Abstract—Machine simulation of human vision has been a subject of intensive research for scientists and engineers due to the numerous challenges associated with it. This paper presents a potential application of a novel biologically-inspired and wavelet-based model for face recognition using Support Vector Machine as classification algorithm. The biological knowledge about the distribution of light receptors, cones and rods, over the surface of the retina, and the way they are associated with the nerve ends for pattern vision forms the basis for the design of this model. A combination of classical wavelet decomposition and wavelet packet decomposition is used for simulating the functional model of cones and rods in pattern vision. The paper also describes results of experiments performed for face recognition using the features extracted on the AT & T face database (formerly, ORL face database) containing 400 face images of 40 different individuals. A feature vector of size 40 is formed for face images of each person and recognition accuracy is computed using the SVM classifier. Overall recognition accuracy obtained for the AT & T face database found to be very promising.

Keywords—Artificial Light Receptor; Feature Extraction; Face Recognition; Pattern Recognition; Support Vector Machine; Wavelets.

Abbreviations—Classical Wavelet Decomposition (CWD); Decision Directed Acyclic Graph (DDAG); Support Vector Machines (SVM); Vapnik-Chervonenkis (VC); Wavelet-based Artificial Light Receptor Model (WALRM); Wavelet based Artificial Light Receptor Feature Parameter (WALRFP); Wavelet Packet Decomposition (WPD).

I. INTRODUCTION

MACHINE simulation of human vision has been a subject of intensive research for scientists and engineers for the last three decades. However automatic face recognition is yet to achieve a completely reliable performance. There are several challenges involved in automatic face recognition -large variation in facial appearance, head size, orientation, changes in illumination and poses, occlusion, presence or absence of structural components etc are some of them to list. The interest devoted to this work is not only by the exciting challenges associated, but also the huge benefits that a Face-recognition system, designed in the context of a commercial application, could bring. Moreover, wide availability of powerful and low-cost desktop and embedded computing systems has also contributed to enormous interest in automatic processing of digital images and videos in a number of applications - Entertainment, Smart cards Information security, Low enforcement and Surveillance are some of them [Grudin, 2000; Daugman, 2001; Yang et al., 2002; Zhao et al., 2003].

Face recognition lies at the core of the discipline of pattern recognition where the objective is to recognize an image of face from a set of face images. A complete face recognition system generally consists of three stages. The first stage involves detecting and localizing the face in arbitrary images [Haddadnia et al., 2000; Wang & Tan, 2000; Li & Jain, 2004; Haddadnia & Ahmadi, 2004].

The second stage requires extraction of pertinent feature from the localized image obtained in the first stage. Finally, the third stage involves classification of facial images based on derived feature vector obtained in the previous stage. In order to design high accuracy recognition system, the choice of feature extraction method is very crucial. Two main approaches to feature extraction have been extensively used in conventional techniques [Daugman, 2001; Li & Jain, 2004]. The first one is based on extracting structural facial features that are local structures of face images, for example, the shapes of the eyes, nose and mouth. The structure based
approaches deals with local information rather than global information, and, therefore is not affected by irrelevant information in an image. However, because of the explicit model of facial features, the structure-based approaches are sensitive to unpredictability of face appearance and environmental conditions. The second method is statistical-based approach that extracts features from the entire image and, therefore uses global information rather than local information.

