Gender Face Classification using Continous Wavelet Transform and Linear Discriminant Analysis

Nisar Hundewale
College of Computers and Information Technology
Taif University
Taif, Saudi Arabia
nisar@computer.org

ABSTRACT: *When we look at a face, we easily identify that person's gender, expression, personality, age, and charisma. Gender classification such as classifying human face is only challenging for computer, but even hard for human in some cases. In this paper, a new Novel approach is proposed to recognize gender from the face image. Continuous Wavelet Transforms are used for features selections for each face images of male and female. These selected features will be used to classify the face images of each Gender using Support Vector Machine (SVM). This Paper use ORL database contain 400 images include both Male and Female Gender. The experimental result shows that the proposed approach (Continuous wavelet Transform (1-D) and Linear Discriminate Analysis achieves excellent classification accuracy (100%).*

Keywords: Face Gender Classification, Feature Selection, Continuous Wavelet Transform (CWT), Linear Discriminate Analysis (LDA)

Received: 9 August 2011, Revised 25 October 2011, Accepted 3 November 2012

1. Introduction

Gender classification using facial images has been in the field of research now days and it is quite interesting. Humans are very good in differentiating the gender from facial images. Social Behavior and human interaction is mainly depending upon on the gender of the person with whom he/she they plan to interrelate. Luckily, human being has the unique capacity of classify gender analyzing simply ones face and exclusive character of conveying personality, emotions, age and lot of other vital information. Even if the face of the human is damaged, to find the gender symptoms, we can identify the gender with very high accuracy [1]. More recently automated gender classification from facial images has gained much interest in computer vision, machine language and Image processing community. The Rapid development progress of this gender classification research area is due to the fast growing field of Internet, electronic commerce, electronic banking systems, more human computer interaction, and demographic research, and security and surveillance applications. It can also bump up other important areas like Image /video indexing, retrieval, passive demographic data collection, vision based human monitoring , human robot interaction , face recognition, face detection , age ,traditional determinations are some other important application , where gender classification play a major role.

Automatic gender classification is more challenges now and it is mainly difficult because of the inherent variability of human faces due to different image formation process in terms of image quality, photometry, geometry, occlusion etc [1]. Gender classification has been studied less. Gender classification has been especially interesting for psychologists but automatic gender classification has applications also in other fields, for example, in demographic data collection [2]. Automatic gender classification is also a useful preprocessing step for face recognition since it is possible to halve, in case of equal amount of both genders, doing the face classification before the face recognition will make the face recognition process almost twice as fast. In addition, separate face recognizers can be trained for the genders and in this way increase the face recognition accuracy. This has been successfully experimented in facial expression recognition [3]. Erno Mäkinen and Roope Raisamo provide the guidelines for classification with automatic detection and aligned faces [4].

Many types of gender classification methods are available appearance – based method or holistic approach, geometric or feature based approach, hybrid approach [1]. In appearance based approach, the whole image or specific regions in a face images is used to generate feature vector. Feature based approach need to localize differential components such as eyes, nose, eyebrows etc. In hybrid approach perceives both local features and whole face. Many techniques have been taken to classify facial images based on gender. This paper works out on the particular approach using Continues Wavelet Transform (CWT) and Linear discriminate analysis for classifying the gender of the facial images.

To analyze all the features describing an image and to detect gender of the images, it is important to extract all the available gender information from the image. It can be helpful to analyze the image at different resolution levels. Wavelet transform is an ideal tool to analyze images of different gender. It discriminates among several spatial orientations and decomposes images into different scale orientations, providing a method for space scale representation. The general principles of wavelet transforms have been described elsewhere [5]. Wavelet functions [5] can be used to select the important features for gender classification. 1-D wavelet transform have been applied to gender images computerized schemes with varying success, but usage of 2-D is less. Many authors have developed computerized methods to classify gender face images. Our technique performs over well in images containing variations in lighting and facial expression, pose angles, aging effects etc. More over it is less time consuming process.

LDA is used as a tool for data classification, function approximation, etc. due to its generalization ability and has found success in many applications. LDA is a supervised learning algorithm. LDA searches for the project axes on which the data points of different classes are far from each other while requiring data points of the same class to be close to each other.LDA encodes discriminating information in a linearly separable space using bases that are not necessarily orthogonal.

In section 1, we introduce the goals of the paper. Section 2 describes the proposed technique. The Feature Selection using Continuous Wavelet Transform is presented in Section 3. Section 4 discusses classification and prediction. Finally, some experimental results with discussion and conclusions are given in section 5 and 6 respectively.

