Ear Biometrics Based on Geometrical Method of Feature Extraction

Michał Choraś

Institute of Telecommunication, University of Technology and Agriculture,
ul. Prof. Kaliskiego 7, 85-796, Bydgoszcz, Poland.
`chorasm@mail.atr.bydgoszcz.pl`

Abstract. Biometrics identification methods proved to be very efficient, more natural and easy for users than traditional methods of human identification. In fact, only biometrics methods truly identify humans, not keys and cards they posses or passwords they should remember. The future of biometrics leads to passive physiological methods based on images of such parts of human body as face and ear. The article introduces to ear biometrics and presents its advantages over face biometrics in passive human identification systems. Then the geometrical method of feature extraction from human ear images in order to perform human identification is presented. The proposed method is invariant to rotation, translation and scaling due to coordinates normalization and placing the major reference point in the centroid. The feature extraction algorithm consists of two steps, so that in the process of classification two feature vectors for each ear image are used.

1   Introduction

There are many known methods of human identification based on image analysis. In general, those biometrics methods can be divided into behavioural and physiological regarding the source of data, and can be divided into passive and invasive biometrics, regarding the way the data is acquired.

The first class is based on the behavioural features of human actions and it identifies people by how they perform something. The most popular of such methods is voice verification. Other methods are basically based on the dynamics of specific actions like making the signature, typing on the keyboard and simply moving or walking. Those methods are not that natural and they require users to take part in the process of identification by repeating specific actions, every time they are examined. Physiological (anatomical) biometrics methods are based on the physiological features of humans thus they measure and compare features of specific parts of human body in the process of identification. So far the main interest is in the head and the hand with face, eye and fingerprint features being the most important discriminants of human identity. The major advantage of physiological biometrics is that it is passive and the implemented systems work only with the acquired images of specific body parts. All the user has to do is to place his face or ear in front of the camera or alternatively touch the sensor with his fingers and wait for the identification process to take place. But some systems can verify the identity of

humans even without their cooperation and knowledge, which is actually the future of biometrics. Such methods include well-examined face recognition but one of the most interesting novel approaches to human passive identification is the use of ear as the source of data. The article presents some advantages of ear biometrics over face in human identification.

Human ears have been used as major feature in the forensic science for many years. Recently so called earprints, found on the crime scene, have been used as a proof in over few hundreds cases in the Netherlands and the United States [9].

Fig. 1. Some examples of "easy" ear images from our database

There are many advantages of using ear in comparison with the face. Firstly, ear does not change during human life, and face changes more significantly with age than any other part of human body. Face can also change due to cosmetics, facial hair and hair styling. Secondly, face changes due to emotions and expresses different states of mind like sadness, happiness, fear or surprise. In contrast, ear features are fixed and unchangeable. Moreover, the colour distribution is more uniform in ear than in human face, iris or retina. Thanks to that fact, not much information is lost while working with the greyscale or binarized images, as we do in our method. Ear is also smaller than face, which means that it is possible to work faster and more efficiently with the images with the lower resolution. Ear images cannot be disturbed by glasses, beard nor make-up. However, occlusion by hair or earrings is possible. Some examples of "easy" and "difficult" ear images from our database are shown in Figs 1 and 2, respectively.

   The first, manual method, used by Iannarelli in the research in which he examined over 10000 ears and proved their uniqueness, was based on measuring the distances between specific points of the ear [11]. The major problem in ear identification systems is discovering automated method to extract those specific, key points. Another well-known method by Burge and Burger [4] is based on building neighbourhood graph from Voronoi diagrams of the detected edges. Hurley et al. [10] introduced a method based on energy features of the image. They proposed to perform force field transformation in order to find energy lines, wells and channels, but their feature vector consisted of only two parameters. Another method used by Victor at al. [17], in the experiment comparing ear and face properties in order to successfully identify humans in various conditions, was based on PCA. Their work proved that ear images are a very suitable source of data for identification and their results for ear images were not significantly different than for face images. The method, however, was not fully automated, since the reference points had to be manually inserted into images.
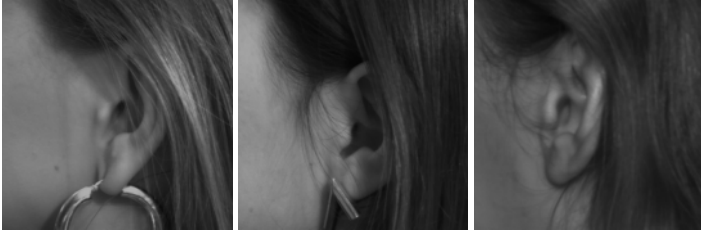
Fig. 2. Some examples of "difficult" ear images from our database

We propose a straightforward method to extract features needed to classification. Our method represents the geometrical approach, but it is fully automated and no manual operations are needed.

