (7)$$

where α_0 is a real value in [0..1]. It is important to remark that the previous fuzzy formulation makes it possible to obtain discrete bounded snake energy values in a natural way.

After the adjustment, we use a number of iterations for adjusting the snake to the ear contour and also the percentage of coincident points between the snake and the contour as features to discriminate between accepted and rejected ear images. A 2-layer perceptron neural network (NN) was used as classifier. This network has two input neurons corresponding to the values of the two previous computed adjustment measures. The number of neurons in the hidden layer was experimentally set to 10 units and the output layer had only one neuron. If the result produced by the classifier is greater than a threshold, then the test ear image is considered genuine and it is accepted; otherwise, it is rejected.

4. EXPERIMENTS

This section describes the databases used for the experimentation and the ear detection and verification tests carried out.

4.1 Considered Databases

To test the proposed ear detection and verification methods, we used three different image databases created by the authors: grayscale, color and near infrared (NIR) databases. The grayscale database is composed of 66 images corresponding to 12 different subjects. All images were captured with a resolution of 640×480 pixels, at the same distance of the subjects with similar illumination conditions. No preprocessing on the images was applied.

The color database is composed of 60 images corresponding to 10 different subjects (i.e. 6 images per person). The images were captured with a SONY DSLR-A200 camera at an original resolution of 3,872×2,592 pixels. Different from the previous database the subjects were placed at a varying distance from the camera and the illumination conditions were different among the images. First, some of the images were manually cropped to contain only the facial region (i.e. neck region and clothes were eliminated). Next, the images were rescaled to a spatial resolution of 953×638 using a bilinear interpolation. After this preprocessing, all database images kept a similar ear resolution.

The NIR database is composed of 150 images corresponding to 25 subjects (i.e. 6 images per person) with a resolution of 480×640 pixels. All the images were taken with a JAI AD-80CL camera. An miniflood 100-LED 855nm 60º infrared illumination system was used when capturing three images of each subject (i.e. illuminated) while the remaining three ones were captured without using this system (i.e. non-illuminated or captured with natural illumination conditions). Different from the other databases, the images on this one contain a more reduced region of the facial profile (and not the complete face).All these images were first rescaled at

resolution 174×232 pixels and then an adaptive equalization was applied on each image to increase its contrast (using the adaphisteq command of MATLAB). The six images corresponding to each individual were captured under these conditions: illuminated and non-illuminated with the subject looking up, illuminated and non-illuminated with the subject looking front, illuminated and non-illuminated with the subject looking down, respectively. The goal is to study on the influence of ear rotation and illumination conditions on the detection and verification tasks on this database.

Figure 7 shows different images of the three databases used in the experiments.

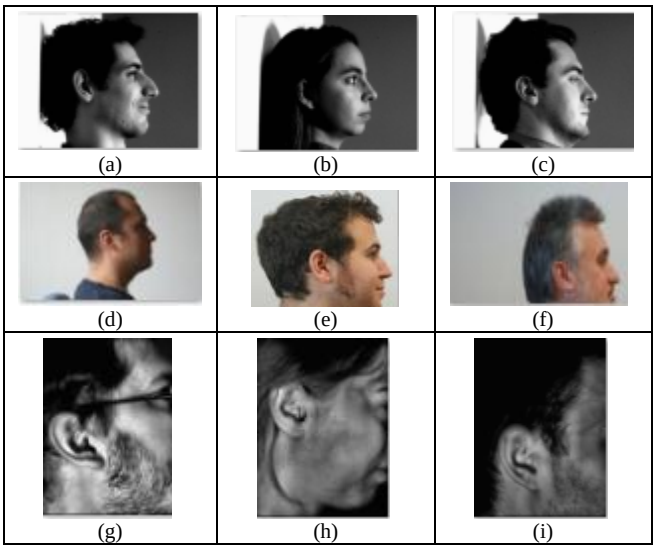

Figure 7. Different sample images (before preprocessing them) of each of the databases used in the experiments: grayscale database (top row), color database (middle row) and NIR database (bottom row).

4.2 Ear Detection Results

To evaluate the accuracy in the ear automatic detection task, the test images were manually segmented. The overlap between both rectangular segmented ear regions (i.e. automatic and manual) was measured. This measure of accuracy in the ear detection is computed as the minimal value between the percentage of the automatically extracted ear region that coincides with the manually segmented ear region, and the percentage of the manually segmented ear region that coincides with the automatically segmented region. Figure 8 shows on a sample face both automatically and manually segmented ear regions. For this example, the minimum value of these percentages is: *min*(100,74) =74%, which is an acceptable result for the automatic detection step.