There have been a lot of popular attempts towards automated face recognition which kept the research in the area active and vibrant. Some of them are Eigenfaces (PCA based approach) [Turk & Pentland, 1991; Moon & Phillips, 2001], Independent Component Analysis (ICA) [Bartlett et al., 2002], Linear Discriminant Analysis (LDA) [Cootes & Taylor, 1997], a specific kind of genetic algorithm called Evolutionary Pursuit (EP) [Liu & Wechsler, 2000], Elastic Bunch Graph Matching (EBGM) where faces are represented as graphs, with nodes positioned at fiducial points [Wiskott et al., 1999] Kernel Methods which are a generalization of linear methods [Yang, 2002] like KPCA, KLDA, KICA etc., Trace transform, a generalization of the Radon transform [Kadyrov & Petrou, 2001], Active Appearance Model (AAM) is an integrated statistical model which combines a model of shape variation with a model of the appearance variations in a shape-normalized frame [Cootes et al., 2000], Hidden Markov Models (HMM) [Nefian & Hayes, 2000], and Support Vector Machine (SVM) [Guo et al., 2000]. The authors have already made a fruitful attempt for modeling Light Receptors, cones and rods, using wavelets in the reference [Kabeer & Narayan, 2009]. And classification results using Artificial Neural Network (ANN) has also been found promising. The present paper is an attempt to use the method with SVM classifier which resulted in better recognition accuracy.

The paper is organized as follows. In Section 2 Wavelet-based Artificial Light Receptor Model (WALRM) for feature extraction method is described. In Sec. 3, Support Vector Machine for face recognition is discussed. Section 4 presents the simulation experiment conducted using AT & T face database and reports the recognition results obtained using SVM classifier. Finally, Sec. 5 gives the conclusions and direction for future research.

II. WAVELET BASED ARTIFICIAL LIGHT RECEPTOR MODEL

Pattern vision is afforded by the distribution of light receptors over the surface of the retina. There are two classes of receptors, called cones and rods. The cones in each eye number between six and seven million, and are located primarily in the central portion of retina. These cones help humans to resolve fine details they see around, largely because each one is connected to its own nerve end. On the other hand, rods are very huge in number when compared to cones. Several rods are connected to a single nerve end, which in turn reduces the amount of detail carried by these receptors. Figure 1 shows this arrangement of rods and cones in retina and biological signal passing structure from retina to the brain. This association of rods and cones with the nerve ends forms the basis for the design of the model in this study.

The model is simulated using a combination of Classical Wavelet Decomposition (CWD) and Wavelet Packet Decomposition (WPD). Each face image is described by a subset of band filtered images containing wavelet coefficients. The elements from these coefficients matrices are subjected to simple statistical operations and the results are organized in such a fashion similar to the arrangements of rods and cones in retina giving compact and meaningful feature vectors. Figure 2 shows the block diagram for the entire recognition system using Wavelet based Artificial Light Receptor Feature Extraction Model.

Figure 1: Rods and Cones in Retina and Biological Image Signal Passing Structure

Figure 2: Face Recognition System using Wavelet based Artificial Light Receptor Model for Feature Extraction
2.1. Feature Extraction Process

The feature extraction process consists of three stages. In the first stage one component of Wavelet based Artificial Light Receptor Feature Parameter (WALRFP) is created by subjecting the face image to undergo CWD recursively to decompose it into fifth level of resolution (fifth level has been found to be optimum experimentally as illustrated by Table 1). Therefore, the approximation matrix at this level of resolution is significantly small representative of the original image and carries enough information content to describe face image characteristics coarsely. This matrix can be considered as analogous to an image formed in retina at cone area. We call this functional unit in our model as Wavelet face image characteristics coarsely. This matrix can be considered as analogous to an image formed in retina at cone area. We call this functional unit in our model as Wavelet Cones. And, each element in the matrix is sent to separate units (nerve ends) as the case may be with human visual system.

Let \( A_k \) represents this approximation matrix at decomposition level \( k \), which can be written as:

\[
A_k = \begin{bmatrix}
A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\
A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{m,1} & A_{m,2} & \cdots & A_{m,n}
\end{bmatrix}
\]

Then, the first component of Wavelet based Artificial Light Receptor model driven Feature vector, \( V_{kc} \), is given by,

\[
V_{kc} = \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} [A_{ij}]
\]

In the second stage, the Wavelet rods are used to extract the other component of Wavelet based Artificial Light Receptor feature vector by decomposing each face image using WPD to its fifth level (for the same reason stated before) of resolution. Then we find the best level of wavelet packet decomposition tree. The first coefficient matrix at the best level tree contain enough information to represent the given input face image without loss of much facial features. Let \( \mu \) represent mean of one row in this coefficient matrix then the second component of the WALRFP vector, \( V_{hr} \), is given by, \( V_{hr} = \{\mu_i\} \), \( \forall i, i = 1, 2, 3 \ldots m \) (number of rows in the best level coefficient matrix).

