Person Identification Using Ear Biometrics

Md. Mahbubur Rahman, Md. Rashedul Islam, Nazmul Islam Bhuiyan,
Bulbul Ahmed, Md. Aminul Islam

Computer Science and Engineering Discipline, Khulna University,
Khulna-9208, Bangladesh.
mahbubcse@yahoo.com, rashedislam742@hotmail.com, nazmul0226@yahoo.com

Abstract

In this paper, a method to recognize human using ear biometrics has been proposed. Traditional image-based approaches to person identification almost exclusively employ front views of the individual's face. Ear on the other hand, has a more uniform distribution of color, so almost all information is conserved when converting the original image into gray scales. When employing face or lip as a biometric, changing of their appearance with the expression of the subject creates problem but in case of ear the shape and appearance is fixed. We extract the features from an ear image by measuring geometric relations between predetermined points (lengths, angles). Here, the geometrical structures observed from pixel value distances are used for the successful recognition of objects. Recognition of objects is based on these geometrical structures. Experimental result reveals that the proposed identification method can achieve almost 89% accuracy. The complexity of this process is less since the method does not require to consider the inner boundary of ear.

Key words: *Biometrics, ear, recognition, identification, and false alarm rate.*

1. Introduction

Detecting objects from images is one of the most essential tasks of vision systems. In addition to this, the stability of tracking systems greatly depends on the detection of targets. A new class of biometrics based upon ear features was introduced for use in the development of passive identification systems by Alfred Iannarelli [1]. Identification by ear biometrics is promising because it is passive like face recognition, but instead of the difficulties to extract face biometrics, it uses robust and simply extracted biometrics like those in fingerprinting. The ear is a unique feature of human beings. Even the ears of "identical twins" differ in some respects. There are persons in crime laboratories that assume that the human external ear characteristics are unique to each individual and unchanging during the lifetime of an adult. Over the years, suggestions have been made in the occasional literature that the shapes and characteristics of human ear are widely different and may be in fact sufficiently variant such that it is possible to differentiate between the ears of all individuals. Unfortunately, this "individuality" has apparently been taken for granted but has never been empirically established. Techniques that allow computers to understand the shape of a human ear in images and video sequences can be used in a wide range of applications. In certain domains it suffices to recognize a few different shapes, observed always from the same viewpoint.

The most famous work among ear identification is made by Alfred Iannarelli [1] when he gathered up over ten thousands ears and found that all were different [2]. In the set of Five hundreds ears only four characteristics were needed to state the ears unique [5]. The performance is not significantly different between face and ear; for example 69.3% versus 72.7% respectively in one experiment [7]. Ear biometrics based on ear form are averagely permanent than other possible identification system e.g. fingerprint, hand geometry etc.

Alfred Iannarelli's second study was for researching identical and non-identical twins. These cases support the hypothesis about ear uniqueness. Even the identical twins had similar, but not identical, ear physiological features [2]. Alfred Iannarelli has created a twelve-measurement system that is named "Iannarelli System". He used only the right ear of the people. To normalize the picture, the distances between each of twelve numbered areas are measured and assigned an integer value [1]. But it had some problem that is generated by Burge and Burger [2]. It is not suitable for machine vision because of the difficulty of localizing the anatomical points. If the first point is not accurate then none of the measurements are useful. A biometric specialized company Bromba GmbH has also shown a table with respect to the comfort, accuracy and costs [Table 1].

The multiple identification method, which combines the result from several neural classifiers using feature of outer ear points, information obtained form ear shape and wrinkles, and macro features extracted compression network [6]. Voronoi diagram is used to determine its curve segment [2]. The force field transformations are used for ear recognition [4]. The image is treated as an array of Gaussion attractors that act as the source of the force field. The comparison between ear and face is a dimensionality reduction technique in which variations in the data set are preserved [3]. Most of the above literature indicates that the variability between ears is so large that it might be possible that ears are unique, and moreover possibly uniquely distinguishable on a limited number of features or characteristics.

Table 1: The best method has the o-symbol and the worst least.