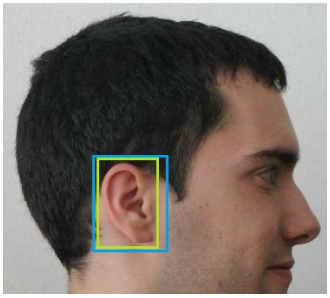

Figure 8. Example of ear detection on a facial profile image: automatic result achieved by the algorithm (blue rectangle) and manual segmented result (green rectangle).

This accuracy percentage is directly related to the distance in pixels between the centers of the circular arcs obtained through the application of the Hough transform (as shown by Fig. 4). Both the center positions can be computed for automatic and manual detections. After some experimentation, we determined that an ear region is correctly detected on a facial image when the distance between these centers is smaller than a threshold $th = r/3$, where r represents the radius value (see again Fig. 4). Using the previous threshold value, the correct detection results were computed on each of the three databases considered. Table 2 summarizes these results which correspond to the ear detection accuracy for each database.

Table 2. Accuracy of the proposed ear detection method on the databases: respective numbers (#) and percentages (%) of correct detections.

Database (# images)	Correct Detections (#)	Correct Detections (%)
Grayscale (66)	58	87.88
Color (60)	47	78.33
NIR (150)	96	64.00

Best detection results were achieved using the Grayscale database that presents more uniform conditions with respect to position, scale and illumination. Due to the higher variability of the images contained in the NIR database (i.e. rotated vs. non-rotated images, LED illuminated vs. non-illuminated images), the corresponding detection results were worse. By comparing our ear detection results with those presented in the recent survey by Abaza et al [Abaza et al, 2011] on other datasets (of different sizes), the accuracy reported varies between 79% and 100%. Our best results on the Grayscale database are comparable to the ones published in this survey. However, our results on the NIR database that presents the highest variability with respect to the image conditions, are a bit worse than those reported in the survey.

4.3 Ear Verification Results

Ear verification experiments were carried out on the three databases considered in this study. Only the correctly detected patterns were used for this purpose. After the ear detection task and the normalization of the images, we obtain the ear contour subimages (as described in Section 3) with a spatial resolution of 124×200 pixels, and these subimages were stored as .png files.

Figure 9 shows the fuzzy partitions associated to the input system variables (defining the snake adjustment energy E_{snake}), the corresponding fuzzy sets defining the labels 'small' and 'large' for each variable, and the considered values of parameters of each fuzzy set in our model images (see again Section 3). The control points of these fuzzy partitions where chosen through experimentation. A 2-layer perceptron classifier was trained during 1,000 iterations using the back-propagation algorithm. As training patterns, we used the measures corresponding to the snakes' adjustments from around 25% individuals of each database and the images of the remaining individuals were used for the tests. In this ear verification problem, the classes correspond to each of the individuals in the databases. One ear image of each person was used to create his/her ear snake model and the remaining ones for verification purposes.

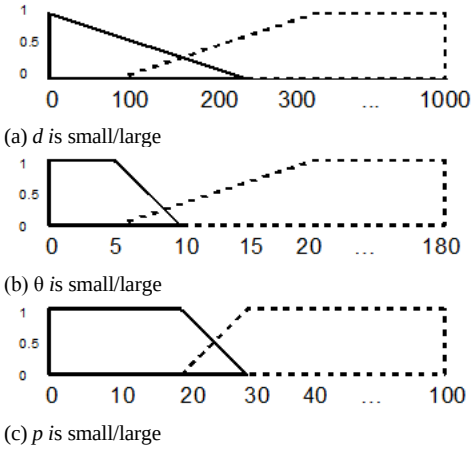

Figure 9. Fuzzy partitions corresponding to the input variables for the simplified fuzzy inference system describing the snake adjustment energy at each control point of the snake. Note that: (a) d (distance) is given in pixels, (b) θ (angle) is given in degrees, and (c) p (proportions) is a percentage. The final snake adjustment energy E_{snake} is a normalized real value between 0 and 1.

