Wavelet Based Face Recognition using ROIs and k-NN

K M Poornima
JNN College of Engineering
Shimoga, Karnataka

Ajit Danti
JNN College of Engineering
Shimoga, Karnataka

S K Narasimhamurthy
Kuvempu University, Shankaraghatta Shimoga, Karnataka

Abstract- In this paper, human face recognition of still images has been proposed. The proposed system involves five steps: face detection by AdaBoost face detector, region of interest (ROI) extraction, feature extraction using discrete wavelet transform (DWT), dimensionality reduction by employing independent component analysis (ICA) and classification using k-Nearest Neighborhood (k-NN) classifier. Experiments were conducted on "Faces94" database by choosing 40 classes where, each class contains 5 images. The proposed system exhibits a recognition rate of about 83.5%.

Keywords – Face detection, ROI, DWT, ICA, Face Recognition.

I. INTRODUCTION

Face recognition is one of the most successful areas of research in image analysis and computer vision. It is an interesting and challenging problem of Pattern recognition and also one of the biometric systems. A human face provides ample of facial information that could make a face recognition system work effortlessly. But, building a computer algorithm for machines to work similar to human ability in recognizing a human face is a very difficult task. Although many techniques have been proposed by researchers, no technique could achieve full accuracy.

Face recognition is one of the emerging and the most successful areas of research from the past several years. One of the first researchers on face recognition was Woodrow W. Bledsoe in 1960 who worked on AI-related contracts. A primary application of face recognition is authentication to safeguard the data, which was used in biometrics. Face recognition ranges from still images to running video sequences. Recognition of faces in still images presents many technical challenges, and these are overcome by video sequences. There evolved many techniques and technologies for effective recognition of face but could not achieve 100% efficiency. Evaluating face recognition system is application-specific. It is difficult task to have a technique that works well for several applications. The applications of face recognition technology include law enforcement, Security/Counter terrorism, Legislature, Gaming Industry, Correctional institutions/ prisons and others.

Face recognition methods include Holistic approaches, Feature Based Method (structural matching), Hybrid Method, Multimodal Method and Spatiotemporal Method. Holistic approaches [8] attempt to identify faces using global representations, i.e., descriptions based on the entire image rather than on local features of the face. The main advantage is that they do not destroy any of the information in the images by concentrating on only limited regions or points of interest, but it is computationally expensive and do not perform effectively under large variations in pose, scale and illumination, etc. Holistic method can be further categorized. Principal-component analysis (PCA) is a successful low dimensional reconstruction of faces using KL or PCA projections. Many face recognition techniques have been developed using PCA: eigenfaces, which use a nearest neighbor classifier; feature-line-based methods, which replace the point-to-point distance with the distance between a point and the feature line linking two stored sample points; Fisherfaces which use linear/Fisher discriminant analysis (FLD/LDA); SVM methods, which use a support vector machine as the classifier and higher order statistics, independent-component analysis (ICA) is argued to have more representative power than PCA, and hence may provide better recognition performance than PCA.

Feature-based approaches [8] (Structural matching), first process the input image to identify and extracts facial features such as the eyes, mouth, nose, etc., as well as other fiducially marks, and then compute the geometric relationships among those facial points, thus reducing the input facial image to a vector of geometric features. Standard statistical pattern recognition techniques are then employed to match faces using these measurements. In structural matching methods, earlier methods used the width of the head, the distances between the eyes and from the eyes to the mouth, etc. or the distances and angles between eye corners, mouth extrema, nostrils, and chin top. One such method is Hidden Markov Model (HMM) [2] which uses strips of pixels that cover the forehead, eye, nose, mouth, and chin. The advantage is, it is relatively robust to position, invariant to size, orientation and/or lighting including the compactness of representation of the face images and high speed matching. But this method drawbacks in automatic feature detection and make arbitrary decisions about which features are important.

In Hybrid Method [8], system uses both local features and the whole face region to recognize a face, a machine recognition system should use both. The modular eigenfaces approach by Pentland [6], uses both global eigenfaces and local eigenfeatures. The use of hybrid features by combining eigenfaces and other eigenmodules is explored. Though the holistic approach provides a quick recognition method, the discriminant information that it provides may not be rich enough to handle very large databases. This insufficiency can be compensated for by local feature methods.