In this paper, Continuous Wavelet transform is applied for features selection thereby to reduce the processing time, and performs over well in images containing variations in lighting and facial expression, pose angles, aging effects of the gray scale images. Also, lower wavelet coefficients can eliminated to reduce the background noises in the images [8]. We employ a novel technique based on the application of one-dimensional continuous wavelet transform to perform feature selection. Later the LDA classifier is applied to make better classification among gender images.

2. Experimental Setup

In this section, we describe our experimental setup, which explain about facial dataset and the proposed novel technique with raw data, and analyze the informative features in images containing variations in lighting and facial expression, pose angles, aging effects etc and finding out the wavelet coefficient using Continuous wavelet transform in gender facial images. One dimensional wavelet transform performs better in images containing variations in lighting and facial expression, pose angles, aging effects etc. Moreover it is less time consuming process. Classification and Prediction is done with the Linear discriminate analysis. It classifies the gender as male and female with various constrains and do better prediction.

2.1 Dataset Description

The paper uses the image dataset called ORL Database. The ORL database totally consists of 400 gray scale images representing male and female gender. This images contains variations in lighting, facial expressions, pose, angles, age effects information. In this work, we collect 400 face images out of which 350 faces are male and rest 50 images are female.

2.2 The Proposed Technique

The proposed technique includes application of 1-D Continuous wavelet transform (CWT) on the facial images.

The resultant wavelet coefficients are sorted, and take 100 coefficients in file. During the training phase, the LDA is trained with the dataset that includes variables pertaining to the corresponding coefficients with label 0 for male and label 1 for female images. Similarly, during testing phase, we randomly select coefficient and apply the LDA classification model on the testing coefficients from the CWT of test image set. Finally, LDA prediction rate is calculated in terms of Mean Squared Error (MSE).Thus the classification of facial images is achieved in two main steps. In the first step, features are selected from the gender facial images. In the second step, the selected features with higher wavelet coefficients with label 0 or 1 are classified using LDA classifier.

The steps of the proposed technique are as follows:

Step 1: Read an image one by one.

Step 2: Convert the image into single dimensional array.

Step 3: Apply the 1-D Continuous wavelet transform (CWT).

Step 4: Take the coefficient of all images, which is consider along with the label (0 for male, and 1 for female)

Step 5: Repeat the Steps 1 though Step 4 for all the images.

Step 6: Split the data from Step 5 in training and testing set by random selection of rows.

Step 7: Train the LDA using the training data from Step 6.

Step 8: Test the LDA using the test data from Step 6.

Step 9: Calculate the Classification and prediction rate.

Figure 1. ORL Database of 400 images

3. Feature Extraction Using Wavelet Transform

The feature extraction method that we adopted is Continuous Wavelet Transform (CWT). In This Novel based Method, 1-D continuous wavelet transform allows an input image to be decomposed into a set of independent coefficients corresponding to each one dimensional wavelet basis. We use continuous wavelets to make no redundancy in the information represented by the wavelet coefficients, which leads to efficient representation. Also, it provides exact reconstruction of the original image. Wavelet coefficient represents the "*degree of correlation*" (or similarity) between the image and the mother wavelet at the particular scale and translation. Thus, the set of all wavelet coefficients gives the wavelet domain representation of the image.

After decomposition of the image, the details coefficients can be threshold.

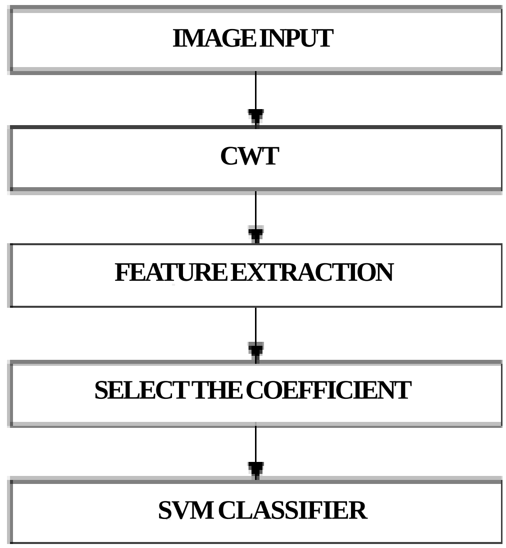

Figure 2. A Block Diagram for the Proposed Methods

3.1 Wavelet Transform

The wavelet transform [10] is a decomposition of an image onto a family of functions called a wavelet family, in which all of the basic functions (called wavelets) are derived from scaling and translation of a single function that is called the mother wavelet (or analyzing function). There exist many types of mother wavelets and associated wavelets [7]. Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals. Wavelet transforms are classified into discrete wavelet transformsn (DWTs) and continuous wavelet transforms (CWTs). Both DWT and CWT are continuous-time (analog) transforms. They can be used to represent continuous-time (analog) signals. CWTs operate over every possible scale and translation whereas DWTs use a specific subset of scale and translation values or representation grid.

