APPEARANCE BASED MULTIMODAL RECOGNITION SYSTEM WITHOUT SUBJECT'S COOPERATION USING A NEW METHOD NAMED MAG

Raja .A.S[1], JosephRaj .V[2]
[1]Research Scholar, Sathyabama University, Jeppiar Nagar, Chennai,Tamil Nadu, INDIA
[2]Professor, Kamaraj College, Thoothukudi, Tamil Nadu, INDIA
E-mail: [1]csehod@capeitech.org

Abstract

In this manuscript we present a new multimodal biometric system named MAG for the detection and recognition of face, ear and hand geometry. We propose a new feature extraction metric (Arity) and a modified classifier (Grep) as a part of MAG method. We use Multi-objective genetic algorithm for the combinatorial search space optimization of ear, face and hand geometry. We show that our proposed MAG method improves the performance and robustness of recognition when compared to methods proposed in literature. We apply the proposed method to a variety of datasets and show the results.

Keywords: Multimodal Biometrics, Multi Objective Genetic Algorithm, Arity feature extractor, Grep classifier (MAG)

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 co-operation is in high demand by security and intelligence agencies requiring a robust person identification system. Many face 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 multimodal system is a combination of fingerprint, face, iris and ear (for instance) or any other combination of biometrics. Multimodal biometrics can be used to overcome some of the limitations of a single biometric. For instance, it is estimated that 5% of the population does not have legible fingerprints [1], a voice could be altered by a cold and face recognition systems are susceptible to changes in ambient light and the pose of the subject.

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.

Chang et al.[10] used PCA on face and ear, with a manual land marking method. With the largest dataset of 111 subjects, they achieved a combined recognition rate of 90%.Rahman and Ishikawa[11] also used PCA for combining face and ear, they used profile images and manually extracted features. On a dataset of 18 subjects of profile face and ear, the recognition rate was 94.44%. Middendorff and Bowyer[12] used PCA/ICP for face/ear, manually annotating feature landmarks. On a 411 subject dataset they were able to achieve a best fusion rate of 97.8%. Yuan eral.[13] used s FSLDA (full-space linear discriminant analysis) algorithm on 75 subject database with 4 images each (USTB) and on the ORL database of 75 subjects, achieving a best recognition rate of 98.7%.

Here we provide a new approach for appearance-based multi-modal recognition that does not require a subject's cooperation called as MAG method. We propose a new feature extraction metric (Arity) and a modified classifier (Grep) as a part of MAG method. We use Multi-objective genetic algorithm for the combinatorial search space optimization of ear, face and hand geometry. We show that our proposed MAG method improves the performance and robustness of recognition when compared to methods proposed in literature. MAG stands for the 3 keywords that are part of this method - Multi-objective genetic algorithm (M), Arity (A) and Grep (G).

The remainder of this paper is organized as follows: In section 2 we discuss an object identification technique suitable for face, ear & hand geometry. In section 3 we discuss PCA. In this section we present our new feature extraction metric called Arity and compare the significance between these methods. In section 4 we present the modified quadratic classifier named Grep. In section 5 we discuss the multi objective problem formulation involving ear, face and hand geometric using multi objective genetic algorithm. In section 6 we present the overall MAG approach as a flow chart. In section 7 we discuss the results obtained using this MAG method. Paper concludes with conclusion and future direction.

II. OBJECT DETECTION

We extract the regions of interest using a Haar like features based object detector provided by the open source project OpenCV library [14]. 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 face, human hand geometry and the human ear. The Haar like object detector was originally proposed by Viola and Jones [15] and later extended by Lienhart and Maydt [16].

A. Face Detection

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

B. Ear Detection

To create the ear detector we also used 1000 positive samples and 2500 negative images. The positive images were scaled to a size of 16x24 to reflect the rectangular dimensions of the ear. The ear detector worked well with a few falsely detected ears, the problem was overcome by selecting the larger detected object.