Multimodal Method [8] is used for Video based face recognition. This method performs the faceless identification by making use of multimodal cues that is, analyzing body motion characteristics. Moreover, using multimodal cues offers a comprehensive solution to the task of identification that might not be achievable by using face images alone. This system works on video and audio based person authentication. Spatiotemporal Method makes use of both spatial and temporal information. Although it is best to use both temporal and spatial information for face recognition, existing spatiotemporal methods have not yet shown their full potential.

The rest of the paper is organized as follows. Proposed system architecture is given in section II. Experimental results are presented in section III. Concluding remarks are given in section IV.

II. Proposed System Architecture

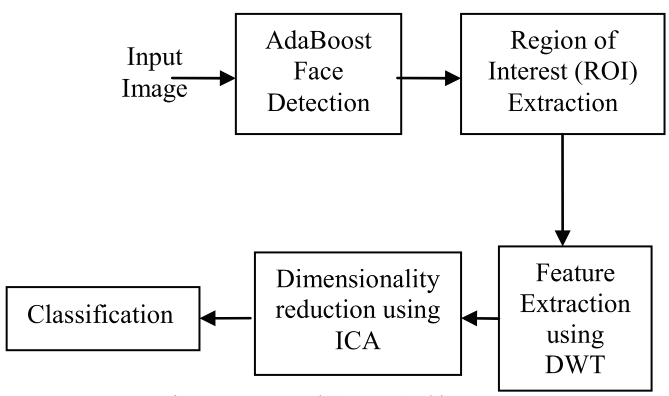

Figure 1. Proposed System Architecture

The proposed system architecture is shown in figure 1. The system is implemented in two stages: training and testing. In training stage first, images are read and AdaBoost face detector is used for detecting the face region. Region of interest (ROI) such as left eye, right eye, nose and mouth are extracted. To extract ROIs, the face image is divided in horizontal and vertical directions. Features are extracted from these ROI's using discrete wavelet transform (DWT). The obtained features size is reduced by employing dimensionality reduction methods such as independent component analysis (ICA). Same steps are repeated in testing stage. Finally, classification is done using k-nearest neighborhood (k-NN) classifier and recognition rate is evaluated.

A. AdaBoost Face Detection

AdaBoost (Adaptive boosting) is a machine learning algorithm which is formulated by Yoav Freund and Robert Schapire. AdaBoost, a Boosting algorithm is used to improve the performance of face detection. AdaBoost is adaptive in the sense that subsequent classifiers built are weak in favor of those instances misclassified by previous classifiers. In this work, AdaBoost face detector used is a "tree- based 20 X 20 gentle AdaBoost frontal face detector", which was invented by Rainer Lienhart. AdaBoost face detector [7] includes: sub-window, Haar-like feature, classifier and classifier cascade. Initially an image is divided into sub-windows on which the detection algorithm is applied. Haar-like features is used to obtain features from an image. Haar-like features have scalar values that represent differences in average intensities between two rectangular regions. Classifiers assigns each input value to a given set of classes, the classes are “Face” and “Not-Face”. Threshold of the sums and difference of rectangular regions of data produced by any feature detector is used by classifiers, which includes Haar wavelets of rectangular gray-scale image values. Classifier Cascade is an algorithm for rapid detection in a sub-window, where a positive result from a classifier triggers the next classifier in line. AdaBoost constructs a “strong" classifier as a linear combination of many “weak" classifiers. Figure 2 shows the basic scheme of AdaBoost. AdaBoost [7] is having two main goals: i) Selecting a few set of features which represents a possible face and, ii) Train a strong final classifier with a linear combination of these best features (Haar- like feature).

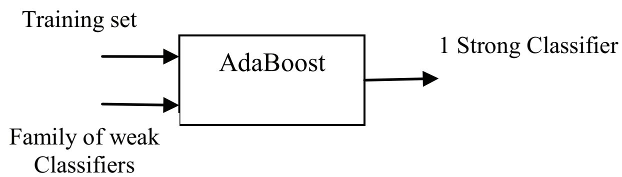

Figure 2. Basic scheme of AdaBoost.

