A Brief Approach to the Ear Recognition Process

Pedro Luis Galdámez, María Angélica González Arrieta, and Miguel Ramón Ramón

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

Abstract. This paper offers an approach to biometric analysis using ears for recognition. The ear has all the assets that a biometric trait should possess. Because it is a study field in potential growth, this study offers an approach using SURF features as an input of a neural network with the purpose to detect and recognize a person by the patterns of its ear, also includes, the development of an application with .net to show experimental results of the theory applied. Ear characteristics, which are a unchanging biometric approach that does not vary with age, have been used for several years in the forensic science of recognition, thats why the research gets important value in the present. To perform this study, we worked with the help of Police School of Ávila, Province of Spain, we have built a database with approximately 300 ears.

Keywords: Neural Network, SURF Algorithm, Ear Recognition.

1 Introduction

The ear has been used as a means of human recognition in forensic activities for a long time. During the investigation of various crime scenes, earprints commonly been used to identify a suspect especially when there is no information of fingerprints. A recognition system based on images of the ears is very similar to a typical face recognition system, however, the ear has some advantages over the face, for example, their appearance does not change due to expression and is little affected by the ageing process, its color is usually uniform and the background is predictable.

Although the use of information from ear identification of individuals has been studied, is still an open question by specifying and determining whether or not the ear can be considered unique or unique enough to be used as a biometric. Accordingly, any physical or behavioural trait can be used as biometric identification mechanism provided which is universal, that every human being possesses the identifier, being distinctive and unique to each individual, invariant in time, finally measurable automatically or manually, the ear accomplish all these characteristics.

S. Omatu et al. (eds.), *Distributed Computing and Artificial Intelligence, 11th International Conference,* Advances in Intelligent Systems and Computing 290,
DOI: 10.1007/978-3-319-07593-8_54,

2 State of Art

The first technique known to detect the ears is raised by Burge and Burger [17] who have made the process of detection using deformable contours with the observation that initialization contour requires user interaction. Therefore, the location of the ear is not fully automatic. Meanwhile Hurley et al. [7] used the technique of force field, this process ensures that it is not required to know the location of the ear to perform recognition. However, only applies when he has the image specifies the ear without too much noise. In [19], Yan and Bowyer have used manual technique based on two previous lines for detection, where you take a line along the border between the ear and face while another line crosses up and down the ear.

The strategy proposed by Alvarez et al. [15] uses ovoid contour of the ear where the limit is estimated by fitting the shape of an ear in the image using a combination of snake line and ovoid models, like the proposal of Burge and Burger this technique requires an initialization. Ansari and Gupta [20] presented a process based on the outer ear helices edges, they use 700 samples collected at IIT Kanpur, the strategy only relies on the outer helix curves. Yuan and Mu [16] have proposed a skin-color and contour information technique, they perform the ear detection considering ear shape elliptical and fitting an ellipse to the edges to get the accurate ear position. Attarchi et al. [21] have shown an ear detection process based on the edge map. It relies on the hypothesis that the longest path in edge image is the ear outer boundary.

A. Cummings [2] shows a strategy using the image ray transform which is capable of highlighting the ear tubular structures. The technique exploits the helix elliptical shape to calculate the localization. Kumar et al [1], have introduced a proposal where uses skin segmentation and edge map detection to find the ear, once they find the ear region apply an active contour technique [22] to get exact location of ear contours, the technique has been tested on 700 ear images. Like these techniques there are many other significant proposals.

The most used technique for face recognition [18], principal component analysis (PCA), is also suitable for use in ear recognition. The first application of PCA to ear recognition was by Victor et al. [3] they used PCA to perform an comparative analysis between face and ear recognition concluding that the face performs better than the ear. However, Chang et al. [14] also have accomplished a comparison using PCA and found that ears provided similar performance to faces, they concluded that ears are essentially just as good as faces for biometric recognition. There are many proposals to solve the problem, in this paper only has done a small review from some of them, the next section introduce an attempt to solve the ear recognition problem in a practical way, applying the concepts studied for 2D images to develop an application which allow to perform the ear recognition in real-time video.