C. Hand Geometry Detection

To create the hand geometry detector we also used 1000 positive samples and 2500 negative images. The positive images were scaled to a size of 16x24 to reflect the rectangular dimensions of the hand geometry. The hand geometry detector worked well with a few falsely detected hand geometry, the problem was overcome by selecting the larger detected object. Some of the sample images are shown in Fig. 1.

Fig. 1. Sample Home made dataset Images

III. FEATURE EXTRACTION

A. PCA

Let X be a d-dimensional feature vector. In our case, d is equal to the number of pixels of each face image. The high dimensionality of X is a well-known problem for the design of a good face recognition algorithm. Therefore, methods for reducing the dimensionality of this image space are required. To this end, principal component analysis (PCA) is widely used.

Principal component analysis [4, 17] is defined by the transformation:

$$y_i = W_{PCA}^t x_i \qquad [1]$$

where $x_i \in X \subseteq \Re^d, i = 1,...,n$ (n samples). WPCA is a d-dimensional transformation matrix whose columns are the eigenvectors related to the eigenvalues computed according to the formula:

$$\lambda e_i = Se_i \qquad [2]$$

where S is the scatter matrix (i.e. the covariance matrix):

$$S = \sum_{i=1}^{n}(x_i - m).(x_i - m)^t, m = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad [3]$$

This transformation is called Karuhnen-Loeve transform. It defines the d-dimensional space in which the covariance among the components is zero, because the covariance matrix is diagonal. The eigenvalues correspond to the variances of each component in the transformed space. After ordering the eigenvalues by increasing order, it is possible to consider a small number of principal components exhibiting the highest variance. The principal components of the transformed space are also called the most expressive features, and the eigenvectors related to the most expressive features are called eigenfaces.

B. Arity

We propose a new feature extraction metric and call it as Arity vector.

Let G = (V, E) be a directed graph without multiple edges that represents a image, where the vertices $v \in V$ denote pixels. Relationships between these entities are made up of directed edges $e \in E$ An ordered pair of pixels {i,j}| ij and i,j $\in$ V, is connected iff there is at least one path from vertex I to vertex j in G. A arity is given by n connected vertices and the way they are connected with each other. The particular coherence which is described by a pattern is always based on all edges that exist between n vertices. The entirety of all distinct n-vertex patterns in G is then given by $P^n = \{P_1^n, P_2^n, ..., P_i^n\}$ where P_i^n is the i-th pattern consisting of n vertices. Actually, each pattern P_i^n represents a set of isomorphic connected subgraphs which have the same structural properties and differ only in the participating vertices. Accordingly, a pattern P_i^n comprehends a set of instances, i.e., $P_i^n = \{P_{i,1}^n, P_{i,2}^n, ..., P_{i,j}^n\}$ and each instance is a unique subgraph $P_{i,1}^n(V_{i,j}^n, E_{i,j}^n)$ of G, with the subset of vertices $V_{i,j}^n \subseteq V$ and the subset of edges $E_{i,j}^n \subseteq E$ (Figure 1). The edges in $E_{i,j}^n$ are only incident to vertices in $V_{i,j}^n$ and we denote them as the intrinsic edges of the pattern instance $P_{i,j}^n$ Other edges, $e \in E \setminus E_{i,j}^n$ do not contribute to the coherence of the vertices. $V_{i,j}^n$ Moreover, these extrinsic edges are part of the environment of $P_{i,j}^n$ which describes how the pattern instance is embedded into the image.