3.2 Continuous Wavelet Transform (CWT)

A continuous wavelet transform (CWT) [11] is used to divide a continuous-time function into wavelet. The CWT has the ability to decompose complex information and patterns into elementary forms. The continuous wavelet transform possesses the ability to construct a time-frequency representation of a signal that offers very good time and frequency localization. The continuous wavelet transform of a continuous, square - integreable function $x(t)$ at a scale $a > 0$ and translational value $b \in R$ is expressed by the following integral:

$$X_W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \chi(t)\, \psi'\left(\frac{t - b}{a}\right) dt \qquad (1)$$

Where, $\psi(t)$ is a continuous function in both the time domain and the frequency domain called the mother wavelet and represents operation of complex conjugate. The main purpose of the mother wavelet is to provide a source function to generate the daughter wavelets which are simply the translated and scaled versions of the mother wavelet. To recover the original signal $x(t)$, inverse continuous wavelet transform can be exploited.

$$x(t) = \int_{0}^{\infty} \int_{-\infty}^{\infty} \frac{1}{a^2} (a, b) \frac{1}{\sqrt{|(a)|}}\, \tilde{\psi}\left(\frac{l - b}{a}\right) db\, da \qquad (2)$$

are Where, $\tilde{\psi}(l)$ is the dual function of $\psi(t)$. And the dual function should satisfy

$$\int_{0}^{\infty} \int_{-\infty}^{1} \frac{1}{|a^3|}\, \psi\left(\frac{c1 - b}{a}\right) \tilde{\psi}\left(\frac{l - b}{a}\right) db\, da\, \delta(t - t1) \qquad (3)$$

Sometimes, $\tilde{\psi}(t) = C_{\psi}^{1}\, \psi(t)$, Where,

$$C_{\psi} = \frac{1}{2} \int_{-\infty}^{+\infty} \frac{|\psi(\zeta)|^2}{|\zeta|}\, d\zeta \qquad (4)$$

is called the admissibility constant and ψ is the Fourier transform of ψ. For a successful inverse transform, the admissibility constant has to satisfy the admissibility condition: $0 < c_{\psi} < +\infty$.

It is possible to show that the admissibility condition implies that $\psi(0) = 0$, so that a wavelet must integrate to zero [8].

The advantage of using wavelet-based coding in image compression is that it provides significant improvements in picture quality at higher compression ratios over conventional techniques. Since wavelet transform has the ability to decompose complex information and patterns into elementary forms, it is commonly used in acoustics processing and pattern recognition. Edge and corner detection, partial differential equation solving, transient detection, filter design, Electrocardiogram (ECG) analysis, texture analysis and business information analysis. Continuous Wavelet Transform (CWT) is very efficient in determining the damping ratio of oscillating signals (e.g. identification of damping in dynamical systems). CWT is also very resistant to the noise in the signal.

4. Classification and Prediction

4.1 Fuzzy c – means

The Fuzzy C-means algorithm [13], also known as fuzzy ISODATA, is one of the most frequently used methods in pattern recognition. Fuzzy C-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. It is based on the minimization of objective function to achieve a good classification. "*J is a squared error clustering criterion, and solutions of minimization are least squared error stationary point of j*".

$$j_m = \sum_{i=1}^{k} \sum_{j=1}^{c} u_{ij}\, |x_i - c_j|^2 \qquad (5)$$

Where '*m*' is any real number greater than 1, is the degree of membership of in the cluster '*j*', is the d-dimensional measured data, is the dimension center of the cluster and is any norm expressing the similarity between any measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of member ship and the cluster centers.

$$u_{ij} = \frac{1}{\sum_{k-1}^{c} \left[\frac{||x_i - c_j||}{||x_i - c_k||} \right]^{\frac{2}{m-1}}} \qquad (6)$$

The iteration will stop when

$$max_{ij} - (|\, u_{ij}^{k+1} - u_{ij}^{k} |) < \in \qquad (7)$$

Where $\in$ is the termination criterion between 0 & 1, whereas *k* is the iteration steps. This procedure converges to a local minimum or a saddle point of j_m.

