RESEARCH ARTICLE   OPEN ACCESS

3D Ear Recognition System Using Neural Network Based Self Organizing Maps

M.Sathish Babu[1], Assistant Professor

Email: sathish.cse24@gmail.com
Department of Computer Science and Engineering, Cape Institute of Technology.

ABSTRACT
In this manuscript we present a new 3D Ear recognition system using neural network based self organizing maps . In this manuscript we propose a new adaptive unimodal biometric framework based on self organizing maps for the recognition of individuals using ear. We show that the proposed framework helps to improve the performance and robustness of recognition when compared to some standard methods in literature. The major focus of our approach is to keep the framework adaptive and robust, thereby, capable of being used in a wide variety of environments. Moreover we also discuss some new directions on which SOM shall be effectively used in biometrics community. We show all our findings with experimental results. We apply the proposed method to a dataset contains 4000+ images and showthe results.
Index Terms - 3D Ear recognition system, Self Organizing Map(SOM).

I.INTRODUCTION

The increase of terrorism and other kinds of criminal actions, such as fraud in e-commerce, increased the interest for more powerful and reliable ways to recognize the identity of a person [1], [2]. To this end, the use of behavioral or physiological characteristics, called biometrics, is proposed. Biometrics is best defined as measurable physiological and or behavioral characteristics that can be utilized to verify the identity of an individual [1].

The recognition of individuals without their full cooperation is in high demand by security and intelligence agencies requiring a robust person identification system. Many biometric recognition algorithms have been proposed so far [3], [4], [5], [6], [7], [8]. Algorithms related to recognition of ear, hand geometry, iris, voice recognition have also been proposed (See Handbook of Biometrics [9]).

A typical biometric system usually consists of that specific biometric detection scheme followed by an extraction methodology (which shrinks the dimensionality of useful information) and then a classifier to make the appropriate decision.

A common approach for 3D ear recognition is the use of registration techniques to perform range image matching. The Iterative Closest Point (ICP) [10] algorithm, or one of its variants, is usually sought to accomplish this task. The Mean Squared Error (MSE), minimized during the convergence process, is then used to compute the similarity between two ear images [11], [12], [13], [14]. This approach can also be employed with deformation

techniques to model facial expressions, minimizing its effects on ear recognition [15], [16], [17].

For a survey of works related to 3D and multimodal 2D+3D ear recognition the reader is encouraged refer to [18], [19].

We present a framework for ear recognition using only 3D information (range images) as input. We provide a new approach that does not require a subjects cooperation called as SOM method. In our experiments we use the FRGC v2 database, the largest available database of 3D ear images, composed of 4,007 images from 466 different subjects [20].

II. OBJECT DETECTION/PREPROCESSING

We extract the regions of interest using a Haar like features based object detector provided by the open source project OpenCV library [21]. This form of detection system is based on the detection of features that display information about a certain object class to be detected. Haar like features encode the oriented regions in images whenever they are found, they are calculated similarly to the coefficients in Haar wavelet

transformations. These features can be used to detect objects in images, in this case the human ear. The Haar like object detector was originally proposed by Viola and Jones [22] and later extended by Lienhart and Maydt [23].

To create a ear detector we used 1000 positive ear samples and 2500 negative samples. The positive samples were scaled to the same size of 24x24; yielding the best and fastest results. The ear detector worked very well, detecting all ears, with a few false detections.

After the ear region is segmented, eight feature points are detected to extract rigid regions of the ear and improve the matching process; the inner right and left eye corners, the right and left nose corners, the nose tip and base, lip top and base.

Fig. 1. Segmented regions from a same ear: (1) the circular and (2) the elliptical areas around nose, (3) the upper head, including eyes, nose and forehead, (4) the entire ear region and (5) elliptic area around lips.

In this work, five regions of the ear are considered (see Fig. 1): (1) the circular and (2) the elliptical areas around nose, (3) the upper head, including eyes, nose and forehead, (4) the entire ear region and (5) elliptic area around lips.