Our method is divided into image normalization, contour extraction (edge detection), calculation of the centroid, coordinates normalization and 2 steps of geometrical feature extraction, as described in the next section. We treat the centroid as the specific point in our method, even though it is not a specific point within the ear topology. Our method consists of the following steps (Fig. 3.):

−   contour detection,
−   binarization,
−   coordinates normalization,
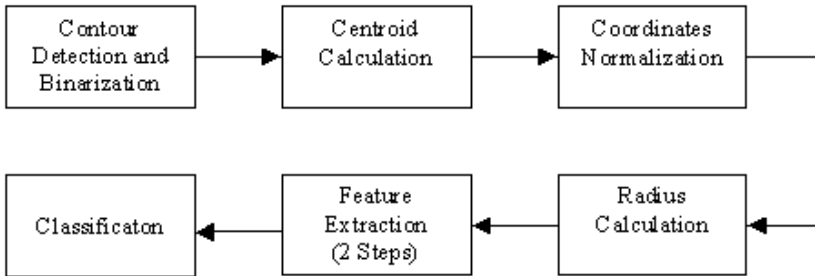−   feature extraction (2 steps),
−   classification.

Fig. 3. The block diagram of our method

2   Contour Detection

First we perform the edge detection step. In our case it is a crucial operation, since it is obvious that lines are the most prominent features that could be obtained from the ear image, and our goal is to detect major outer and inner curves of the earlobe. We tested many known methods as Canny operator, Sobel filters and CWT (Complex Wavelets) but we propose another method, which proved to be most convenient in our experiments.

We propose to use the local method which examines illumination changes within the chosen window $n \times n$. We usually use 3×3 window and we divided the image into many overlapping regions of that size.

For each of those regions we calculated mean μ and standard deviation σ of the pixel intensity values in 8-pixel neighbourhood.

$$\mu = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} I(i, j) \tag{1}$$

$$\sigma = \sqrt{\frac{1}{(n^2 - 1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(I(i, j) - \mu\right)^2} \tag{2}$$

Then we perform decision if the centre pixel of the examined region belongs to the line or to the background. For the maximum value of the pixel intensity I_H, and the minimum value of the pixel intensity in the region I_L, we calculate the difference $S(i, j)$ such as:

$$S(i, j) = I_H - I_L \tag{3}$$

and we compare it to certain threshold value. Even though thresholding is one of the basic operations of image processing, there is always the major problem in selecting appropriate threshold value.

We propose the usage of mean and standard deviation of pixel intensities in calculation of the threshold value $T(i, j)$ used in contour detection as given in equation 4:

$$T(i, j) = \mu - k\sigma \tag{4}$$

where k is a certain value.
Then the rule for the contour detection is:

$$g(i, j) = \begin{cases} 1 & \text{if} \ \ S(i, j) \geq T(i, j) \\ 0 & \text{if} \ \ S(i, j) < T(i, j) \end{cases} \tag{5}$$

In result we obtain the binary image $g(i, j)$ with the detected contours. Moreover, the constant k allows to adjust and change the sensivity of the edge detection algorithm.

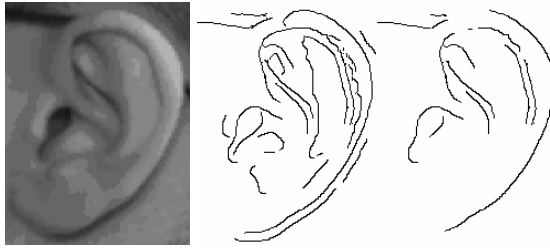An example of the edge detection algorithm is shown in the Fig. 4.

Fig. 4. Result of edge detection algorithm for two different values of k

3   Normalization

Given the binary image $g(i,j)$, we obtain the centroid which later becomes the reference point for feature extraction. Because the features for a recognition algorithm should be invariant to ear translation and scale change, the coordinates normalization is performed. Therefore we normalize coordinates, such that the centroid becomes the centre of the image. Suppose that the image with pixel coordinates (i,j) undergoes geometric transformations to produce an invariant image with coordinates (x,y). This transformation may be expressed as:

$$[x,y,z] = [i,j,1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -I & -J & 1 \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sigma_i} & 0 & 0 \\ 0 & \dfrac{1}{\sigma_j} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

where:
- I, J – centroid,
- σ_i, σ_j – standard deviation of i and j respectively.