4.2 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) [6], [7], [12] is used for classification and dimensionality reduction. Linear Discriminant Analysis easily handles the case where the within-class frequencies are unequal and their performances have been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. LDA providing better classification compared to Principal Components Analysis. The prime difference between LDA and PCA is that PCA does more of feature classification and LDA does data classification. In PCA, the shape and location of the original data sets changes when transformed to a different space whereas LDA doesn't change the location but only tries to provide more class separability and draw a decision region between the given classes.

The mathematical operations involved in LDA, the global feature preservation technique is analyzed here. The fundamental operations are:

1. The data sets and the test sets are formulated from the patterns that are to be classified in the original space.

2. The mean of each data set μ_i and the mean of entire data set μ are computed.

$$\mu = \sum_i p_i \mu_i \tag{8}$$

where p_i is priori probabilities of the classes.

3. Within-class scatter S_w and the between-class scatter S_b are computed using:

$$S_w = \sum_i p_j * (cov_j) \tag{9}$$

$$cov_j = \prod_i (x_j - \mu_i) \tag{10}$$

Note that S_b can be thought of as the covariance of data set whose members are the mean vectors of each class. The optimizing criterion in LDA is calculated as the ratio of between-class scatter to the within-class scatter. The solution obtained by maximizing this criterion defines the axes of the transformed space.

The LDA can be a class dependent or class independent type. The class dependent LDA requires L-class L separate optimizing criterion for each class denoted by C_1, C_2, C_L and that are computed using:

$$C_j = (cov_j)^{-1} S_b \tag{11}$$

The transformation space for LDA, WLDA is found as the Eigen vector matrix of the different criteria defined

5. Results and Discussion

We used a set of 400 images of male and female that belongs to ORL Database. All the experiments are carried with low resolution images and no cross validation. Its grey scale images. The algorithms were all coded in MATLAB. The average size of training set is 200 male and female images and the average set of testing is 200 images of male and female. Here the continuous wavelet transform is applied to take the coefficients. The Figures 3 and 4 shows the Coefficient of Male and Female Gender. Later LDA is used to separates two or more classes of objects or events such as male and female.