In the third stage we combined \( V_{kc} \) and \( V_{hr} \) to form the final wavelet based Artificial Light Receptor Model feature vector \( V_{WALRFP} \).

\[
V_{WALRFP} = \bigcup [V_{kc}, V_{hr}]
\]

As the wavelet cone feature component is of size 12 and wavelet rod component is 28 the estimated \( V_{WALRFP} \) dimension is constraint to forty.

Feature vectors of representative samples had been generated from the database at different decomposition levels, and these feature vectors were subjected to classification using \( k-NN \) classifier - comparatively faster classification algorithm with lesser accuracy. Table 1 gives the classification results on this representative subset using \( k-NN \) classifier.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Resolution Level</th>
<th>Feature Size</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1</td>
<td>2604</td>
<td>25</td>
</tr>
<tr>
<td>2.</td>
<td>2</td>
<td>672</td>
<td>33</td>
</tr>
<tr>
<td>3.</td>
<td>3</td>
<td>196</td>
<td>56</td>
</tr>
<tr>
<td>4.</td>
<td>4</td>
<td>70</td>
<td>70.5</td>
</tr>
<tr>
<td>5.</td>
<td>5</td>
<td>40</td>
<td>81.5</td>
</tr>
<tr>
<td>6.</td>
<td>6</td>
<td>32</td>
<td>46</td>
</tr>
<tr>
<td>7.</td>
<td>7</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>8.</td>
<td>8</td>
<td>29</td>
<td>30</td>
</tr>
</tbody>
</table>

Analysis of table 1 shows that feature vector generated at resolution level 5 is better than feature vectors at other resolution levels. This analysis lead us to decide the features at resolution level 5 is optimal for recognition.

Figure 3: Feature Vector Generated for Ten Face Images of the First Person in AT & T Face Database using the Artificial Light Receptor Model

Figure 3 shows feature vector graph obtained from face database plotted for ten samples of first person in AT&T face database along with the mean curve. The graphs obtained for different samples of same person are found to be quite similar while the graphs for different persons are highly distinguishable.

III. SUPPORT VECTOR MACHINE (SVM) FOR HUMAN FACE IMAGE CLASSIFICATION

Basically, SVM is a linear machine with some nice properties. In the context of pattern classification, the main idea of a support vector machine is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. Basically, the support vector machine has the inherent ability to solve a pattern-classification problem in a manner close to the optimum for the problem of interest. Moreover, it is able to achieve such a remarkable performance with no problem domain knowledge built into the design of the machine. The machine achieves this desirable property by following a principled approach rooted in statistical learning theory, structural risk minimization. This induction principle is based on the fact that the error rate of a learning machine on test data (i.e., generalization error report) is bounded by the sum of the training-error rate and...
the term that depends on the Vapnik-Chervonenkis (VC) dimension [Vapnik, 1982; 1998].

In the case of separable patterns, the support vector machine produces a value of zero for the first term and minimizes the second term. Accordingly, the support vector machine can provide a good generalization performance on pattern classification problems despite the fact that it does not incorporate problem-domain knowledge [Burges, 1998]. This attribute is unique to support vector machines.

In the present study, an effort is made to build a recognition system using the support vector machine for human face recognition. A notion that is central to the construction of the support vector learning algorithm is the inner-product kernel between a “support vector” \(x_i\) and the vector \(x\) drawn from the input space. The support vectors consist of small subset of the training data extracted by the algorithm. We used WLRM based features discussed in the previous section for human face recognition. The recognition results show that this method is more efficient and can be adopted for developing a complete FR system for human face images.