III. PROPOSED APPROACH (SOME)

We use SOM for ear recognition and hence we call the approach as SOME. The set of input data in our case refers to the set of images that is used; the observations refer to the pixels present in each image. First we apply SOM ear separately. In this case for ear, the dimensionality of the input vector is 3600 (this is because of the normalized size of the ear image that is used −60x60 size). The output space is an array (for ear) of p by q neurons (nodes) topologically connected following a kind of geometrical rule (a rectangular topology has been used). In our case p=11 and q = 11 for ear. With the same setup, we do a supervised mode SOM analysis (where we use some images for training and some images for testing). In the end (SOME approach), In other words, in SOME we get multiple layers (as opposed to supervised SOM where there are only two layers X and Y). of ear layers and determine the optimum weightage for the recognition experiment under consideration. All these interesting experimental results obtained using SOM in unsupervised mode, supervised mode, super organized mode are explained in the next section.

IV.EXPERIMENTS & RESULTS

As mentioned earlier, this paper uses the Ear dataset obtained from [44]. There are 107 subjects. Each subject has 3 images for ear.

The first experiment which was performed was to find the total number of output nodes which are required. Unsupervised SOM was ran over the given 107 subjects related to ear dataset. In the plot shown in Figure.2 the background color of a unit corresponds to the number of samples mapped to that particular unit; one shall observe that they are reasonably spread out over the map (one unit is empty for ear; no samples have been mapped to them). The plot in Figure. 3 shows the mean distance of objects, mapped to a particular unit, to the codebook vector of that unit. A good mapping should show small distances everywhere in the map. These show that the numbers of output nodes which are chosen (11x11) are good enough for our purpose.
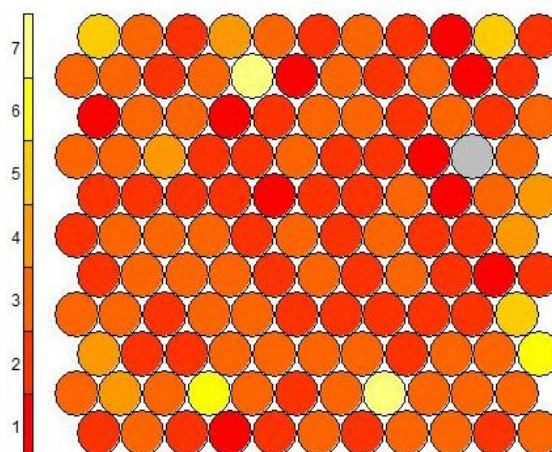

Fig. 2 Counts plot of the map obtained from the ear dataset. Empty units are depicted in gray. The color in each cell represents the number of ear
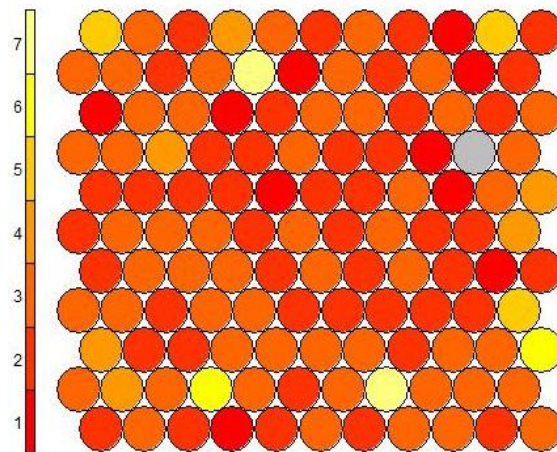
Fig. 3 Shows the quality of the mapping; the biggest distances between x and mi vectors are found in the bottom left of the map for ear