Following the logic of, we denote the arity of a pattern instance $P_{i,j}^n$ as how essential for all connections within a image it is. To quantify this significance we eliminate all edges of a pattern instance (i.e., its intrinsic edges $E_{i,j}^n$ and measure how this affects the number of connected ordered pairs of vertices in the network. An ordered pair of vertices (i, j)|i ≠ j and i, j ($P_{i,j}^n$) as the fraction of those initially connected pairs of vertices in a network which become disconnected if the intrinsic edges of the pattern instance $P_{i,j}^n$ are removed from the network.

$$Arity(P_{i,j}^n) = 1 - \frac{N'}{N} \quad [4]$$

In Eq. 1 N is the total number of ordered pairs of vertices in a graph G = (V, E) that are connected by at least one directed path of any length. It is supposed that N > 0, i.e., there exists at least one edge in the network that links two different vertices. N' is the number of ordered pairs of vertices in the subgraph G' = (V, E') of G where E' = E/ $E_{i,j}^n$ Therefore, G' is the subgraph of G that results from removing the intrinsic edges of the pattern instance $P_{i,j}^n$ from G. The arity index of a pattern instance ranges between 0 and 1, whereas zero indicates that the removal of its intrinsic edges does not disconnect vertices within the network and one denotes the cases when no pair of vertices is connected any more.

Usually, several instances of a particular pattern can be found in a network. For estimating the arity of the pattern itself the impact of its representatives has to be considered. We find that the average arity index of all instances of a pattern reflects this appropriately and define

$$Arity(P_i^n) = Arity(P_{i,j}^n) = \frac{1}{j}\sum_{j=1}^{j} Arity(P_{i,j}^n) \quad [5]$$

as the arity index of a pattern $P_{i,j}^n$ that consists of J instances. With it Eq. 2 also states the arity of a randomly chosen instance of the pattern $P_{i,j}^n$

We use this arity vector and its average across the traing set as the basis. The pattern size that we choose is size 3, hence size of vector is 13

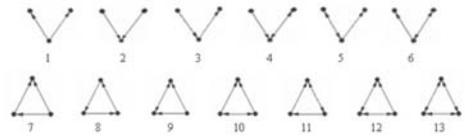

Fig. 2. Arity Vector

C. Comparison between PCA and Arity

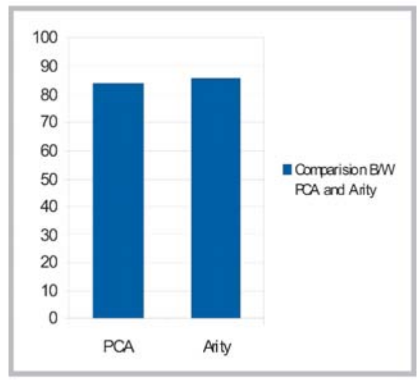

Fig. 3. Comparison between PCA and Arity

Based on a comparative experimental study with the sample dataset we find that Arity gives better performance when compared to PCA. It shall be intuitively argued that the computational effort required for Arity is less when compared to PCA.

IV. GREP CLASSIFIER

A quadratic classifier is used in machine learning and statistical classification to separate measurements of two or more classes of objects or events by a quadric surface.

Statistical classification considers a set of vectors of observations x of an object or event, each of which has a known type y. This set is referred to as the training set. The problem is then to determine for a given new observation vector, what the best class should be. For a quadratic classifier, the correct solution is assumed to be quadratic in the measurements, so y will be decided based on

$$X^T AX + b^T X + c \quad [6]$$

In the special case where each observation consists of two measurements, this means that the surfaces separating the classes will be conic sections (i.e. either a line, a circle or ellipse, a parabola or a hyperbola).

We use a slightly modified form of quadratic classifier by applying a kernel trick and refer it as Grep classifier by creating a longer measurement vector from the old one by adding all pairwise product of individual measurements. For instance, the vector

$$[x_1, x_2, x_3, ..., x_{13}] \quad [7]$$

would become

$$[x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2, ..., x_{13}^2] \quad [8]$$

V. MULTI-OBJECTIVE OPTIMIZATION

Multi-objective optimization (or programming), [18, 19] also known as multi-criteria or multi-attribute optimization, is the process of simultaneously optimizing two or more objectives subject to certain constraints.

We use a Multiobjective Optimization Evolutionary Algorithm by [20, 21].