3 Ear Recognition System

Most of ear biometric articles have centred their attention on recognition using manually cropped ear images. However, for a robust ear recognition system is desired to detect the ear from the profile face image in an automatic way.

4 Detecting and Tracking the Ear

There exist some techniques which could be used to detect ear automatically. However, These techniques usually can detect the ear only when a profile face image do not contain a noisy or big background around the ear. This section proposes an useful ear localization technique which attempts to solve these issues.

4.1 Ear Localization

OpenCV and its wrapper for .Net framework EmguCV includes different object detectors based on the Viola-Jones framework. Modesto Castellón-Santana et al. [5] have developed a haarcascade classifier to be used with OpenCV to detect left and right ears. This classifier represents a first step to create a robust ear detection and tracking system. The Application is developed in C#.

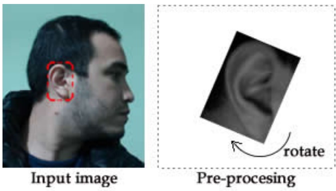

Fig. 1. Ear detection

With the ear identified we proceed to perform the pre-processing task, converting the image to gray scale and we begin the normalization process, first we perform the segmentation of the image applying a mask to extract only the ear, then the image is converted to an edge map using the canny edge filter. If w is the width of the image in pixel and h is the height of the image in pixel, the canny edge detector takes as input an array $w \times h$ of gray values and sigma. The output is a binary image with a value 1 for edge pixels, i.e., the pixel which constitute an edge and a value 0 for all other pixels. We calculate a line between major and minor y value in the edge image to rotate and normalize each image, trying to put the lobule of the ear in the centre. This process is trying to get all the images whose shape is similar to the image to identify. We identify some points on the external shape of the ear and the angle created by the center of the line drawn before and the section in the ear's tragus with the major x value.

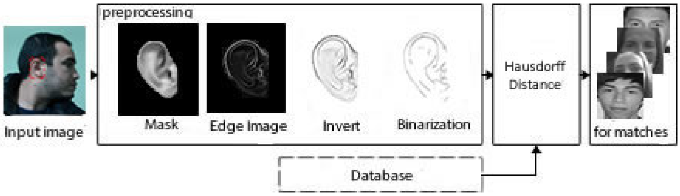

Fig. 2. Image pre-processing

Once the pre-processing is complete we proceed to compute a Match within the database using the contours of the ear form, with this we are trying to reduce the number of candidates. The Hausdorff distance measure used in this document is based on the assumption that the ear regions have different degrees of importance, where characteristics such as tragus, antitragus and helix contour play the most important role in ear form. The algorithm applied is based on what is stated in [13]. It operates basically in the comparison of edge maps. The advantage of using edges to match two objects, is that this representation is robust to illumination change. In this stage we also compute the SURF features to track the ear in the video.

4.2 Tracking the Ear

Speeded Up Robust Features (SURF)[9] is a scale and rotation invariant interest point detector and descriptor. It has been designed for extracting highly distinctive and invariant feature points from images. One of the basic reasons to use SURF for the feature representation is to analyse how the distinctive characteristics works in images, and at the same time is to found more robust with respect to change, taking into account the point of view, rotation, scale, illumination and occlusion [9] as compared to other scale and rotation invariant shape descriptors such as SIFT [6] and GLOH [12]. The result for the feature vectors SURF is the relative measured to the dominant orientation to generate each vector that represent an invariant with respect to rotation of the image.

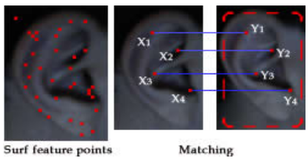

Fig. 3. Example of SURF Features

The way SURF process pairing is using the most proximate neighbour ratio pairing. To get the greatest pairing match for a key-point of a picture inside in another picture is elucidated by detecting the most proximate neighbour in the other key-points from a second picture where the most proximate neighbour is defined as the key-point with the least euclidean distance from the known key-point of the first picture between their characteristic unidirectional matrices. Due to the fact that these SURF vectors are invariant to the image rotation, the process of ear detection combining the previous viola-jones approach with the SURF vectors becomes robust and efficient.