The second experiment which was performed was to do an exploratory analysis using unsupervised SOM. Figure.4   shows the mapping of images related to unsupervised SOM. Each color/shape in the figure is used to represent a particular subject. From the dataset, one shall infer that each subject has 3 ear images related to him which are more or less mapped into different unique cells. Figure 4 reveals this out clearly. For instance in Figure 4, if one looks at the first cell, approximately 3 similar units are mapped onto that cell for ear. The similar units indicate that they belong to the same subject. This explains that even without any training, unsupervised SOM was able to more or less grossly able to put the subjects into different cells. The error rate in grouping in this case was observed to be approximately 27% for ear (out of the 321 images of 107 subjects, 225 went into the appropriate cells which belonged to similar subjects and 96 images did not gets mapped properly).
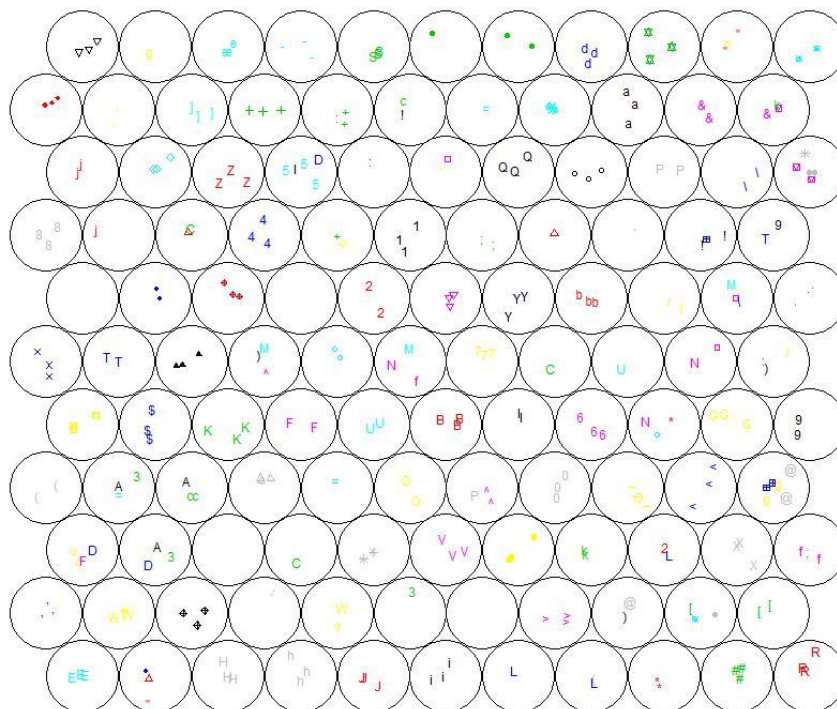

Figure.4 Mapping of the 107 Ear subjects in a eleven-by- eleven SOM

The third experiment that was used is to use the classifier information related to which image belonged to which subject using supervised SOM. In this experiment, the subject has been considered as the dependent variable (variable Y as explained in Section 3) and the pixel values of the image as the independent value (variable X as explained in Section 3). 1 random image from each subject has been chosen for training and the rest of the 2 images of each subject has been used for testing. The weights for X and Y has been varied with supervised SOM and the following characteristics as mentioned in Table 1 has been observed (the weights in a way indicate the relative strength between X and Y for recognizing a subject).

X Weightage	Y Weightage	Error Rate
0.9	0.1	23.5
0.8	0.9	19.8
0.7	0.3	17.2
0.6	0.4	16.6
0.5	0.5	14.7
0.4	0.6	12.9
0.3	0.7	11.2
0.2	0.8	8.6
0.1	0.9	7.9

Table 1. Error rate with supervised SOM by varying X and Y weights for ear

The above Table 1 shows that, if one uses the classification information also (using supervised SOM), then the recognition rate improves significantly (when compared to not using it - as earlier seen with unsupervised SOM). This is true for ear.

The fourth experiment that was done was related to super-organized SOM. We modeled ear related pixel values as the first layer and the class information as the second layer. A weight is associated to every layer to be able to define an overall distance of an object to a unit. We pose an optimization problem to optimize the weights in such a way that the recognition rate is the maximum. Interestingly, this also allows one to easily find out the dominant metric (ear - based on the one which takes a higher weightage). To begin with, we seeded the initial weights to be of extremely low (0) for ear. We noted down the results. We then optimized the weights for ear as explained above and observed the results. The experimental results are presented in Table 2. It seems that ear is a better metric and its seems to give a better recognition rate.

Ear Weightage
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

Table 2. Recognition rates ear weights using super SOM

The fifth experiment that was done was a comparative analysis of SOME with other methods related to multimodal biometrics involving ear. Table 4 shows the comparative results between Ear-PCA (Principal Component Analysis), -Ear-Sequential Float Feature Selection (SFFS) and Self Organizing Map for Ear (SOME).

The size of the training set varied from 1 to 2 images per person and the remaining of the images for each subject form the test set. For the PCA and the SFFS, the experiments that were conducted showed that all the training images during the training phase are classified correctly (Table 3). On the other ear, the SOME could not classify correctly all the training images. Furthermore, Figure. Table 3 shows a greater improvement in the performed experiment with SOME than PCA or SFFS when using one number of training samples for each person. Using SOME with one image per person during training phase gives 3.4% error recognition rate against 11.9% error recognition rate using the PCA, and 9.6% error recognition rate using the SFFS method.