B. Region of Interest (ROI) Extraction

Regions of interest (ROI) are extracted from the detected face by dividing it in horizontal and vertical directions and these ROIs are eyes, nose and mouth. Steps for dividing the detected face region are given below:

a. At first, image is divided into two equal halves in horizontal direction which results in two regions, Top region and Bottom region as shown in figure 3.
b. Top region is then divided horizontally and vertically into two equal halves such that, forehead (region 1) is separated from eyes part. Now, left eye (region 2) and right eye (region 3) are extracted.
c. Bottom region is divided vertically into 3 parts and only nose and mouth region is extracted (region 5).

Extracted ROI’s such as left eye, right eye, nose and mouth are shown in figure 4.

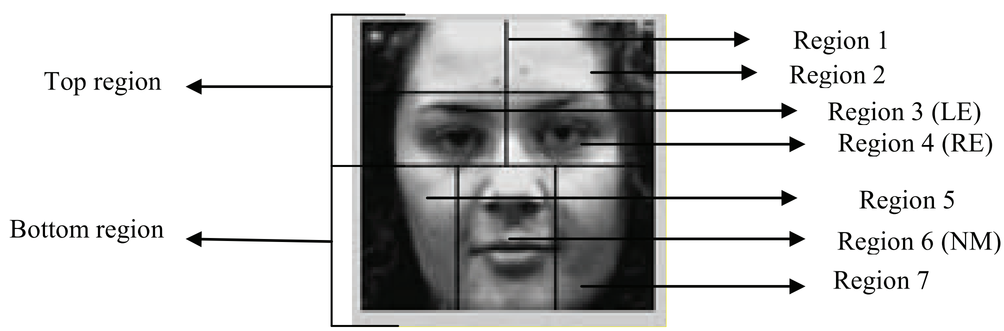

Figure 3. ROI Extraction

Figure 4. Extracted ROIs

C. Feature Extraction using DWT

Feature extraction can be defined as "extracting the information from the image which is most relevant for the classification purposes". A feature extraction method that proves to be successful in one application domain may not be very useful in another application. Wavelet transforms can be used for feature extraction. Wavelet transform of an image yields a set of coefficients which characterizes several physical features of the image. It is a simple and less computationally expensive technique giving more detailed information. One of the main advantages of wavelets is that they offer a simultaneous localization in time and frequency domain. Wavelets have great advantage of being able to separate the fine details in a signal i.e., a very small wavelets can be used to isolate very fine details in a signal, while very large wavelets can identify coarse details. Wavelet transform often compress or de-noise a signal without appreciable degradation.

Discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translations. The DWT [10] of a signal *x* is calculated by passing it through a series of filters. Initially, samples are passed through a low pass filter 'g' with impulse response resulting in a convolution as given in equation (1),

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]\, g[n-k] \qquad (1)$$

The signal is also decomposed simultaneously using a high-pass filter 'h'. We got the detail coefficients from the high-pass filter and approximation coefficients from the low-pass filter. The decomposition was defined as an orthogonal multi-resolution representation called wavelet representation. It was computed by a pyramidal algorithm which was based on convolutions with quadrate mirror filters. However, since half the frequencies of the signal have now been removed, half the samples can be discarded according to Nyquist's rule. The filter outputs are then subsampled by 2 which are represented by equation (2) and equation (3).

$$y_{low} = (x * g) \downarrow 2 \qquad (2)$$

$$y_{high} = (x * h) \downarrow 2 \qquad (3)$$

This decomposition has halved the time resolution since only half of each filter output characterizes the signal. However, each output has half the frequency band of the input so the frequency resolution has been doubled. Further, this decomposition is repeated to increase the frequency resolution and the approximation coefficients are decomposed with high and low pass filters and are then down-sampled. This is represented as binary tree with nodes representing a sub-space with different time-frequency localization. Tree is known as filter bank. Figure 5 shows 2-level filter bank where, x_n is the input image, A_n is the approximation image obtained by low-pass filtering at scale n and D_{ni} is the detail images obtained by high-pass filtering in a specific direction at scale n, for i=1, 2, 3 for vertical, horizontal and diagonal directions. Equation (4) is used for n level decomposition, where * denotes the convolution operator and ↓ 2, 1 (↓ 1, 2) is for subsampling along the rows or columns.