Fig. 4. Tracking Ear using SURF Features

The approach to isolate the ear in the image, the prototype we used for the ear identification should reveal the characteristics of scale and rotation immutability. To calculate such prototype in a suggested method, an invariant shape characteristic to rotation and scale was used. Among numerous scale and rotation invariant shape characteristics, SURF [11] offers respectable distinctive features and at the same time it is robust to variations in viewing circumstances, rotations and scales. SURF denotes a picture by first detecting some exclusive feature points in it and then by describing them with the support of a unidirectional feature descriptor matrix.

5 Ear Recognition Using a Neural Network

The ear image is recreated through the SURF algorithm as a set of salient points, where each on is associated with a vector descriptor. Each can be of 64 or 128 dimensions. The 128 dimensional descriptor vector is considered the more exacting feature based in the knowledge that is always best to represent the image with the most powerful discriminative features possible. A method to obtain a unique characteristic fusion of one sole individual is proposed by combining characteristics acquired from various training instances of the individual. If we have n ear images of an individual for training, a fused prototype is gained by fusing the feature descriptor array of all training images collected, considering the redundant descriptor array only once.

Having all the images processed, a collection was made with tags indicating to whom each image and fusion vector, belongs. These vectors are used as inputs to train the network. In the training algorithm, the unidirectional matrices or vectors of values belonging to an individual, are taken as positive returning 1 as the neuron output assigned to that user and 0 to other neurons.

Fig. 5. Avila's Police School Database

6 Experimental Results

The results obtained in the process of detection and recognition of the ear are presented in this section, Table 1 shows the percentages of accuracy when only using the Viola-Jones classifier included in OpenCV vs the potentiation accomplished by adding the tracking with SURF features. It can be seen that in 2D images or photographs the difference is not so evident, however when the process is done on video, the difference is almost 10 percentage points, and is only done when considering the location of the ear in the video in different pose and lighting conditions. If you take into consideration the time it succeeds in maintaining the object identified, the algorithm combined with SURF tracking is much more accurate because these features allow you to place the image even if it has a 180 degrees event that does not happen with ear.

Table 1. Ear Detection (Haar-Cascade and adding SURF Tracking)

	#Images	*EarLocalization(%)*	
		Haar − Cascade	*With SURF tracking*
2D Images	308	92.53	98.70
Real Time Video	308	86.69	95.13

In Table 2 we can observe the results of the recognition process in normal conditions with controlled lighting. At this stage we have compared the results obtained with traditional algorithms such as PCA and Fisher to check the validity of our work. In this sense the results are encouraging, using SURF features as input of a neural network with different test subjects, we get a recognition percentage higher than traditional algorithms in video, however, the precision decreases to 14% being always greater in a variable range from 10 to 20% compared to the PCA and Fisher when we change the normal conditions.

Summarizing with perspective and illumination in normal conditions, we get 86% of succeed in recognition with PCA, 87% with fisher algorithm, using the neural network with SURF descriptors, the percentage increased to 92%, over more than 300 attempts of different individuals.

Table 2. Normal Conditions

	PCA		*Fisher*		*NeuralNetwork*	
	Positive	*Negative*	*Positive*	*Negative*	*Positive*	*Negative*
Positive	131	49	197	37	269	11
Negative	41	118	38	41	23	107

The parameter settings of the neural network used in this method are dynamic, the output neurons depends on Hausdorff Distance filter stage where the algorithm selects some possible answers to the recognition problem in order to reduce the amount of candidates to solve the problem. The hidden layer is created dynamically, respecting that the number of hidden neurons should be between the size of the input layer and the size of the output layer, should be 2/3 the size of the input layer, plus the size of the output layer; and less than twice the size of the input layer based on the research of Jeff Heaton [8].