Furthermore, our method is also invariant to rotation as all the rotated images of the same object have the same centroid. That is the major reason that we chose the centroid of the image to be the reference point in the feature extraction algorithm. Such approach allows the successful processing of RST queries.

4   Feature Extraction

There are many possible geometrical methods of feature extraction and shape description such as Fourier Descriptors, Delaunay Triangles and methods based on combination of angles and distances as parameters. We propose a 2 step-method that is based on number of pixels that have the same radius in a circle with the centre in

the centroid and on the contour topology. The algorithm for the first step of feature extraction is presented below:

1. we create a set of circles with the centre in the centroid (Fig. 5.)
2. number of circles N_r is fixed and unchangeable
3. we create circles in such a manner that the corresponding radiuses are α pixels longer from the previous radius
4. since each circle is crossed by the contour image pixels we count the number of intersection pixels l_r
5. next we calculate all the distances d between neighbouring pixels, we proceed in the counter-clockwise direction
6. we build the feature vector that consists of all the radiuses with the corresponding number of pixels belonging to each radius and with the sum of all the distances between those pixels $\sum d$
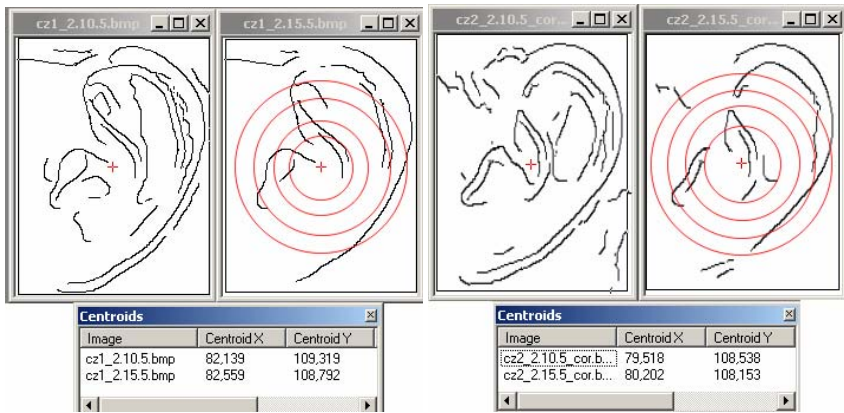

Fig. 5. Binary ear images with the extracted edges (2 values of *k*) and with the centroid marked with a cross. Circles represent the radius values for calculation of number of pixels intersecting each circle. The table below shows the centroid values for each binary image

The algorithm for *Nr=3* is symbolically presented in the Fig. 6.
The general rule for forming the first vector is presented below:

$$V = \left\{ \left[r_{min}, l_{r\,min}, \sum d_{r\,min} \right] \cdots \left[r_{max}, l_{r\,max}, \sum d_{r\,max} \right] \right\} \qquad (7)$$

where:
r – radius length,
l_r – number of intersection points for each radius,
$\sum d$ – sum of all the distances between the intersection points for the considered radius.
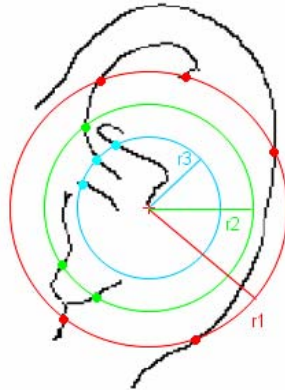
Fig. 6. The symbolic representation of our algorithm for $N_r = 3$

Then in order to enhance the distinctiveness of the extracted features, we build the second vector in the second step of feature extraction. Once again we base upon the created circles with the centre in the centroid. Hereby, we propose to extract the characteristic points for each contour in the normalized coordinates.

For each contour line the characteristic points are:

- contour endings,
- contour bifurcations,
- all the points that cross the created circles (those points are already extracted by the previous algorithm).

In each contour we check the topological properties of every pixel. For each contour pixel g_o we use 3×3 window as in Fig. 7. (left). When $g_0 = 1$, the connected number N_c^8 of g_0 is defined as:

$$N_c^8\left(g_o\right) = \sum_{k=S} \left(\bar{g}_k - \bar{g}_k\, \bar{g}_{k+1}\, \bar{g}_{k+2}\right), \text{ where } S = (1,3,5,7) . \qquad (8)$$
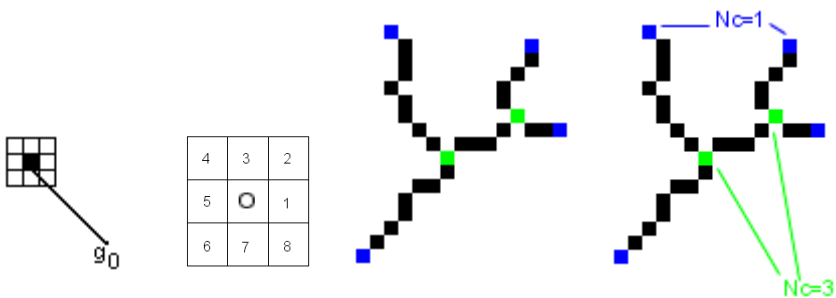