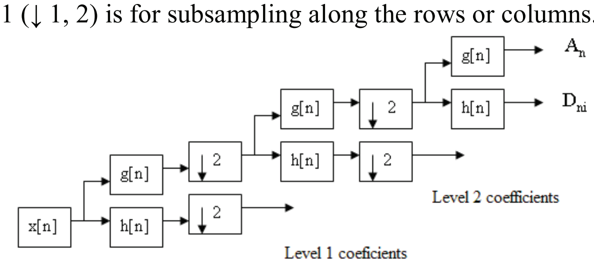

Figure 5. 2-level filter bank.

$$\left.\begin{aligned} A_n &= \left[h_x * \left[h_y * A_{n-1}\right]\downarrow 2,1\right]\downarrow 1,2 \\ D_{n1} &= \left[h_x * \left[g_y * A_{n-1}\right]\downarrow 2,1\right]\downarrow 1,2 \\ D_{n2} &= \left[g_x * \left[h_y * A_{n-1}\right]\downarrow 2,1\right]\downarrow 1,2 \\ D_{n3} &= \left[g_x * \left[g_y * A_{n-1}\right]\downarrow 2,1\right]\downarrow 1,2 \end{aligned}\right\} \qquad (4)$$

In this work, two-level decomposition is performed on the extracted ROIs by using 'db2' wavelet family. Approximation, horizontal, vertical and diagonal details are obtained and these form the feature vectors. The obtained feature vectors are used for classification.

D. Independent Component Analysis (ICA)

Independent component analysis (ICA) [11] is one of the dimensionality reduction methods. Independent component analysis was originally developed to deal with problems related to cocktail-party problem. The goal of independent component analysis (ICA) is to find a linear representation of nongaussian data such that the components are statically independent. Due to such representation of independent component analysis (ICA), its application also includes feature extraction and signal separation. ICA is much more powerful technique than PCA and factor analysis because it finds the underlying factors and sources whenever these classic methods fail to do. A statistical 'latent variables' model can be used to define ICA. As in [4] assume the sum of 'n' linear mixtures of a time signal $x_{1,}\ldots,x_n$ of 'n' independent components which results in x_j. It is given by equation (5),

$$x_j = a_{j1}s_1 + a_{j2}s_2 + a_{j3}s_3 + \ldots + a_{jn}s_n, \qquad (5) \text{ for all } j$$

where x_i and each independent component s_n are the random variables. On dropping the time signal the equation can be denoted in terms of vector-matrix notation. Consider 'X' as a random vector whose elements are mixtures of $x_1, \ldots, x_n$ and 'S' as random vector with elements $s_1, \ldots, s_n$ and matrix with a_{ij} is denoted as 'A'. 'X' is a column matrix, but X^T or the transpose of 'X' (row matrix) is considered. Using this vector-matrix notation, the model is written as equation (6),

$$X = AS \qquad (6)$$

sometimes columns of matrix A are also needed, by denoting them as $a_{i,}$ the model can also be written as equation (7),

$$X = \sum_{i=1}^{n} a_i s_i \qquad (7)$$

The statistical model in equation (7) is called as independent component analysis or ICA model. ICA model is a generative model, describes how the observed data are generated by the process of mixing the components of s_i. The independent components are latent variables which cannot be directly observed and mixing matrix is assumed to be unknown. Only by using the random vector X, both A and S are estimated. After estimating the matrix A, its inverse can also be computed denoted by W, and the independent component are obtained by equation (8),

$$S = WX \qquad (8)$$

ICA Algorithm:

1. Read images from training (x) and testing (y) dataset having equal size (N, N).
2. Convert training(x) and testing (y) column vector as a row vector of size (N^2*1).
3. Concatenate x and y to form a matrix S.
   S = [x; y] of size (2, (N*N))
4. Pre-process by multiplying **S** with generated random matrix A of size 2*2.
   X= A*S
5. Perform transformation by calculating covariance,
   $$Cov(Sx, Sy) = \sum_{i=1}^{n} \frac{(Sx_i - \overline{Sx_i})(Sy_i - \overline{Sy_i})}{n-1}$$
6. Calculate Eigen values e_i and Eigen vectors λ_i and project the values, $Ce_i = \lambda_i e_i$
   Where $C = Cov(Sx, Sy)$
   $$Y = E * D^{-0.5} * E' * X$$
7. Finally, estimate the independent components.