Biometrics Trail	Comfort	Accuracy	Cost
Ear form	00000	0000	00000
Signature	000	0000	0000
Facial geometry	00000000 0	0000	00000
Iris	00000000	000000000	00000000
Retina	000000	00000000	0000000
Fingerprint	0000000	000	0000
Voice	0000	00	00
DNA	0	0000000	000000000
Finger geometry	0000000	000	0000
Keyboard strokes	0000	0	0

In the following sections, we describe the concept of the ear detection. In Section 2, we describe an algorithm for collecting data from an ear images. Finally, we present results.

2. The Overall System

The proposed system for person identification by ear biometrics is shown in Fig. 1. Initially image is acquired through digital camera or other means. The image is converted to gray scale. Then using threshold a binary image obtained. After that Region of Interest (ROI) is chosen which is the rectangular area containing the ear image. Masking is used to detect the edges of the ear image. Using Generalized Hugh Transform (GHT) the features of the ear is extracted. The extracted features along with the subjects id is stored in the database for testing for a match.

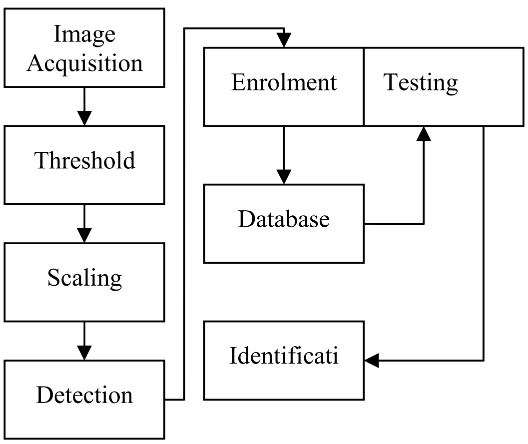

Fig. 1: Person identification using ear biometrics.

2.1 Anatomy of an Ear

The ear can easily identify the person. The ear consists of several parts. Those are helix (Three parts i.e. H. superior, H. anterior, H. posterior), pliegue superior, Foseta, Coneha, Origen, Trago, Canal intertraguiano, pliegue inferior, Fose navicular, Antitrago, Lobule, Zone, Lobule etc (Fig. 2). In the preprocessing step the ear image is resized into a fixed size. Then to distinguish the edges, threshold is applied.

Fig. 2: The anatomy of the ear: 1. Helix Rim 2. Lobule 3. Triangular Fossa 4. Cannal Intertraguiano 5. Concha 6. Crus of Helix 7. Tragus 8. Antihelix 9. Antritragus.

2.1 Threshold

Threshold is a technique that creates a two level binary image. It helps in edge detection. A threshold image $G(x, y)$ can be found as:

$$G(x, y) = \begin{cases} 1 \ \ if \ f(x, y) > T_{xy} \\ 0 \ \ if \ f(x, y) \leq T_{xy} \end{cases}$$

where T_{xy} is the threshold assigned to location (x, y) in the image (Fig. 4). T_{xy} is determined from the histogram of the ear image (Fig. 3).

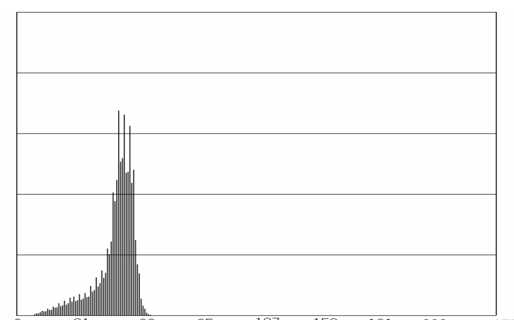

Fig. 3: Histogram of Fig. 2.

Since the boundary of ear has a distance from face, it contains black pixels along their boundary. We do convert all the pixel value that stands between zero to the threshold value into white pixel and remaining are black. Then the image has white pixels across its boundary and others have black. Since maximum area is skin, the peak contains the skin value (Fig. 3). Threshold is determined by using the following algorithm [8].