Here μ1 refers to face, μ2 refers to ear and μ3 refers to hand geometry and our objective is to maximize all these μ1, μ2 and μ3. The pareto front point represent the recognition optimal basis combining all the 3 metrics. We find that this produces very significant recognition rates when compared to other methods discussed.

VI. MAG

The overall diagrammatic self-explanatory view of MAG method is described below:

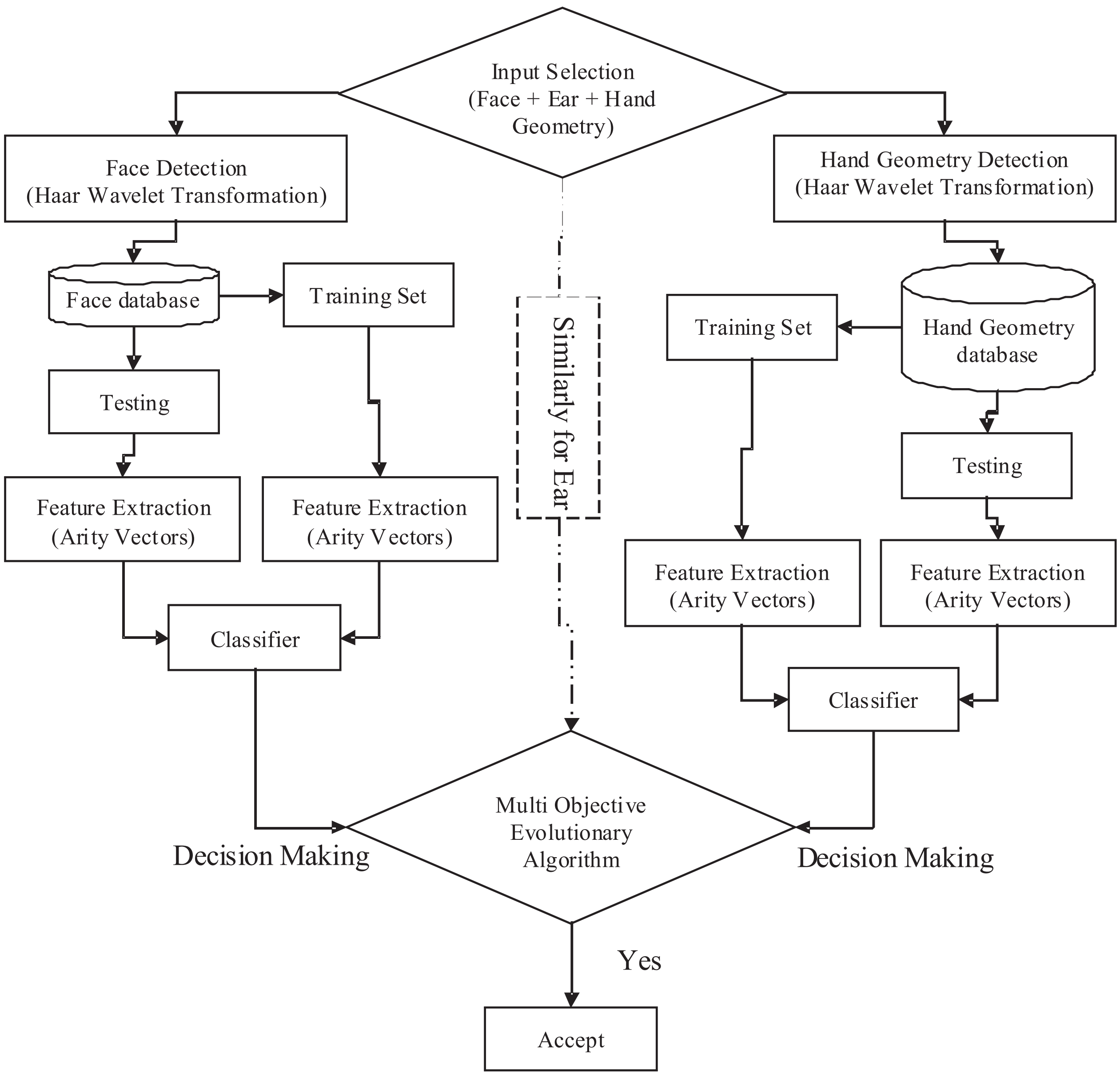

Fig. 4.