As pointed out before, the comparative verification results among databases is carried out using the correctly detected ear images. Figure 10 presents a ROC curve comparing the results achieved on three databases. Best (and similar) results are obtained using the Grayscale and the NIR databases with around 17% of Equal Error Rate (EER) value. The Color database produced the worst results with around 30% of EER value. As it also happens with detection, the contrast on the images has an important influence on the verification results. Once correctly detected the ear region, since the Grayscale and NIR images present a higher contrast

than the Color images, this makes that the corresponding ear contours are better enhanced. This makes more effective the snake adjustment process on these images.

It can also be observed from Figure 10 that the verification results on the NIR database are slightly better than the Grayscale (i.e. the Area Under Curve-AUC- on the NIR database is a bit smaller). By analyzing more in detail the ROC curve of the NIR dataset, an interesting operating point in this curve corresponds to: 5% of FAR (False Acceptance Rate) and 20% of FRR (False Rejection Rate) errors. This point makes the proposed ear-based verification system interesting for an access control application where security is the priority, since the FAR value is low.

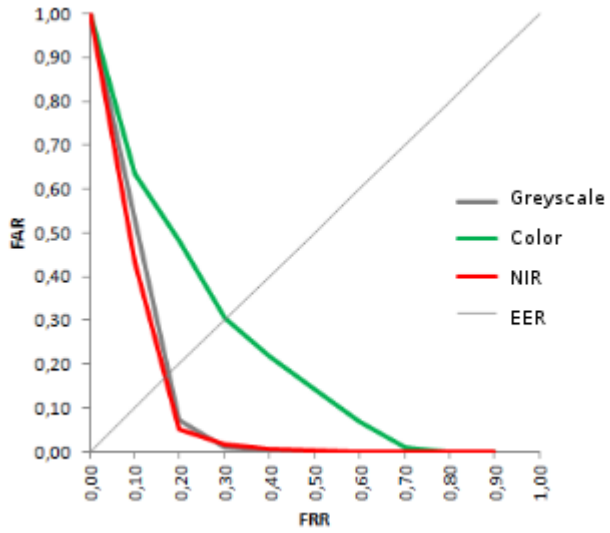

Figure 10. Comparative of achieved verification results on the three databases using their corresponding ROC curves.

5. CONCLUSION

This paper presented a robust and automatic ear detection method for facial profile images. The method has been integrated into a biometric verification system that uses outer ear images as patterns. Robust ear detection requires acceptable segmentation accuracy under different types of variabilities (i.e. illumination, rotations and scale differences). The proposed method combines the application of circular Hough transform with some anthropometric information of the ear to extract a well-adjusted rectangular region containing the ear. Next, the segmented region is used to extract the ear contour. The verification method is based on the fuzzy adjustment of one active contour model (corresponding to a given ear model) to one ear contour (corresponding to a test pattern). Some measures extracted from the snake adjustment are used as features to decide whether or not to accept an individual based on his/her ear template. The complete system has been tested on three different facial profile databases created by the authors (i.e. referred in the paper as Grayscale, Color and NIR databases, respectively). Achieved ear detection results are comparable with other presented in the

literature using different datasets. The best ear verification results were achieved using the NIR database which has images with a higher contrast making it easier to detect the ear contour by the proposed fuzzy snakes.

As future work, we will first try to test our ear detection and verification methods using different standard datasets as those referenced in [Abaza et al, 2011]. Moreover, it is interesting to determine the robustness of our algorithms using images that include the partial occlusion of this human organ (i.e. due to hair or ear-rings).

ACKNOWLEDGEMENT

This research has been partially supported by the Spanish research project TIN2011-29827-C02-01.

REFERENCES.

Abaza, A., et al, 2011. A Survey on Ear Biometrics, *ACM Computing Surveys*, Vol. 45, No. 2.

Alexander, M. and Laubach, L.L., 1968. *Anthropometry of the human ear* (*A photogrammetric study of USAF flight personnel*). Aerospace Medical Research Laboratories, Technical Report AMRL-TR-67-203, Ohio, USA.

Alvarez, l., González, E. And Mazorra, L., 2005. Fitting ear contour using an ovoid model, *Proc. 39th Annual International Carnahan Conference on Security Technology (CCST '05)*, pp. 145 – 148.

Ansari, S. and Gupta, P., 2007. Localization of ear using outer helix curve of the ear. *Proc. of the International Conference on Computing: Theory and Applications*, pp. 688–692.

Blake, A. and Isard, M., 1998. *Active Contours*, Springer-Verlag, New-York, USA.

BMIR, 2009. Biometrics Market and Industry Report 2009-2014. URL: http://www.ibgweb.com/products/reports/bmir-2009-2014. (Accesed: 16-09-12)

Bolle, R.M. et al, 2004. *Guide to Biometrics*, Springer, New York, USA.