1. Select an initial estimate for T_{xy}, which is same as maximum frequency value.

2. Segment the image using T_{xy}. This will produce two groups of pixels: G_1 consisting of all pixels with gray level values $> T_{xy}$ and G_2 consisting of pixels with values $\le T_{xy}$

3. Compute the average gray level values μ_1 and μ_2 for the pixels in regions G_1 and G_2.

4. Compute a new threshold value:

$$T_{xy} = \frac{1}{2}(\mu_1 + \mu_2)$$

5. Repeat steps 2 through 4 until the difference in T_{xy} in successive iterations is smaller than a predefined value 0.5.

Fig. 4: Threshold of Fig. 2.

2.2 Masking

The most common way to look for edges is to run the mask through the image. This procedure involves computing sum of products of the coefficients with the gray levels contained in the region. The response of a $m \times n$ mask at any point (x, y) in the image is by using the following expression:

$$R = W_1 Z_1 + W_2 Z_2 + \ldots\ldots\ldots\ldots W_{mn} Z_{mn} = \sum_{i=1}^{mn} W_i Z_i$$

Where Z_i is the gray level of the pixel associated with mask coefficient W_i.

We used a mask, where W is:

$$W = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -2 & -1 & 0 & 0 \\ 0 & -1 & -2 & -8 & -2 & -1 & 0 \\ -1 & -2 & -8 & 60 & -8 & -2 & -1 \\ 0 & -1 & -2 & -8 & -2 & -1 & 0 \\ 0 & 0 & -1 & -2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}$$

The value of R will be assigned as pixel value to that location of the image. After masking all the edges will be shown in the image (Fig. 5).

Fig. 5: Masking of Fig. 3.

2.3 Detection

First of all Helix Superior, Helix Posterior, lobule and Cannal Intertraguiano are to be detected by using generalized Hough Transform (GHT) and four points or region of points are specified (E, F, G, H of Fig. 9). In case of Helix Superior top most points are taken and right most points for Helix Posterior, bottom most points for

Lobule and left most points for Cannal Intertraguiano. We consider some fixed number of pixels residing in the specified region of points that create edges of a rectangle (A, B, C, D of Fig.9).

The normal representation of a line is

$$X\cos\theta + Y\sin\theta = P$$

Instead of straight lines, however, the loci are sinusoidal curves in the $P\theta$-plane. K Collinear points lying on a line $X\cos\theta_j + Y\sin\theta_j = P_i$ yield K sinusoidal curves that intersect at (P_i, θ_j) in the parameter space. Incrementing θ and solving for the corresponding P gives K entries in accumulator $A(i, j)$ associated with the cell determined by (P_i, θ_i) (Fig. 6).

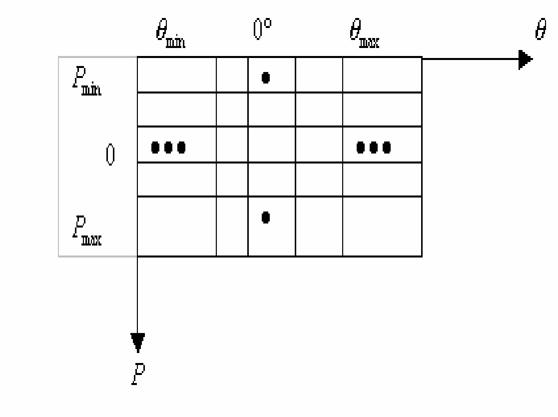

Fig. 6: Subdivision of a line

The range of angle is $\theta \pm 90^0$, measured with respect to the X-axis. Thus a horizontal line has $\theta = 0^0$, with P being equal to the positive X-intercept. Similarly, a vertical line has $\theta = 90°$, with P being equal to the positive Y-intercept or $\theta = -90°$, with P being equal to the negative Y-intercept.

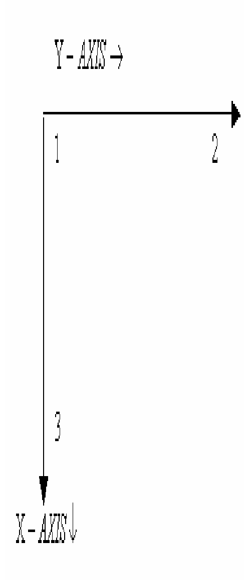

Fig 7.