Number of training images per person	Number of testing images per person	Training phase			Testing phase		
		PCA	SFFS	SOME	PCA	SFFS	SOME
1	2	0	0	4.6	11.9	9.6	3.4
2	1	0	0	4.1	9.2	8.7	2.6

Table 3. Test error recognition rate (%) with varying number of images per person

Table 3 shows that SOME can provide an improvement in error recognition rate when compared to the other approaches based on literature. Interestingly, self organizing maps shall also be used to address some interesting curiosities discussed in the biometrics community in a formal manner. For instance, there has been a curiosity/hypothesis which says that 'ear as a biometric does not change over age when compared to other biometrics like face. If one shall gather images of same subjects at different ages in a similar pose and background and do a supervised SOM across the different ages, one shall find out if ear has been consistently performing when compared to palm print or some other biometric. Most of the results used in this paper are obtained using an open source software framework named statistical R[45]. The archive of results and code used related to this paper is accessible at [46].

V.CONCLUSION

Neural Network based Self Organizing Maps has been used in this paper. The proposed approach SOME has been shown to perform well when compared to some standard methods from literature. This has been done by taking a standard dataset from literature.

REFERENCES

[1] Jain AK, Bolle R, Pankanti S , Biometrics: Personal identification in networked society. Kluwer Academic Publishers, Boston/ Dordrecht/ London, 1999.

[2] Wechsler H, Phillips JP, Bruce V, Folgeman Soulie F, Huang TS, Face recognition From theory to applications. ASI NATO series, vol 163, Springer, Berlin Heidelberg New York, 1997.

[3] Zhao WY, Chellappa R, Rosenfeld A, Philips PJ, Face recognition: A literature survey. UMD CfAR technical report CAR-TR-948, 2000.

[4] Turk M, Pentland A, Eigenfaces for face recognition. J Cognitive Neurosci 3(1):7186, 1991.

[5] Jain A, Flynn P, Ross AA, Handbook of Biometrics. Springer, Heidelberg, 2008.

[6] Wiskott L, Fellous JM, Kru ger N, von der Malsburg C, Face recognition by elastic bunch graph matching. IEEE T Pattern Anal 19(7):775779, 1997.

[7] Kotropoulos CL, Tefas A, Pitas I, Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions. Pattern Recogn 33(12):19351947, 2000.

[8] Penev PS, Atick J, Local feature analysis: A general statistical theory for object representation. NetworkComp Neural 7(3):477500, 1996.

[9] de Vel O, Aeberhard S, Line-based face recognition under varying pose. IEEE T Pattern Anal 21(10):10811088, 1999.

[10] P. J. Besl and N. D. McKay, A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, no. 2, pp. 239-256, 1992.

[11] K. I. Chang, K. W. Bowyer, and P. J. Flynn, Multiple nose region matching for 3D face recognition under varying facial expression. IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 10, pp. 1695-1700, 2006.

[12] J. Cook, V. Chandran, S. Sridharan, and C. Fookes, Face recognition from 3d data using

iterative closest point algorithm and gaussian mixture models in Proc. Int'l Symposium 3D Data Processing, Visualization and Transmission, 2004, pp. 502-509.

[13] T. Faltemier, K. W. Bowyer, and P. J. Flynn, A region ensemble for 3d face recognition. IEEE Trans. Inf. Forensics Security, vol. 3, no. 1, pp. 62-73, 2008.

[14] X. Lu, A. K. Jain, and D. Colbry, Matching 2.5D face scans to 3D models. IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 1, pp. 31-43, 2006.

[15] J. Huang, B. Heisele, and V. Blanz, Component-based face recognition with 3d morphable models. in Proc. Int'l Conf. Audio and Video-based Biometric Person Authentication, 2003, pp. 27-34.

[16] I. Kakadiaris, G. Passalis, G. Toderici, N. Murtuza, and T. Theoharis, Three-dimensional face recognition in the presence of facial expression: An annotated deformable model approach. IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 4, pp. 640-649,2007.

[17] X. Lu and A. K. Jain, Deformation modeling for robust 3D face matching. in Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 2, 2006, pp. 1377-1383.

[18] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, 2D and 3D face recognition: A survey. Pattern Recognition Letters, vol. 28, no. 14, pp. 1885-1906, 2007.