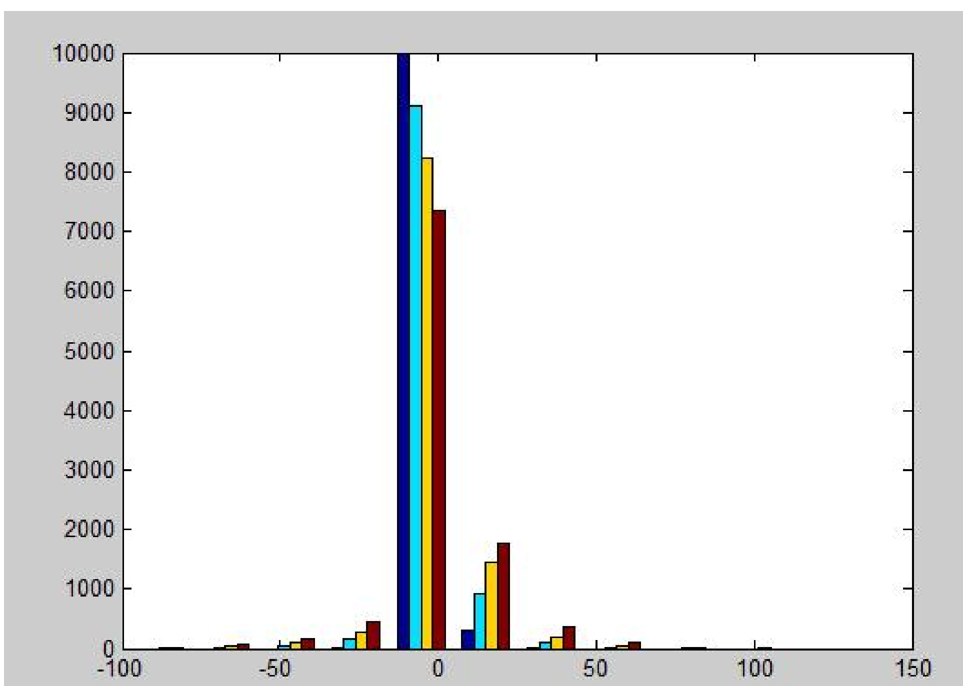

Figure 3. The CWT Coefficient of Male Gender

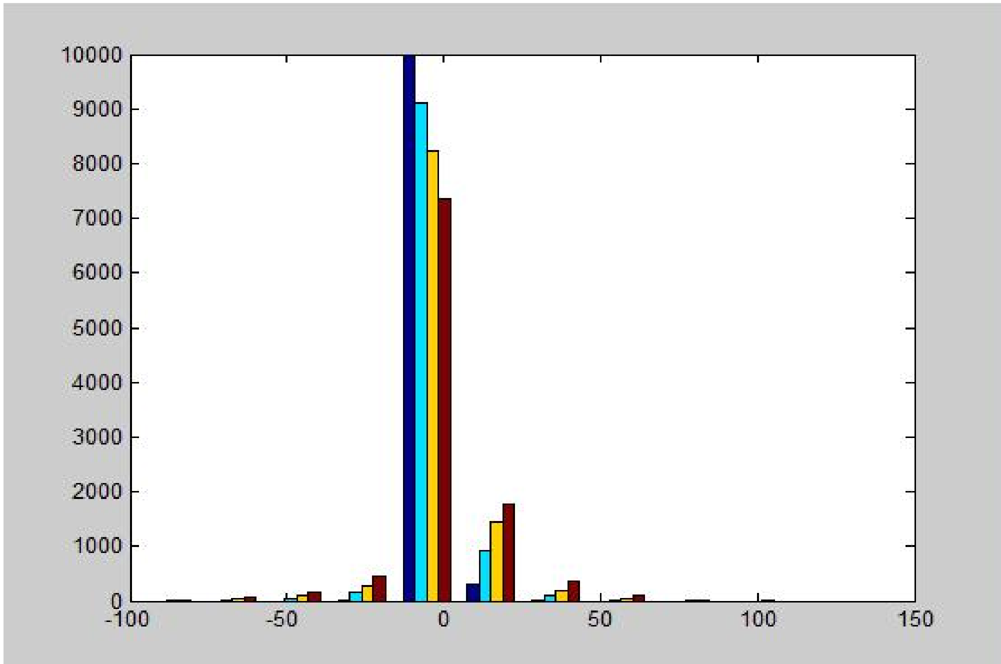

Figure 4. The CWT Coefficient of Female Gender

In Table 1, we present the comparison of techniques, authors to achieve the same goal as this paper. The classification rate of the various techniques are Tamura et al. (1996) is 93 % Using neural networks ,Wiskott et al(1997)is 95.3% using genetic algorithms, Jain and Huang shows the percentage of 99.3% using ICA/LDA. Costene et al. (2002) shows the percentage. The classification rate of our technique is superior to the rest up to techniques. Furthermore there is 0% error rate in our proposed technique. Overall speaking, the proposed novel technique outperforms other techniques in terms of specificity and sensitivity.

Authors	**Methods**	**Percentage**
Tamura et al (1996) [2]	Neural Networks	93.00
Wiskott et al (1997) [2]	Genetic Algorithms(GA)	95.30
Jain et al (2004) [2]	(ICA) / LDA	99.30
Costen et al (2004) [2]	SVM	94.42
Sun et al (2006) [2]	SOM	95.75
Lian et al (2006) [2]	SVM	96.75
Baluja et al (2007) [2]	Adaboost classifier	93.00
The Proposed Method	CWT (1-D), LDA	100.00

Table 1. Prediction Performance Comparison of Related and Proposed Techniques

The classification rate of our technique is superior to the rest up to techniques. Furthermore there is 0% error rate in our proposed technique. Overall speaking, the proposed novel technique outperforms other techniques in terms of specificity and sensitivity.

6. Conclusion and Future Work

An original analysis of algorithm to classifying the gender of male and female is distinguish and computed and verified. The