A relatively new learning architecture is employed in the paper the Decision Directed Acyclic Graph (DDAG) [John C. Platt et al., 2000], which combines many two-class classifiers into a multiclass classifier. For an \(N\)-class problem, the DDAG contains classifiers, one for each pair of classes. DAGSVM operates in a kernel- induced feature space and uses two-class maximal margin hyperplanes at each decision-node of the DDAG. The DAGSVM is substantially faster to train and evaluate than either the standard algorithm or Max Wins, while maintaining comparable accuracy to both of these algorithms.

### 3.1. Multi-Class Classification using SVM

The problem of multi-class classification, especially for systems like SVMs, doesn’t present an easy solution. It is generally simpler to construct classifier theory and algorithms for two mutually-exclusive classes than for \(N\) mutually-exclusive classes. Literatures reveal that constructing \(N\)-class SVMs is still an unsolved research problem. The standard method for \(N\)-class SVMs [Vapnik, 1998] is to construct \(N\)-SVMs. The \(i\)th SVM will be trained with all of the examples in the \(i\)th class with positive labels, and all other examples with negative labels. SVMs trained in this way are referred as 1-v-r SVMs (one-versus-rest). The final output of the \(N\) 1-v-r SVMs is the class that corresponds to the SVM with the highest output value. Unfortunately, there is no bound on the generalization error for the 1-v-r SVM, and the training time of the standard method scales linearly with \(N\). Another method for constructing \(N\)-class classifiers from SVMs is to combine two-class classifiers.

We have used a multiclass learning architecture, called the Decision Directed Acyclic Graph (DDAG) proposed by John C. Platt et al., (2000). The DDAG contains \(N(N-1)/2\) nodes, each with an associated 1-v-1 classifier.

A Directed Acyclic Graph (DAG) is a graph whose edges have an orientation and no cycles. A Rooted DAG has a unique node such that it is the only node which has no arcs pointing into it. A Rooted Binary DAG has nodes which have either 0 or 2 arcs leaving them. Here, rooted Binary DAGs are employed in order to define a class of functions to be used in classification tasks. The class of functions computed by Rooted Binary DAGs is formally defined as follows.

Given a space \(X\) and a set of boolean functions \(F = \{f: \{0,1\}\}\), the class DDAG(\(F\)) of Decision DAGs on \(N\) classes over \(F\) are functions which can be implemented using a rooted binary DAG with \(N\) leaves labelled by the classes where each of the \(N(N-1)/2\) internal nodes is labelled with an element of \(F\). The nodes are arranged in a triangle with the single root node at the top, two nodes in the second layer and so on until the final layer of \(N\) leaves. The \(i\)-th node in layer \(j < N\) is connected to the \(i\)-th and \((i+1)\)-th node in the \((j+1)\)-st layer.

To evaluate a particular DDAG \(G\) on input \(x \epsilon X\), starting at the root node, the binary function at a node is evaluated. The node is then exited via the left edge, if the binary function is zero; or the right edge, if the binary function is one. The next node’s binary function is then evaluated. The value of the decision function \(D(x)\) is the value associated with the final leaf node (see Figure 4(a)). The path taken through the DDAG is known as the evaluation path. The input \(x\) reaches a node of the graph, if that node is on the evaluation path for \(x\). It is referred that the decision node distinguishing classes \(i\) and \(j\) as the \(ij\)-node. Assuming that the number of a leaf is its class, this node is the \(i\)-th node in the \((N-j+1)\)-th layer provided \(i < j\). Similarly the \(j\)-nodes are those nodes involving class \(j\), that is, the internal nodes on the two diagonals containing the leaf labelled by \(j\).

![Figure 4](image-url)

Figure 4: (a) The decision DAG for finding the best class out of four classes. The equivalent list state for each node is shown next to that node. (b) A diagram of the input space of a four-class problem. A 1-v-1 SVM can only exclude one class from consideration.

