EAR Recognition using Artificial Neural Networks

Akshay Sanjay Borse, Ammen Sanjay Siingh, Arshad Harun Attar, Nuzhat F Shaikh

Abstract— Authentication of a person is to ascertain his/her identity is an important problem in the society. Among various physiological traits(biometrics) , ear has gained much popularity in recent years as it has been formed to be the reliable biometrics for human recognition .Ear recognition consists of the two important steps:-

I. Ear Detection

II. Ear Recognition

Ear Detection causes out the segmentation of the ear from the profile face. In this project we have decided to work on an available database and implement the neural network for the classification of the person from the specific image.

Index Terms—Ear recognition, SURF matching, Neural Networks, Concept Learning.

I. INTRODUCTION

Human ears offer some distinct advantages over other biometric modalities: they have a wealthy of structural features that are permanent with increasing age from about 8 to 70 years old, and they are not affected by the expression variations Current ear recognition approaches have exploited how to use 2D ear image and 3D ear model for human identification. At present, 3D ear recognition performs well in illumination variation or pose variation, but it needs expensive computation and special equipment's, most of the resent works on ear recognition are focused on 2D images because using 2D images is more consistent with deployment in surveillance or other planar image scenarios. Recent research on ear recognition in 2D can be categorized as follows: ear recognition under controlled environment, ear recognition with pose variation and ear recognition under partial occlusion. In constrained environment, the proposed ear recognition methods perform well. But in real applications, human. Ears will be occluded in some scenarios. So ear recognition under occlusion, pose variation or with noised images are unavoidable problems. A direct comparison is made with other recognition techniques for ear recognition. The proposed method in the paper achieves better recognition performance.

II. PROBLEM DEFINITION

Amidst many biometric identification systems, ear has received special light in the recent years due to its reliability as a unique biometric trait. Here, we have proposed a system which uses twelve features of the ear. They represent manually measured distances among various feature points.

Manuscript Received on November 2014.

Akshay Borse, Computer Engineering, University of Pune, Modern Education Society's College of Engineering , Pune , India.

Ammen Siingh, Computer Engineering, University of Pune, Modern Education Society's College of Engineering, Pune , India.

Arshad Attar, Computer Engineering, University of Pune, Modern Education Society's College of Engineering, Pune, India.

This paper suggests that ear can be used as a biometric system using a limited number of features. The decidability index of the ear is found to be more than that of the face but less than that of iris. The decidability index represents the separation of genuine and false scores for a biometric system. The characteristics responsible for making ear popular as a biometric system are given below:

1. Ear has a uniform color distribution; also it doesn't depend on expressions like face.
2. Ear's shape remains constant from the age of 8 to 70.
3. Ear is easy to detect and localize as the background remains constant.
4. Ear's size is much larger than other systems like fingerprints, iris, etc.
5. It's not affected by cosmetics and spectacles.
6. It's a passive trait and doesn't need cooperation of the person.

In spite of ear having so many favorable properties, the low accuracy rate of ear recognition systems has kept it far from regular usage. Here, we try to change that and improve ear recognition efficiency by coming up with a modified approach. Any ear recognition system has two steps:

1. Ear Detection
2. Ear Recognition.

Ear Detection is the segmentation of the ear from the profile face picture. Recognition is the recognition of the with respect to the ear that has been segmented. Poor contrast and illumination, noise, poor database and probe image are the few problems faced in this system. We propose a 2D ear recognition technique from an already segmented database. We use three parallel image enhancement techniques to neutralize the problems of poor contrast and illumination. We use the local features of the ear to handle poor registration of database and probe images.

III. VARIOUS EAR RECOGNITION METHODS

The present ear recognition techniques in 2D can be classified into: appearance based techniques, force field transformation based techniques, geometric features based techniques and neural network based techniques. Principal Component Analysis (PCA), Independent Component Analysis (ICA) fall under appearance based techniques. PCA uses the training database to discover the orthogonal vectors which represent the max variance with minimum reconstruction mean square error; it drops the initial eigen-vector thinking that it represents the illumination changes in the given image. When present features are less, then PCA is observed to be working more efficiently than ICA. PCA and ICA can only be used when the images are captured in controlled environment. They do not give any invariance and thus are require very accurate registered database for efficient and consistent results.