Algorithm	Percentage
CWT + FUZZY -C- MEAN	43
CWT + LDA	100

Table 2. Prediction Performance Comparison of Proposed Techniques

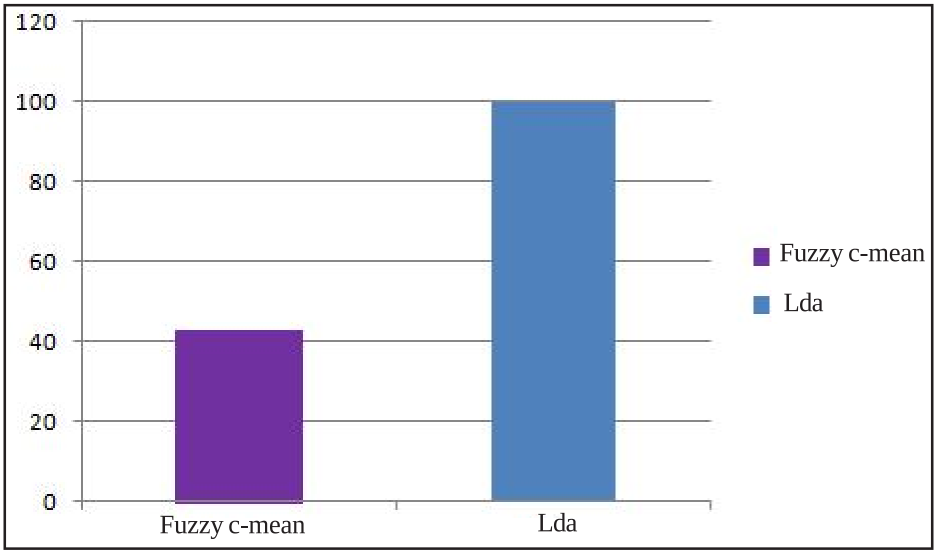

Figure 5. Prediction Performance Comparison plot for Proposed Techniques

images contain variations in lighting and facial expression, pose angles, aging effects etc and finding out the wavelet coefficient using Continuous wavelet transform in gender facial images. CWT is constructed to define the feature of the face gender. Linear Classifier is developed for classification. After Applying CWT, The resulted Coefficient is the binary data: 0 for Male Gender and 1 for Female Gender. The Linear Discriminate classifier exhibits superior efficacy of pattern recognition, in comparison to other classifier. The classification rate is raised up to 100% and there is no error rate, if CWT and LDA are both employed. Finally, LDA shows the lowest MSE in Classification and Prediction, and it is also compared with fc- mean technique, which show poor result. The computational effort for achieving 100% accuracy is due to minimal retrieval of features and low resolution images, training and testing strategy.

In Future, We have to develop the Method to classification the face gender and to adjust or identify the Threshold automatically using a Novel Clustering Algorithms.

References

[1] Michael, C., Mangini. (2004). Irving Biederman. Making the ineffable explicit: estimating the information employed for face classifications. *Cognitive Science* 28, 209–226.

[2] Jain, A., Huang, J., May 2004. Integrating independent components and linear discriminate analysis for gender classification. *In:* Proc. International Conf. on Automatic Face and Gesture Recognition (FGR'04), p. 159–163.

[3] Saatci, Y., Town, C. (2006). Cascaded classification of gender and facial expression using active appearance models. *In:* Proc. 7th Internat. Conf. on Automatic Face and Gesture Recognition (FGR'06), April, p. 393–400.

[4] Erno Mäkinen, (2008). Roope Raisamo. An experimental comparison of gender classification methods. *Pattern Recognition Letters* 29, 1544–1556.

[5] Daubechies I. (1992). Ten Lectures on Wavelet, SIAM, p. 167-213, Philadelphia.

[6] Fahim Mannan, 260 266 294, School of Computer Science, McGill University, Classification of Face Images Based on Gender using Dimensionality Reduction Techniques and SV. www.cim.mcgill.ca/~fmannan/comp652/comp652_report_FM.p df

[7] Samarasena Buchala, Neil Davey, Ray, J.Frank, Tim, M.Gale. Dimensionality Reduction of Face Images for Gender Classification. www.informatik.unitrier.de/~ley/db/.../a.../Frank:Ray_J=. htm.

[8] Samarasena buchala, tim m.gale, neil davey, ray j.frank, kerry foley, global and feature based gender classification of faces: a comparison of human performance and computational models. citeseerx.ist.psu.edu/viewdoc download?doi=10.1.1.59.rep.

[9] en.wikipedia.org/wiki/Wavelet

[10] en.wikipedia.org/wiki/Wavelet transform

[11] en.wikipedia.org/wiki/Continuous_wavelet_transform.

[12] en.wikipedia.org/wiki/Linear discriminate analysis.

[13] Hasassnpour, H., Dehghan, H. (2010). Neonate facial gender classification using PCA and fuzzy clustering, Biomedical Engineering (ICBME), 17th Iranian Conference.

Author Biography

Dr. Nisar Hundewale received his Ph.D. in Computer Science from Georgia State University, USA. He has worked at National Institutes of Health (NIH), USA, as post-doctoral Fellow. Currently, he is an Assistant Professor at Taif University. His research interests are Algorithms, Machine Learning, Bioinformatics, Distributed Computing, and Networking.