[19] K. W. Bowyer, K. Chang, and P. J. Flynn, A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. Computer Vision and Image Understanding, vol. 101, pp. 1-15, 2006.

[20] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, Overview of the face recognition grand challenge. in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005, pp. 947-954.

[21] CVlibrary, http://sourceforge.net/rojects/opencvlibrary/

[22] Viola P, Jones M, Rapid object detection using boosted cascade of simple features. In: Proceedings of IEEE Computer Vision and Pattern Recognition, 2001.

[23] Lienhart R, Maydt J, An extended set of Haar-like features for rapid object detection. In: Proceedings of IEEE International Conference on Image Processing, pp. 900903, 2007.

[24] L. Silva, O. R. P. Bellon, and K. Boyer, Robust Range Image Registration Using Genetic Algorithms and the Surface Interpenetration Measure, ser. Machine Perception and Artificial Intelligence. World Scientific Publishing, 2005, vol. 60.

[25] L. Silva, O. R. P. Bellon and K. L. Boyer, Robust range image registration using the surface interpenetration measure and enhanced genetic algorithms. IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, pp. 762-776, 2005.

[26] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing. Science, vol. 220, no. 4598, pp. 671-680, 1983.

[27] Jain AK, Bolle R, Pankanti S (eds) (1999) Biometrics: Personal identification in networked society. Kluwer Academic Publishers, Boston/Dordrecht/London

[28] Wechsler H, Phillips JP, Bruce V, Folgeman Soulie F, Huang TS (eds) (1997) Face recognition – From theory to applications. ASI NATO series, vol 163, Springer, Berlin Heidelberg New York.

[29] Zhao WY, Chellappa R, Rosenfeld A, Philips PJ (2000) Face recognition: A literature survey. UMD CfAR technical report CAR-TR-948

[30] Turk M, Pentland A (1991) Eigenfaces for face recognition. J Cognitive Neurosci 3(1):71–86

[31] Jain A, Flynn P, Ross AA, Handbook of Biometrics, Springer, Heidelberg (2008)

[32] Wiskott L, Fellous JM, Kru¨ ger N, von der Malsburg C (1997) Face recognition by elastic bunch graph matching. IEEE T Pattern Anal 19(7):775–779

[33] Kotropoulos CL, Tefas A, Pitas I (2000) Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions. Pattern Recogn 33(12):1935–1947

[34] Penev PS, Atick J (1996) Local feature analysis: A general statistical theory for object representation. Network–Comp Neural 7(3):477–500

[35] de Vel O, Aeberhard S (1999) Line-based face recognition under varying pose. IEEE T Pattern Anal 21(10):1081–1088

[36] Z. Mu, L. Yuan, Z. Xu, D. Xi, S. Qi, Shape and structural feature based face recognition, in: Sinobiometrics, LNCS, vol. 3338, 2004, pp. 663–670, 2004.

[37] DeVel O, AeberhardS,Line-based face recognition under varying pose. IEEE T Pattern Anal, 21(10):10811088, 1999.

[38] G. Lu, D. Zhang, K. Wang, Palmprint recognition using eigenpalms features, Pattern Recognition Letters, 24 (9–10), 1473–1477, 2003.

[39] X. Wu, D. Zhang, K. Wang, Fisherpalms based palmprint recognition, Pattern Recognition Letters, 24, 2829–2838, 2003.

[40] W.K. Kong, D. Zhang, W. Li, Palmprint feature extraction using 2-D Gabor filters, Pattern Recognition, 36 (10),2339–2347, 2003.

[41] Rahman M M, Ishikawa S (2005), Proposing a passive biometric system for robotic vision. In: Proc. 10th International Symp. on Artificial Life and Robotics,.

[42] Middendorf C, Bowyer K W (2008), Multibiometrics using face and ear. In: Handbook of Biometrics, pp. 315341. Springer, Heidelberg,.

[43] Yuan L, Mu Z -C, Xu X -N. (2007), Multimodal recognition based on face and ear. In:International Conference on Wavelet Analysis and Pattern Recognition, ICWAPR2007, vol. 3(2-4), pp. 12031207.

[44] Face and Palmprint Database Source Link, http://www4.comp.polyu.edu.hk/ csajaykr/

[45] R Project http://www.r-project.org/

[46] Archive of software code, dataset and experimental results http://www.capeitech.org/research/asr/SOME/