Force field based techniques transform the ear into a force field and uses it to extract features. In this, the image is considered to be a collection of mutual attracting particles that act as a source of Gaussian force field. This technique has proven itself to be more efficient than PCA when the images are not properly registered. Burge and Burger [1] have come up with an ear recognition technique using its geometric features. The ear has been represented using a nearest neighborhood graph whereas template comparison has been done through sub graph matching. In this technique, shape feature vector and structural feature vector are used to extract the features of the outer and inner ear respectively. The first well known technique used for ear detection is due to Burge and Burger [4]. It has detected ears with the help of the deformable contours. But contour initialization in this technique needs user interaction. As a result, ear localization is not fully automatic so Hurley et al.[5] have used force field technique to get the ear location properly. This technique claims that it does not require exact ear localization for ear recognition. However, it this technique is only applicable when a small background is present in ear image. In [6] Yan and Bowyer have used manual technique based on two-line landmark to detect ear where one line is taken along the border between the ear and the face while other line is considered from the top of the ear to the bottom. The 2D ear localization technique proposed by the Alvarez et al [7] uses the ovoid and active contour (snake) [11] models. Ear boundary is estimated by fitting the contour of an ear in the image by combining snake and ovoid models. This technique requires the initial approximated ear contour to execute and hence cannot be used in the fully automated ear recognition system. There is no mathematical evaluation of the technique. Yan and Bowyer [8] have proposed the another technique by considering a predefined sector from the nose tip as the probable ear region. It first computes the ear pit using the curvature information obtained from 3D data and uses its boundary to initialize active contour which detects the ear boundary. It fails if the ear pit is occluded. It produces 78.79% correct ear segmentation when only color information is used for active contour conversion. Ansari and Gupta [9] have presented the ear detection technique based on edges of the outer ear helices. The accuracy of the technique is reported to be 93.34% on 700 sample images collected at IIT Kanpur. This technique solely relies on the parallelism between the outer helix curves and does not use any structural information present in the inner part of the ear and hence, it may fail if the helix edges are poor. Yuan and Mu [10] have proposed the technique based on the skin-color and contour information. It detects ear by roughly estimating the ear location and thus by improving the localization using the contour information. It considers the ear shape elliptical and fits an ellipse to the edges to get the accurate position of the ear. There is no quantitative evaluation reported for this technique. Another ear localization technique which exploits the elliptical shape of the ear has been proposed it has been successfully tested on 252 images of 63 individuals selected from XM2VTS and 942 image pairs of 302 subjects of UND database. For XM2VTS database which is relatively small and has less complex images, this technique has achieved 100% detection rate. However for UND database which contains the complex images, it has offered only 91% detection rate. Moreover, the assumption of considering ear shape as an elliptical for all the subjects may not be true and hence, may not help in detection of the ear. So in this paper we propose a new and efficient technique for the ear recognition in 2D images.

IV. PROPOSED TECHNIQUES

1. IMAGE ENHANCEMENT:

As discussed earlier, our method uses two parallel image enhancement algorithms to remove the problems of poor illumination and contrast from the images. These are given below:

Images used for the ear recognition may be affected by various factors i.e. noise, contrast, illumination. There arises the need of enhancement of the images before the recognition of the images. This section describes the algorithms to be used for the enhancement of the images. It briefly focuses on the three enhancement techniques namely

A. Adaptive Histogram Equalization

B. Non-Local Filter Means

A. Adaptive Histogram Equalization:

This technique divides the image into tiles which are non-overlapping and performs histogram equalization on each tile, thereby improving the contrast of the image.

Let $I \in R^{a\times b}$ of size a×b size to be enhanced. It is divided into tiles and then stitched together to get an overall enhanced image.This algorithm is used to improve the contrast of the images.

Input: Ear image $I \in R^{a\times b}$ of size $a \times b$.