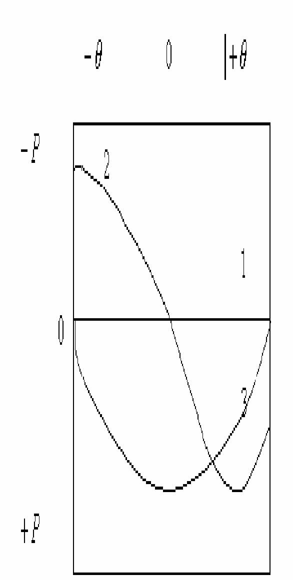

Fig 8.

Fig. 7: Some points (1, 2, 3) of an image. Fig. 8: Mapping points in Fig. 7 onto the $P\theta$ plane.

The intersection point of the curves corresponding to points 1 and 2 (Fig. 7) in the XY – image, lie on a straight line passing through the origin $(P = 0)$ and oriented at 0^0 (Fig. 8). Similarly the intersection point of the curves corresponding to points 1 and 3 (Fig. 7), lie on a straight line passing through the origin $(P = 0)$ and oriented at $\pm 90^0$ (Fig. 8).

Approach towards edge linking is as follows:

1. Examine the counts of the accumulator cells for high pixel concentrations.
2. Examine the relationship between pixels in a chosen cell.

The concept of continuity in this case is based on computing the distance between disconnected pixels identified during traversal of the set of pixels corresponding to a given accumulator cell. A gap at any point is significant if the distance between that point and its closest neighbor exceeds a certain threshold.

Using GHT, we find some fixed amount of pixels that lie in a horizontal or vertical line. We searched from upper position to find the top most line and from lower end to find bottom most line and from left end to find leftmost line and right end to find rightmost line.

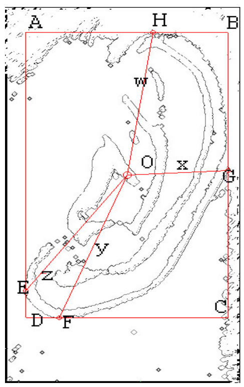

Fig. 9: Topmost, bottommost, leftmost and rightmost points of a detected ear image.

The cross-point *O* of two diagonals of that particular rectangle is determined. The distances from this point to topmost point, rightmost point, bottommost point, leftmost point of ear are determined in accordance with alignment. Suppose those distance are *w, x, y, z* (Fig. 9). The ratio of these distances is chosen for identification. Let w', x', y', z' are distances at the time of identification and *w, x, y, z* are distances at the time of enrollment. Since the subject and environment and subject changes over time a limited tolerance must be permitted. If permitted tolerance is *t* then the task is to determine whether

$$\left(\frac{x}{w}:\frac{y}{w}:\frac{z}{w}:\right)-\left(\frac{x'}{w'}:\frac{y'}{w'}:\frac{z'}{w'}\right)=\pm t$$

Besides alignment (i.e. angle with horizon or vertical) of line OH, OG, OF and OE may also be used as passive biometrics choices and if this choices is incorporated with the previous biometric choice then it is obvious that identification success rate will be high. After edge detection those four points will be detected and find out the cross point of the diagonals.

3. Experimental Result

We tested the method on 350 samples of 100 persons by day variation taken from various camera viewpoints, and an image size of 250x230.

Table 2: Experimental Result

Test Case No.	Test criteria	Person	Sample	Detection rate	False Alarm Rate
One	Same day	100	350	90%	3%
Two	Day variation	100	300	88%	4%
Three	Light variation	100	300	87%	5%

Table 2 shows the measured performance. It shows that the recognition rate is the highest for the same day (test case one). Here out of the hundred subject's images, only eighty images were selected and their extracted data are stored in the database. Then three hundred and fifty samples are given for identification. In the next test case two hundred and forty image of the subjects in the training set and sixty samples of the subjects not in the training set. Test case three have the same criteria except it was collected in different lighting condition. The detection ratios under different conditions are shown in Fig. 10.