E. k-Nearest Neighborhood (k-NN)Classifier k-Nearest Neighborhood (k-NN) classification method is used in this work. k-NN was proposed by Cover and Hart [3], is a non-parametric method for classifying objects based on closest training samples in the feature space. In k-nearest neighborhood, an object is classified by majority votes of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. If k=1, then the object is simply assigned to the class of its nearest neighbor. k-NN is also known as “Memory based classification”, “Lazy Learning technique”, “Example-Based classification” or “Case-Based classification”. A lazy learner like k-NN [5] uses the training set directly to classify an input when an input is given. When using a k-nearest neighbor algorithm on an input with ‘n’ attributes the input is classified by taking a majority vote of the k, where k is some user specified constant, “closest” training records across all ‘n’ attributes. Here “closest” means the distance an attribute is away from the same attribute of the training set, evaluated using some specified similarity metric. k-NN is widely used in the field of pattern recognition, object recognition, text recognition, ranking models and event recognition applications.

Consider each sample x in a data set having ‘n’ attributes which combine to form an n-dimensional vector as, x = $(x_1, x_2, \ldots, x_n)$. These ‘n’ attributes are considered to be the independent variables. Each sample also has another attribute, denoted by y the dependent variable, whose value depends on the other ‘n’ attributes of x. Assume that y is a categorical variable, and there is a scalar function f which assigns a class, y = f(x) to every such vectors. Then, set a class of T, such that vectors are given together with their corresponding classes as: x(i), y(i) for i = 1,2, . . . , T.

Algorithm:

1. For each test point, x to be classified, find the k nearest samples in the training data.
2. Set a class of T, such that x(i),…, x(T) , where i=1, …, T.
3. Classify the point x according to the majority vote of their class labels.

III. Experimental Results

The proposed system is evaluated by conducting the experiments on Faces 94 database. Sample images are shown in figure 6. “Faces94” database contains frontal images of female, male and male-staff classes. There are 20 classes of female dataset, 113 classes of male dataset and 20 classes of male-staff dataset. Each class has 20 images, out of which some images are occluded by mustache and some by spectacles. For experimenting the proposed system, 40 classes from “Faces94” database are chosen. For training and testing, 40 class each and each class containing 10 images are considered. Out of 10 images 5 images from each class are used for training and remaining are used for testing. Out of 40 classes, 13 classes are taken from female dataset, 21 classes are from male dataset and 6 classes are from male-staff dataset. Therefore, total 400 images are considered, out of which 200 images are used for training and remaining 200 images are used for testing.

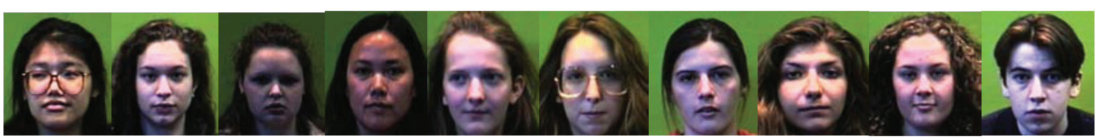

Figure 6 Sample images of Faces94 database.

A. Face detection

AdaBoost face detector is applied on the original image shown in figure 6.1(a). The detected face region is bounded by the rectangle as shown in figure 6.1(b). This detected face region is extracted (figure 6.1(c)) and is used for region of interest (ROI) extraction.