Output: Enhanced Image $I \in R^{a\times b}$ of size $a \times b$.

1. Divide image I into tiles T_i, $i=1\ldots n$, each of the size $\alpha \times \beta$, where $\alpha \leq a, \beta \leq b$ and $n = \frac{a \times b}{\alpha \times \beta}$.
2. **for** i=1 to n do
3. Enhance the contrast of the tile T_i.
4. **end for**
5. Obtain I_e by stitching all the values of T_i s,$i=1\ldots n$. Bilinear interpolation is used to remove all artificial induced boundaries that may occur whiling stitching T_i, $\forall\ i$ with its neighbours.
6. Return I_e.

B. Non Local Means Filter:

This technique is used for removing noise from the image. It assumes that for every small window in the image, many similar windows can be found in the image which can be used for noise reduction. This algorithm is used for reducing the noise in the images.

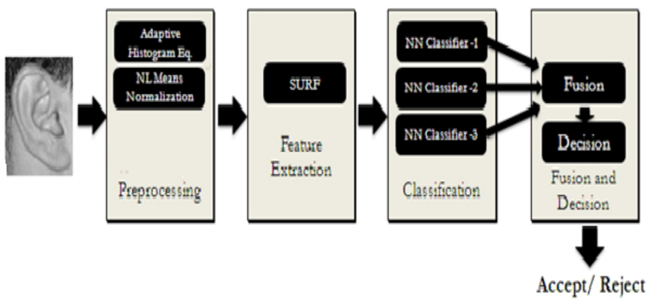

Figure 1. Basic Structure of the System

Input: Ear image $I \in R^{a\times b}$ of size $a \times b$.

Output: Enhanced Image $I \in R^{a\times b}$ of size $a \times b$.

1. **For all pixels** $p = (x,y) \in I, x \in [1 \ldots a], y \in [1 \ldots b\,]$ ***do***
2. $I_d(p) = \sum_{z \in I} w(p,z) I(z)$

   Where $w(p,z) = \frac{1}{Z(p)} e^{\frac{-G\sigma||I(\Omega p) - I(\Omega z)||_{22}}{h_2}}$

   $Z(p) = \sum_{z \in I} e^{\frac{-G\sigma||I(\Omega p) - I(\Omega z)||_{22}}{h_2}}$

   And $G\sigma\ ||.||^2_2$ denotes the Gaussian weighted Euclidian's Distance.
3. **end for**.
4. Return I_d.

2. FEATURE EXTRACTION:

SURF is used to extract the features of the ear in our technique. SURF can catch the properties of spatial organization. It uses salient feature points to represent an image. With every point, there's a descriptor vector of 128 dimensions. These features are then fused to get a fused template.

Input: Two sets of descriptor vectors $D^1 = \{D^1_1, D^1_2, \ldots, D^1_n\}$ and $D^2 = \{D^2_1, D^2_2, \ldots, D^2_n\}$ corresponding to n and m key points of the images I_1 and I_2 to be matched and matching threshold is $\tau\ \varepsilon\ (0,1)$.

Output: Matching score N stating number of matching descriptor vectors in two images.

1. N ← 0
2. **for** i=0 to n do
3. **for** j=0 to n d
4. Compute *distance[j]=Euclidian _Distance(D^1_i, D^2_j);*
5. **end for**
6. Compute [Sort_Dist,Orginal_Index]=Sort(distance) where Sort(.) is a function which sorts distance array in ascending order and returns sorted distance values in array Sort_Dist and their corresponding original index values of the distance array in an Original_Index.
7. **If** $\frac{Sort\_Dist[1]}{Sort\_Dist[2]} \leq \tau$ then
8. Descriptor vector D^1_i of Image I_1 matches to the descriptor of $D^2_{Original\_Index[1]}$ of Image I_2 where Original_Index[1] is the index of the matched descriptor I_2.
9. N ← N+1
10. **end if**
11. **end for**
12. Return matching score N.

C. ARTIFICIAL NEURAL NETWORKS

Here the ear prediction problem is tackled with the help of artificial neural network. An ANN is a mathematical representation of the human neural architecture, representing its "learning" and "generalization" capabilities.