Fig. 7. Characteristic points detection algorithm

We search for the contour beginning in the area A, such that: $r_i \geq A > r_{i-1}$. We begin the search for $r_i = r_{max}$, which means that we start our algorithm in the most outer circle. Then we search in all other circles heading towards the centre in the centroid. If we come across any point with $N_c^8 = 1$, we check if it is already stored in the feature vector and if not, we store it as the ending point and we trace its contour. Points with $N_c^8 = 1$ and $N_c^8 > 2$ are the ear contour endings and the contour bifurcation points respectively. Those points are marked as E and B in the Fig. 8. For each contour we also extract the intersections with the circles created earlier. For each contour intersecting the circles we store all the intersections coordinates i and the number of such intersections N_I as presented in Fig. 8 (right) and Eq. 9
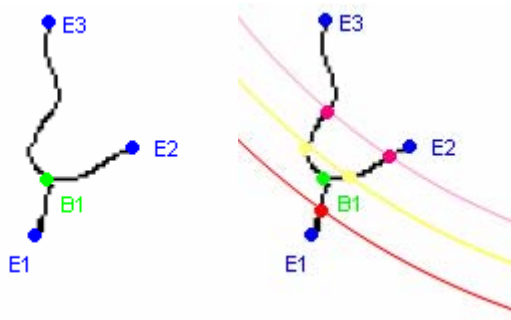

Fig. 8. The symbolic representation of the second step of feature extraction algorithm

The general rule for forming the second vector for each contour is presented below. First we store the number of endings, bifurcations and intersection points, and then we store all the coordinates of those points for all the extracted and traced contours. For C contours in a given image we obtain:

$$F = \left\{ \left[\left(N_E, N_B, N_I \right) \left(e_1, \cdots, e_{N_E}, b_1, \cdots, b_{N_B}, i_1, \cdots, i_{N_I} \right) \right]_1 \cdots \right.$$
$$\left. \cdots \left[\left(N_E, N_B, N_I \right) \left(e_1, \cdots, e_{N_E}, b_1, \cdots, b_{N_B}, i_1, \cdots, i_{N_I} \right) \right]_C \right\} \tag{9}$$

where:

N_E – number of endings in each contour,

N_B – number of bifurcations in each contour,

N_I – number of points intersecting the circles,

e – coordinates of endings,

b – coordinates of bifurcations,

i – coordinates of intersections in each contour.

5   Classification

For each image stored in the database we have two vectors F_{ref} and V_{ref}. For each input ear, we acquire many images under different angles to the camera.
The algorithm for recognition of an input image is following:

1. for the fixed number of circles, the feature vectors V and F of the input image are obtained
2. for each radius, we search the database feature vectors V_{ref} that have the same number of intersections l_r for the corresponding radiuses
3. the vectors with number of intersections $(l_r \pm 1)$ are also accepted, allowing the difference of one pixel on each circle
4. in the next step we check if the difference within the distance sum $\sum d$ for all the extracted vectors is less than a certain threshold value
5. if none of the vectors V_{ref} are found for the input image, the input image is rejected
6. if the number of intersecting points l_r is accepted and the difference within the distance sum $\sum d$ is less than a certain threshold value we check the contour-topology vector F
7. we first search for the same triples $\left(N_E, N_B, N_I\right)$ of the input contour-topology vector F with the reference contour vectors F_{ref}
8. then for the images with the same triples $\left(N_E, N_B, N_I\right)$ we check if the coordinates of the stored points are the same
9. if the corresponding coordinates of those vectors refer to the same points , the algorithm finds the winner of classification.