Burge, M. and Burger, W., 2000. Ear biometrics in computer vision, *Proc. of the International Conference on pattern Recognition (ICPR'00)*, Barcelona, Spain, Vol. II, pp. 2822–2826.

Castrillón-Santana, M., Lorenzo-Navarro, J. and Hernández-Sosa, D., 2011. A Study on Ear Detection and Its Applications to Face Detection, LNCS. Springer, Vol. 7023, pp 313-322.

Chang, K. et al, 2003. Comparison and combination of ear and face images in appearance-based biometrics, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), pp. 1160–1165.

Chen, H. and Bhanu, B., 2007. Human ear recognition in 3D, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4), pp. 718–737.

Colliot, O., Camara, O. and Bloch, I., 2006. Integration of fuzzy spatial relations in deformable models: Application to brain MRI segmentation, *Pattern Recognition* 39, pp. 1401-1414.

Farkas, L.G. and Munro, I.M., 1987. *Anthropometric Facial Proportions in Medicine*. Charles C Thomas, Springfield, Illinois, USA.

Frejlichowski, D. and Tyszkiewicz, N., 2010. "The West Pomeranian University of Technology Ear Database – A Tool for Testing Biometric Algorithms", In: A. Campilho and M. Kamel (Eds.): ICIAR 2010, Part II, Lecture Notes in Computer Science, vol. 6112, pp. 227–234. URL: http://ksm.wi.zut.edu.pl/wputedb

Goudelis, G., Tefas, A. and Pitas, I., 2008. Emerging biometric modalities: a survey, *J Multimodal User Interfaces* 2, pp. 217–235.

Hajsaid, E., Abaza, A. and Ammar, H., 2008. Ear segmentation in color facial images using mathematical morphology. *Proc. of the 6th Biometric Consortium Conference* (BCC).

Hurley, D.J., Nixon, M.S. and Carter, J.N., 2005. Force field energy functionals for ear biometrics, *Computer Vision and Image Understanding* 98(3), pp. 491–512.

International Biometrics Group (IBG) Report 2009-2014. URL: http://www.ibgweb.com/products/reports/bmir-2009-2014.

Islam, S.M.S., Bennamoun, M. and Davies, R., 2008. Fast and Fully Automatic Ear Detection Using Cascaded AdaBoost, *Proc. IEEE Workshop on Applications of Computer Vision* (WACV '08), pp. 1-6.

Kumar, R., Selvam, P. and Rao, K.N., 2009. Pattern Extraction Methods for Ear Biometrics: A Survey", *Proc. World Congress on Nature & Biologically Inspired Computing (NaBIC'09)*, Coimbatore, India, pp. 1657-1660.

Li, S.Z. and Jain, A.K. (editors), 2009. *Encyclopedia of Biometrics*, Springer, New York, USA.

Pei, S.-C. and Horng, J.-H., 1995. Circular arc detection based on Hough transform, *Pattern Recognition Letters* 6, pp. 615-625.

Pflug, A. and Busch, C., 2012. Ear Biometrics: A Survey of Detection, Feature Extraction and Recognition Methods, *IET Biometrics* 1(2), pp. 114-129.

Prakash, S., Jayaraman, U. and Gupta, P., 2009. A Skin-Color and Template Based Technique for Automatic Ear Detection, *Proc. Intl. Conf. on Advances in Pattern Recognition* (ICAPR'09), pp. 213-216.

USTB Databases, University of Science and Technology in Beijing (China), Ear Recognition Laboratory.

URL: http://www1.ustb.edu.cn/resb/en/aboutus/lab.htm

UND Databases, University of Notre Dame (USA), Computer Vision Research Laboratory. URL: http://cse.nd.edu/~cvrl/CVRL/CVRL_Home_Page.html

Vélez, J. Sánchez, A. and Fernández, F., 2009. Improved Fuzzy Snakes Applied to Biometric Verification Problems, *Proc. Int. Conf. Intelligent Systems Design and Applications (ISDA '09)*, Pisa, Italy, pp. 158-163.

Velez, J., et al, 2009b. Fuzzy shape-memory snakes for the automatic off-line signature verification problem, *Fuzzy Sets and Systems* 160, pp. 182-197.

Viola P. And Jones, M., 2001. Rapid Object Detection using a Boosted Cascade of Simple Features, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR'01), v.1, pp. 511-518.