VII. RESULTS

We applied the MAG method to around 100 home made datasets and made lot of studies. The results are shown in diagrammatic form as below:

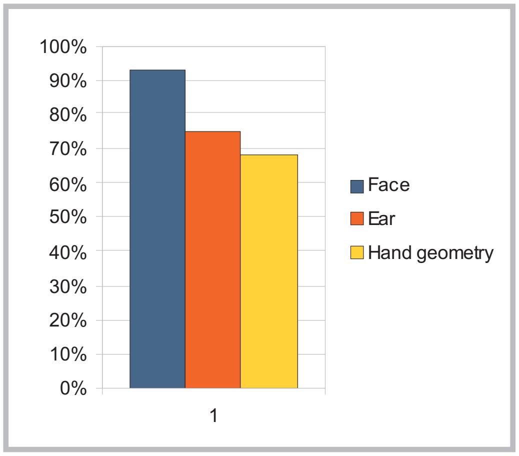

Fig. 5. Graph of unimodal recognition rates for face, ear and hand geometry

The unimodal recognition rates were not that good as shown in the above graph (although face recognition scores individually better). We did a weighted sum analysis of face/ear/hand geometry over the home made datasets and analyzed the results which are tabulated in the following figures.

Table 1. Combined face/ear normalized weighted sum recognition rates using Euclidean distance

Weight(face/ear)	DataSet
(1.0/0.0)	93.60%
(0.9/0.1)	95.40%
(0.8/0.2)	98.40%
(0.7/0.3)	91.60%
(0.6/0.4)	96.80%
(0.5/0.5)	94.80%
(0.4/0.6)	90.40%
(0.3/0.7)	91.90%
(0.2/0.8)	95.20%
(0.1/0.9)	85.50%
(0.0/1.0	78.50%

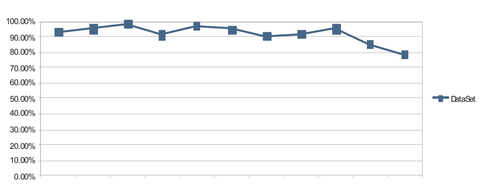

Fig. 6. Recognition rates for different face/ear weights using normalized sum

Table 2. Combined face/hand geometry normalized weighted sum recognition rates using Euclidean distance

Weight(face/hand)	DataSet
(1.0/0.0)	78.20%
(0.9/0.1)	81.00%
(0.8/0.2)	73.80%
(0.7/0.3)	85.20%
(0.6/0.4)	76.40%
(0.5/0.5)	79.80%
(0.4/0.6)	80.50%
(0.3/0.7)	79.40%
(0.2/0.8)	74.20%
(0.1/0.9)	70.60%
(0.0/1.0	69.40%

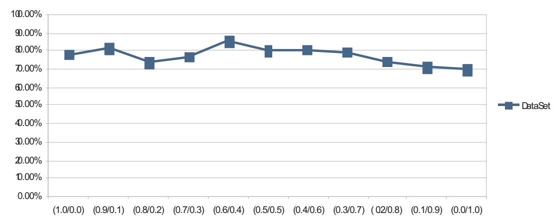

Fig. 7. Recognition rates for different face/hand geometry weights using normalized sum

Table 3. Combined ear/hand geometry normalized weighted sum recognition rates using Euclidean distance

Weight(ear/hand)	DataSet
(1.0/0.0)	65.30%
(0.9/0.1)	63.45%
(0.8/0.2)	66.50%
(0.7/0.3)	68.00%
(0.6/0.4)	71.50%
(0.5/0.5)	60.70%
(0.4/0.6)	67.60%
(0.3/0.7)	59.80%
(0.2/0.8)	58.50%
(0.1/0.9)	59.62%
(0.0/1.0	57.20%