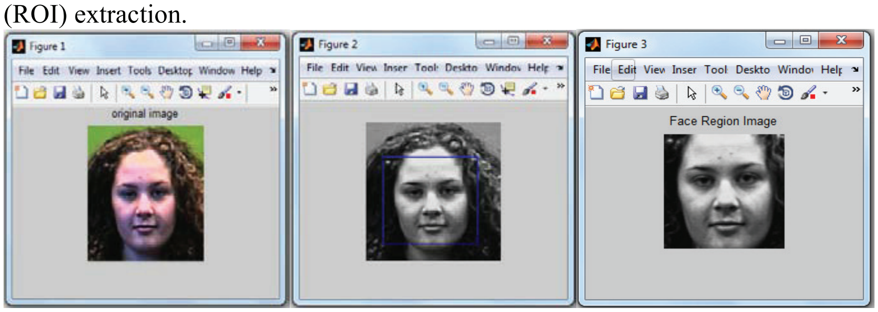

Figure 6.1(a) Original image. Figure 6.1(b) After applying AdaBoost face detector. Figure 6.1(c) Resultant extracted face.

B. Region of interest Extraction

To extract region of interest, the detected face is divided into 6 parts in horizontal and vertical directions, as shown in figure 6.2(a). In the top region, left eye and right eye are extracted which is shown in figure 6.2(b) and figure 6.2(c). In Bottom region, only nose and mouth region is extracted which is shown in figure 6.2(d).

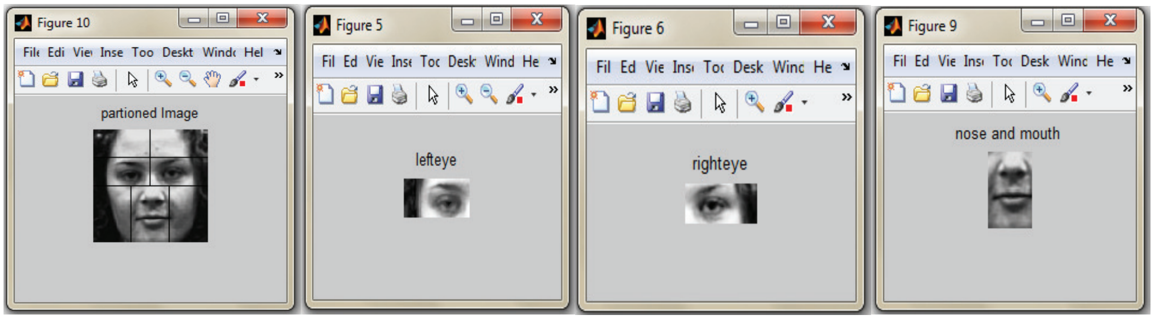

Figure6.2 (a) ROIs Extraction Figure6.2 (b) Extracted Left eye Figure6.2(c) Extracted right eye. Figure6.2(d) Extracted nose and mouth region.

C. Feature extraction

On performing 2-level decomposition using discrete wavelet transform on each of the extracted ROIs, approximation, horizontal, vertical and diagonal details are obtained which are shown in figure 6.3. The feature vectors of left eye, right eye and nose and mouth are obtained.

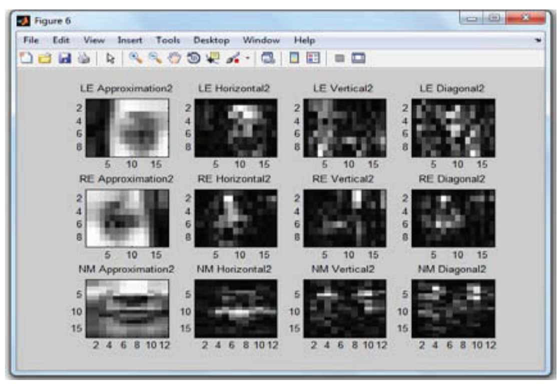

Figure 6.3 Approximation, horizontal, vertical and diagonal details for extracted ROIs.

D. ICA Implementation

Independent component analysis (ICA) is used for reducing dimensions of feature vectors. On applying ICA for training and testing feature vectors with size 2040*200, the size got reduced and is then projected by multiplying it with the original feature matrix.

E. k-NN Classification

Classification is done by classifier which consists of patterns of desired images. The pre-calculated values of training set are compared with the test values to classify the images by using k-nearest neighborhood classifier. Recognition Rate (RR) is used as a means of measuring the accuracy of recognition. The recognition rate is evaluated by using equation (9),

$$\text{Re}\,cognition\ rate = \frac{Number\ of\ \ face\ images\ correctly\ classified}{Total\ number\ of\ \ face\ images} \qquad (9)$$

On evaluating, the recognition rate about 83.5% is obtained when k=1, recognition rate for different k values is shown in table 1 and the corresponding plot is shown in figure 6.4.