6   Experimental Results and Future Work

We perform our experiments on our own database of collected ear images. At the moment of writing, the database consists of over 240 images, but we are still adding more ear images of different type. For each person included in the experiments, we collected 2 left ear images, first with the camera perpendicular to the head and the second, with the camera within the specified angle of 30 degrees. We divided the database to several sets of images concerning their quality and degree of complexity. So far we have only experimented with images of very high quality and with the ideal conditions of recognition, without illumination changes. For such "easy" images from our database we obtained error-free recognition.
In further experiments we work with the "difficult" images and with the changing conditions of the image acquisition. In order to achieve satisfactory results with such complex images (Fig. 2) we are improving the contour detection algorithm, so that long, straight line-contours of glasses and artificial contours of earrings and hair are eliminated before applying feature extraction algorithm. We also think that the feature

vectors should be enriched with more geometrical features in order to better distinguish ear identity. Moreover, we search for other than geometrical features describing ear images, such as energy and shape parameters. We try to discover, which features are the most significant in determining ear identity, so that we will be able to weight them properly in the process of building hybrid vectors of features of different type.

7   Conclusions

In the article we proposed a human identification method based on human ear images. We believe that human ear is unique and has many advantages over other biometrics methods. We proposed invariant geometrical method in order to extract features needed to classification. First we perform contour detection algorithm, then size normalization. Thanks to placing the centre of the new coordinates system in the centroid, our method is invariant to rotation, translation and scaling, which will allow RST queries. The centroid is also a key reference point in the feature extraction algorithm, which is divided into 2 steps. In the first step, we create circles centred in the centroid and we count the number of intersection points for each radius and the sum of all the distances between those points. All those points are stored in the first feature vector corresponding to the radiuses. In the second step, we use the created circles, but hereby we count the intersection points for each contour line. Moreover, while tracing the contour lines, we detect the characteristic points like endings and bifurcations. Together with the intersection points for each contour, we store them in the second feature vector corresponding to contour topology. Then we perform classification, basing on the simple comparison between the input image feature vectors, and all the vectors from the database. So far we have obtained very good results, however we still continue our research in order to improve our method and add more parameters to the feature vectors.

References

[1]   Ashbourn J., Biometrics - Advanced Identity Verification, Springer-Verlag 2000.
[2]   Beveridge J.R., She R., Draper B.A., Givens G.H., Parametric and Nonparametric Methods for the Statistical Evaluation of Human Id Algorithms, Workshop on Evaluation Methods in Computer Vision, 2001.
[3]   Bowman E., Everything You Need to Know about Biometrics, Technical Report, Identix Corporation, 2000.
[4]   Burge M., Burger W., Ear Biometrics, Johannes Kepler University, Linz, Austria 1999.
[5]   Burge M., Burger W., Ear Biometrics for Machine Vision, 21 Workshop of the Austrian Association for Pattern Recognition, Hallstatt, 1997.
[6]   Canny J., A Computational Approach to Edge Detection, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, 679-698, 1986.
[7]   Choras M., Human Identification Based on Image Analysis – New Trends, Proc. Int. IEEE Workshop Signal Processing'03, pp. 111-116, Poznan 2003.
[8]   Danielsson P. E., Ye Q. Z., Rotation-Invariant Operators Applied to Enhancement of Fingerprints, Proc. 8th ICPR, Rome 1988.

[9]   Hoogstrate A.J., Heuvel van den H., Huyben E., Ear Identification Based on Surveillance Camera's Images, Netherlands Forensic Institute, 2000.

[10]   Hurley D.J., Nixon M.S., Carter J.N., Force Field Energy Functionals for Image Feature Extraction, Image and Vision Computing Journal, vol. 20, no. 5-6, 311-318, 2002.

[11]   Iannarelli A., Ear Identification, Forensic Identification Series, Paramont Publishing Company, California 1989.

[12]   Jain A., Bolle R., Pankanti S., Biometrics: Personal Identification in Networked Society, Kluwer Academic Publishers, 1999.

[13]   Jain L. C., Halici U., Hayashi I., Lee S. B., Tsutsui S., Intelligent Biometric Techniques in Fingerprint and Face Recognition, CRC Press International Series on Computational Intelligence, 1999.

[14]   Kouzani A.Z., He F., Sammut K., Towards Invariant Face Recognition, Journal of Information Sciences 123, Elsevier 2000.

[15]   Lai K., Chin R., Deformable Contours: Modeling and Extraction, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 17, no. 11, 1084-1090, 1995.

[16]   Safar M., Shahabi C., Sun X., Image Retrieval By Shape: A Comparative Study, University of Southern California, November 1999.

[17]   Victor B., Bowyer K.W., Sarkar S., An Evaluation of Face and Ear Biometrics, Proc. of Intl. Conf. on Pattern Recognition, I: 429-432, 2002.

[18]   Zhang D., Automated Biometrics – Technologies and Systems, Kluwer Academic Publishers, 2000.