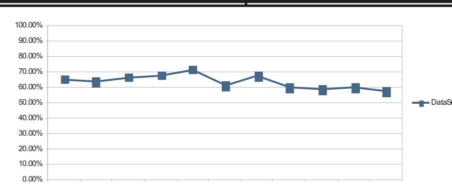

Fig. 8. Recognition rates for different ear/hand geometry weights using normalized sum

We finally used a multi objective genetic algorithm approach to find the pareto optimal recognition of face/ear and hand geometry, the results of which is as below:

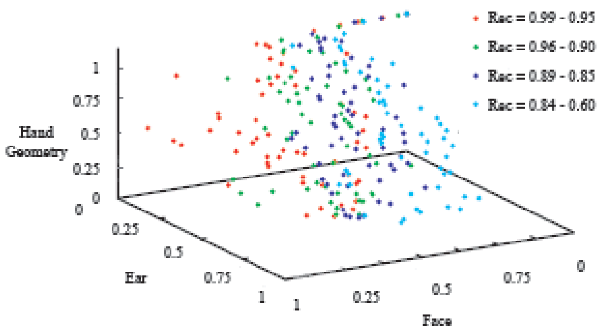

Fig. 9. A 3D-Pareto front using multi objective genetic algorithm

VIII. CONCLUSION

A new feature extraction metric (named Arity) has been proposed. We proposed a new feature extraction metric (Arity) and a modified classifier (Grep) as a part of MAG method. We used Multi-objective genetic algorithm for the combinatorial search space optimization of ear, face and hand geometry. We showed that our proposed MAG method improves the performance and robustness of recognition when compared to methods proposed in literature. The biometric importance of the suggested approach relies on its capacity to quantitatively evaluate the arity of each pixel group in the context of other pixels in a given image: that is how a given pixel group can be regulated by means of its reorganization, i.e., removing a pixel group and restoring the pixel group.

Future Directions of this work include using game theory, cooperative theories, for the multiobjective/mutimodal part. Different environmental scenarios with variability of environmental parameters like lighting, scale and facial expression are also interesting studies that are planned in near future.

REFERENCES

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

[2] 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

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

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

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

[6] 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

[7] 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

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

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

[10] Chang, K., Bowyer, K., Sarkar, S., Victor, B., 2003, Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics. IEEE Transactions on PatternAnalysis and Machine Intelligence 25(9), 1160–1165

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

[12] Middendorf, C., Bowyer, K.W., 2008, Multibiometrics using face and ear. In: Handbook of Biometrics, pp. 315–341. Springer, Heidelberg

[13] 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. 1203–1207

[14] http://sourceforge.net/projects/opencvlibrary/

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

[16] Lienhart, R., Maydt, J., 2002, An extended set of Haar-like features for rapid object detection. In: Proceedings of IEEE International Conference on Image Processing, pp. 900–903

[17] Duda RO, Hart PE, Stork DG., 2001, Pattern Classification, Wiley, USA

[18] Steuer, R.E., 1986, Multiple Criteria Optimization: Theory, Computations, and Application. New York: John Wiley & Sons, Inc. ISBN 047188846X.

[19] Sawaragi, Y. Nakayama, H. and Tanino, T., 1985, Theory of Multiobjective Optimization (vol. 176 of Mathematics in Science and Engineering). Orlando, FL: Academic Press Inc. ISBN 0126203709.

[20] Deb, K., 2001, Multi-Objective Optimization using Evolutionary Algorithms John Wiley & Sons Coello Coello, C. A.; Lamont, G. B. & Van Veldhuizen, D. A., 2007, Evolutionary Algorithms for Solving Multi-Objective Problems Springer

Raja .A.S Head, Department of Computer Since & Engineering - Cape Institute of Technology - Tirunelveli, has 5 years of Teaching Experience. His research interest is biometrics.