Table 1 Recognition Rate for different values of k

kvalue	Recognition Rate in %
1	83.5
3	79.0
5	77.0
7	76.0
9	75.5
11	74.0
13	74.0
15	74.0

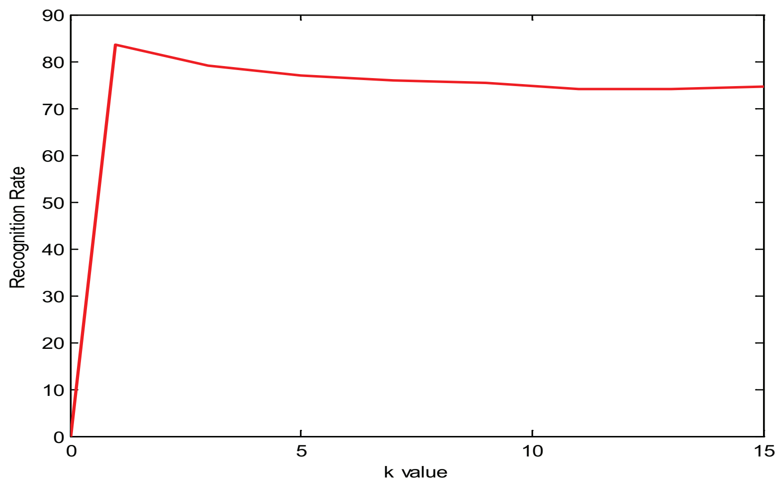

Figure 6.4 Plot of Recognition rate against different k value

IV.CONCLUSION

Face Recognition is one of the fastest growing biometric technologies to be used in real-time applications as it requires lesser user co-operation when compared to other biometrics like fingerprint and iris recognition. In this work, a face recognition system is implemented by extracting ROIs, using discrete wavelet transform for feature extraction and k-Nearest Neighborhood (k-NN) classifier for efficient classification. A Dimensionality reduction technique such as independent component analysis (ICA) is being used. Experiments were conducted on "Faces94" database containing female, male and male-staff classes by choosing 40 classes and 10 images for each class. Out of 10, 5 were used for training and remaining 5 were used for testing. The proposed system achieved about 83.5% of recognition rate when k value is 1.Recognition rate decreases as k value increases which is shown in figure 6.

REFERENCES

[1] Ajit Danti, K.M. Poornima, Narasimhamurthy, "Detection of Multiple Faces in Color Images using Haar Wavelets", International Conference on VLSI, Communication & Instrumentation (ICVCI), Proceedings published by International Journal of Computer Applications (IJCA), 6-11,2011.
[2] Ara V. Nefian and Monson H. Hayes, "Face Detection And Recognition Using Hidden Markov Models", IEEE Trans. ICIP, Proceedings 141 - 145 vol.1, 1998.
[3] T.M. Cover and P.E. Hart, "Nearest neighbor pattern classification", IEEE transaction in information theory, Vol 13, pp. 21-27, Jan 1967.
[4] Hyvarinen and Erkki Oja, "Independent Component Analysis: Algorithm and Applications", Neural Network computation, 13(4-5): 411-430, 2000.
[5] Nitin Bhatia and Vandana, "Survey of Nearest Neighbor Techniques", International Journal of Computer science and Information Security, Vol.8, No.2, 2010.
[6] A.Pentland and Matthew Turk, "Eigenaces for Recognition", journal of cognitive neuroscience, volume 3, number 1, page no71-86, 1991.
[7] K.T.Talele, Sunil Kadam and Atul Tikare. Article: Efficient Face Detection using Adaboost. *IJCA Proceedingson International Conference in Computational Intelligence (ICCIA2012)* iccia(10), March 2012.
[8] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey", ACM Computing Surveys, Vol. 35, No. 4, December 2003.
[9] https://en.wikipedia.org/wiki/Discrete_wavelet_transform
[10] https://en.wikipedia.org/wiki/Independent_component_analysis