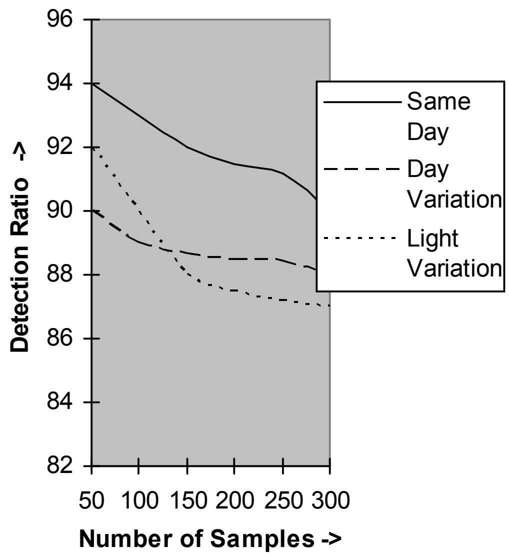

Fig. 10: Comparison between different test conditions.

3.1 Comparison with other system

In our proposed method, detection ratio is overall eighty nine percent, which is better than Iannarelli's and Burger M. and Burger W's methods, which were sixty nine percent and seventy three percent respectively. They were the most successful in the field of ear detection. The main thing that is making difference between is that we are considering the alignment that helps in better detection ratio, which was not used before. In case of pose variation there may be change of distance but alignments help to detect accurately. Here inner boundaries are discarded, because they most often erroneous by the gathering of oil, dirt, wax etc (Fig. 12) which guide to wrong information. For this reason complexity of the method is less than Iannarelli method in which twelve measurements were taken (Fig. 11) and less than Burger M. and Burger W who also considered these inner curves.

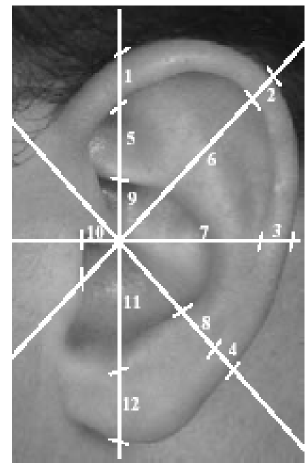

Fig. 11: Iannarelli's Measurement.

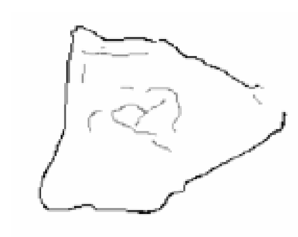

Fig. 12: Noise.

4. Conclusions

In this paper, we have described a system that tracks and detects ear features simply and robustly. First of all, appropriate threshold value is identified and then ear boundary is detected. Then edge linking is done. Data taken from the ear image is compared with the database. Our ear detection algorithm is quite simple and, hence, has low computation complexity and can be applied in many real-time applications. In selected populations (e.g., those with short hair as in the defense industry), it is especially applicable. Currently we are working to enhance the identification rate of the system using multi-biometrics.

References

[1] Iannarelli, A., *Ear Identification,* Forensic Identification Series, Paramount Publishing Company, Fremont, California (1989).

[2] Burger, M. and Burger, W. "Ear biometrics", In *Biometrics: Personal Identification in Networked Society*, ed. Jain A. et al., Kluwer Academic Publishers (1998).

[3] Victor, B., Chang, K., Bowyer, K.W., and Sarker, S., "An evaluation of the face and ear biometrics", IEEE *International Conference on pattern Recognition,* pp. 492-43 (2002).

[4] Hirschi, "Identifizierung von Ohrab-drücken"*, Kriminalistik* 24, pp. 75-79 (1970).

[5] Imhofer, R., "Die Bedeutung der Ohrmuschel für die Feststellung der Identität", *Archiv für die Kriminologie*, Bd. 26, pp. 150-163 (1906).

[6] Moenssens, A., *Scientific evidence in civil and criminal cases*, 4th Ed., Foundation Press, pp. 555-620 (1995).

[7] Chang, K., Bowyer, K.W., Sarker, S., Victor, B., "Comparison and Combination of Ear and Face Machine Image in appearance –Based Biometrics', *IEEE Transaction on pattern Analysis and machine Intelligence*, vol. 25, no, 9 (2003).

[8] Gonzalez, R. C., Woods, R. E., Digital *Image Processing*, 2nd Ed., Pearson Education Inc. (2002).