The DDAG is equivalent to operating on a list, where each node eliminates one class from the list. The list is initialized with a list of all classes. A test point is evaluated against the decision node that corresponds to the first and last elements of the list. If the node prefers one of the two classes, the other class is eliminated from the list, and the DDAG proceeds to test the first and last elements of the new list. The DDAG terminates when only one class remains in the list. Thus, for a problem with \(N\) classes, \((N-1)\) decision nodes will be evaluated in order to derive an answer. Figure 4(a)
shows the decision DAG for finding the best class out of four classes. The equivalent list state for each node is shown next to that node. Figure 4(b) shows a diagram of the input space of a four-class problem. A 1-v-1 SVM can only exclude one class from consideration.

The current state of the list is the total state of the system. Therefore, since a list state is reachable in more than one possible path through the system, the decision graph the algorithm traverses is a DAG, not simply a tree.

Decision DAGs naturally generalize the class of Decision Trees, allowing for a more efficient representation of redundancies and repetitions that can occur in different branches of the tree, by allowing the merging of different decision paths. The class of functions implemented is the same as that of Generalized Decision Trees [Bennett et al., 2000], but this particular representation presents both computational and learning-theoretical advantages.

IV. EXPERIMENTS AND RESULTS

All the experiments were carried out using the AT & T face database, which contains face images of 40 distinct persons. Each person has ten different images, taken at different times. Figure 5 shows five individuals (in five rows) in the AT & T face images.

![Figure 5: Sample Faces Images taken from the AT & T Face Database](image)

There are variations in facial expressions such as open/closed eyes, smiling/non-smiling, and facial details such as with glasses/without glasses. All the images were taken against a dark homogeneous background with the subjects in an up-right, frontal position, with tolerance for some side movements. There are also some variations in scale.

The face image database containing 400 images of 40 different persons were divided into training set and test set by randomly selecting images from the database. The SVM was trained using training set images separately. Then the test set was used to check the recognition accuracy of the method. The results were drawn against each person class in a cumulative manner, and the cumulative recognition accuracy was plotted. The plot in figure 6 shows this. The result shows that there is an overall recognition of 96.65% in AT & T face database for the method proposed. This result obtained was found to be promising when compared to the other methods investigated by the authors.

V. CONCLUSION

This paper presented a robust Wavelet-Based Artificial Light Receptor Model for extracting face image feature vectors. A feature vector of size 40 is formed for face images of each person and recognition accuracy is computed using SVM classifier. Overall recognition accuracy obtained for the AT&T face database is 96.65%. There is significant dimensionality reduction as we used a feature vector of a 40-element size to represent a face image. More effective implementation of multiple classifier system is one of our future research directions and more research is needed to deal with building a fully functional system.

REFERENCES


Dr. V. Kabeer, MCA, Ph.D., is the present head of the Postgraduate Department of Computer Science, Farook College, Kozhikode, which is a College identified with Potential for Excellence by the UGC and affiliated to University of Calicut. He earned his Ph.D. in Information Science and Technology from Kannur University, Kerala, India and his MCA Degree from TKM College of Engineering, Kollam, Kerala, India. Previously, he was Associate Professor at MES College of Engineering, Kuttippuram, Kerala. He has more than 17 publications in his credit which include International and National Journals and Conferences. Also, he has made more than 20 invited talks in the areas of his research in particular and Computer Science & Applications in general. He has served as the editor of the MES Journal of Technology and Management in the year 2010. He is a life member of ISTE.

Zainul Abid T P, B.Tech (Computer Engg.), M.E (Computer Science & Engg.) is a research scholar having more than eleven years of teaching and research experience. He has previously worked in M.E.S college of Engg, Kuttippuram, Kerala and Royal College of Engg, Thrisur, Kerala. He pursued his bachelor degree from Cochin University of Science & Technology and masters from University Visheswaraya College of Engg, Bangalore University.