Neural networks basically are a system of interconnected neurons which consists of 3 parts (as shown in figure 2) they are:

1. Input Layer
2. Hidden Layers
3. Output Layer

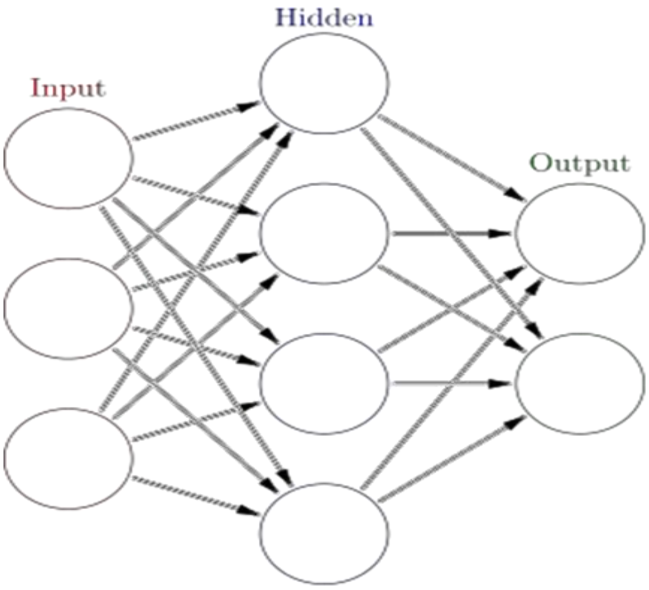

Figure 2. Basic Structure of the Neural Network

Feed forward neural network which is one of the first and simplest of neural network where information only moves in one direction i.e. forward and contains no cycles and loops.

There are two types of Feed forward neural networks:

A. Single-Layer Perception

B. Multi-Layer Perception

Multi-layer perception is fully connected i.e. output from each neuron either input or hidden is distributed to every other neuron in next layer. The number of hidden layers may vary. Back-Propagation is one of the most popular learning method used in multi-layer networks. In this method the output value is compared with some predefined correct value to calculate an appropriate value of some predefined error function. This computed value of error function is then fed back through the network. With the help of this information the algorithm alters the weights of each connection between neuron with the intension of reducing the value of error function. This process is repeated continuously until the value of error function is quite small. At that instance we can say that the network has learned certain function. Number of Hidden layers in the network plays an important role in accuracy for the network. There is no fix law for calculating the number of hidden layers a neural network should have. There are some empirically-derived rules-of-thumb, of these, the most commonly relied on is the optimal size of the hidden layer is usually between the size of the input and size of the output layers. Jeff Heaton [3].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

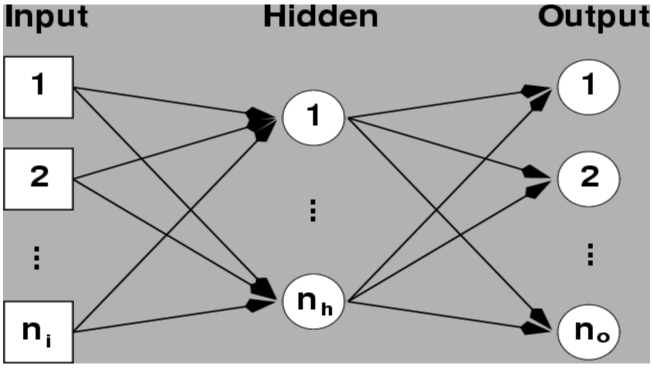

Figure 3. Schema of a Back-Propagation Neural Network (BPNN)

In Figure 3 the BPNN has 3 layers in which all the node in input layer are connected to ever other node in the hidden layer and all the nodes in hidden layer are connected to every other node in next layer (Hidden or output). All the connections between nodes are directed and no nodes from the same layer are connected. Each of these connections have weights and these weights are altered using the BPNN algorithm in the training (learning) process. Also it is important to choose the number of training samples. According to Baum and Haussler [2], to guarantee a level of performance on a set of test samples obtained from the same training data we can place a bound on the training samples.