7 Conclusion and Future Work

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

As future work, the most interesting and useful for the police is to achieve the development of an application not only able to propose candidates from the image of an ear but to achieve the identification and recognition of a criminal using an ear otogram found at a crime scene. The ear otogram is nothing more than a photo print of the ear identical to that obtained from fingerprints. Criminals sometimes place their ear on the door and from there they the police can get the ear photo print. The results of this research are pointing towards that goal, although preliminary, they show a significant progress to approach the final purpose, recognition based on these otograms.

References

1. Kumar, A., Hanmandlu, M., Kuldeep, M., Gupta, H.M.: Automatic ear detection for online biometric applications, pp. 146–149 (2011)
2. Cummings, A., Nixon, M., Carter, J.: A novel ray analogy for enrolment of ear biometrics. In: BTAS 2010, pp. 1–6 (2010)
3. Victor, B., Bowyer, K., Sarkar, S.: An evaluation of face and ear biometrics. In: ICPR 2002, vol. 1, pp. 429–432 (2002)
4. Carreira-Perpinan, M.A.: Compression neural networks for feature extraction: Application to human recognition from ear images. Master's thesis, Faculty of Informatics, Technical University of Madrid, Spain (1995)
5. Castrillón-Santana, M., Lorenzo-Navarro, J., Hernández-Sosa, D.: An Study on Ear Detection and Its Applications to Face Detection, pp. 313–322 (2011)
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
7. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field feature extraction for ear biometrics. Computer Vision and Image Understanding, 491–512 (2005)
8. Heaton, J.: Introduction to Neural Networks for C#, 2nd edn. (2010), `http://www.heatonresearch.com`
9. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (SURF). Computer Vision and Image Understanding, 346–359 (2008)
10. Jolliffe, I.T.: Principal Components Analysis, 2nd edn., New York (2002)
11. Bustard, J., Nixon, M.: 3D morphable model construction for robust ear and face recognition. In: CVPR 2010, pp. 2582–2589 (2010)
12. Krystian, M., Cordelia, S.: A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1615–1630 (2005)
13. Lin, K.-H., Lam, K.-M., Siu, W.-C.: Spatially eigen-weighted Hausdorff distances for human face recognition. Polytechnic University, Hong Kong (2002)
14. Kyong, C., Kevin, W.B., Sudeep, S., Barnabas, V.: Comparison and combination of ear and face images in appearance-based biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1160–1165 (2003)
15. Alvarez, L., Gonzalez, E., Mazorra, L.: Fitting ear contour using an ovoid model. In: Proceedings of IEEE International Carnahan Conference on Security Technology (ICCST 2005), pp. 145–148 (2005)
16. Yuan, L., Mu, Z.-C.: Ear detection based on skin-color and contour information. In: ICMLC 2007, vol. 4, pp. 2213–2217 (2007)
17. Burge, M., Burger, W.: Ear biometrics in computer vision. In: ICPR 2000, vol. 2, pp. 822–826 (2000)
18. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience 3(1), 71–86 (1991)
19. Yan, P., Bowyer, K.W.: Empirical evaluation of advanced ear bio-metrics. In: Conference on Computer Vision and Pattern Recognition, vol. 3, pp. 41–48 (2005)
20. Ansari, S., Gupta, P.: Localization of ear using outer helix curve of the ear. In: Proceedings of the International Conference on Computing: Theory and Applications (ICCTA 2007), pp. 688–692 (2007)
21. Attarchi, S., Faez, K., Rafiei, A.: A new segmentation approach for ear recognition. In: Blanc-Talon, J., Bourennane, S., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2008. LNCS, vol. 5259, pp. 1030–1037. Springer, Heidelberg (2008)
22. Lankton, S., Tannenbaum, A.: Localizing region-based active contours. IEEE Transactions on Image Processing 17(11), 2029–2039 (2008)
23. USTB Database, University of Science and Technology Beijing