V. CONCLUSION

In this paper the proposed method aims at improving the ear recognition and optimization system expected to have higher accuracy and optimized result. The system to be developed uses the combination of algorithm to meet its requirements. The Speed Up Robust Feature (SURF) algorithm follows nearest neighbor approach for optimization and then proceeds to follow different algorithm for normalization, feature extraction, template matching.

Our main job is

1. To provide security.
2. To obtain better accuracy and optimized results.
3. To get false accept and reject rate as low as possible.
4. To compare the algorithm with the system as whole.

VI. FUTURE SCOPE

We plan to create our own database of the students in our college for automatic attendance system. Ours being a developing nation has a major problem of proper identification of people in rural areas below poverty line. Ear recognition can be used in this sphere it being a non-intrusive and cheap mode of identification. In hospitals can be implemented to create medical history databases for all the people and use it across all the hospitals. It can also prove convenient in any emergency medical identification.

REFERENCES

[1] Burge, M., Burger, W., “Ear biometrics in computer vision,” In Proceedings of 15th International Conference on Pattern Recognition, 2000: 822-826W.

[2] E.B. Baum and D. Haussler, ``What size net gives valid generalization?," *Neural Computation*, vol. 1, no. 1, pp. 151-160, 1989.

[3] Jeff Heaton, Heaton Research, Inc. (25 November 2005)”,Introduction to neural networks with java,” ISBN-10: 097732060X, ISBN-13:978-0977320608.

[4] Mark Burge and Wilhelm Burger. “Ear biometrics in computer vision”. In Proceedings of International Conference on Pattern Recognition (ICPR' 00), volume 2, pages 822-826, 2000.

[5] David J. Hurley, Mark S. Nixon, and John N. Carter. “Force Field feature extraction for ear biometrics”. Computer Vision and Image Understanding, 98(3):491-512, 2005.

[6] Ping Yan and KevinW. Bowyer. “Empirical evaluation of advanced ear biometrics”. In Proceedings of International Conference on Computer Vision and Pattern Recognition-Workshop, volume 3, pages 41-48, 2005.

[7] L. Alvarez, E. Gonzalez, and L. Mazorra. “Fitting ear contour using an ovoid model”. In Proceedings of IEEE International Carnahan Conference on Security Technology (ICCST' 05), pages 145-148, 2005.

[8] Ping Yan and K.W. Bowyer. “Biometric recognition using 3D ear shape”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(8):1297-1308, 2007.

[9] Saeeduddin Ansari and Phalguni Gupta. “Localization of ear using outer helix curve of the ear”. In Proceedings of the International Conference on Computing: Theory and Applications (ICCTA' 07), pages 688-692, 2007.

[10] Li Yuan and Zhi-Chun Mu. “Ear detection based on skin-color and contour information”. In Proceedings of International Conference on Machine Learning and Cybernetics (ICMLC' 07), volume 4, pages 2213{2217, 2007.

[11] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. International Journal of Computer Vision, 1(4):321-331, January 1988.

Akshay Borse, pursuing Degree in Computer Science and Engineering From University of Pune, Modern Education Society’s College Of Engineering, Pune, India .

Ammen Siingh, pursuing Degree in Computer Science and Engineering From University of Pune, Modern Education Society’s College Of Engineering, Pune, India.

Arshad Attar pursuing Degree in Computer Science Engineering From University of Pune, Modern Education Society’s College Of Engineering, Pune, India.

Prof. Nuzhat F. Shaikh, obtained her Bachelors in Computer Engineering from University of Pune in 1997. She completed her M.E. in Computer Engineering from the University of Pune in 2004. She is currently a research scholar working for a Ph.D degree at Swami Ramanand Teerth Marathwada University, Nanded. At present, she is working as an Associate Professor in the Department of Computer Engineering at MES College of Engineering, Pune. Her research areas are Image processing, biometrics, Artificial Intelligence and Optimization of Neural Networks. She is the reviewer of many international journals and has published various technical papers at conferences and in reputed journals.