Multiresolution Hybrid Approaches for Automated Face Recognition

Nicholl, P.; Bouchaffra, D.; Amira, A.; Perrott, R.H.;
Queen's Univ. Belfast, Belfast

This paper appears in: Adaptive Hardware and Systems, 2007. AHS 2007. Second NASA/ESA Conference on
Issue Date: 5-8 Aug. 2007
On page(s): 89 - 96
Location: Edinburgh
Print ISBN: 978-0-7695-2866-3
References Cited: 26
Cited by : 1
INSPEC Accession Number: 9864442
Digital Object Identifier: 10.1109/AHS.2007.77
Date of Current Version: 20 August 2007

ABSTRACT

This paper presents an evaluation of three classifiers using the discrete wavelet transform (DWT) as a feature extractor. The thrust is to investigate the impact of DWT with its various filter banks on the HMM, PCA and SHMM classifiers. In addition, we have developed a novel approach that combines the multiresolution feature of the discrete wavelet transform with the local interactions of the facial structures expressed through the structural hidden Markov model (SHMM). Tests have been carried out on the AT&T and Essex face databases, which show that DWT/SHMM outperforms both the DWT/HMM and DWT/PCA with an 8% increase in accuracy.

Multiresolution Hybrid Approaches for Automated Face Recognition

Paul Nicholl
School of Electronics, Electrical Engineering & Computer Science
Queen's University, Belfast, UK
Email: p.nicholl@qub.ac.uk

Abbes Amira
School of Engineering & Design
Brunel University
London, UK
Email: abbes.amira@brunel.ac.uk

Djamel Bouchaffra
School of Engineering & Computer Science
Oakland University, Rochester
MI 48309, USA
Email: dbouchaffra@ieee.org

***Abstract*—This paper presents an evaluation of three classifiers using the discrete wavelet transform (DWT) as a feature extractor. The thrust is to investigate the impact of DWT with its various filter banks on the HMM, PCA and SHMM classifiers. In addtion, we have developed a novel approach that combines the multiresolution feature of the discrete wavelet transform with the local interactions of the facial structures expressed through the structural hidden Markov model (SHMM). Tests have been carried out on the AT&T and Essex face databases, which show that DWT/SHMM outperforms both the DWT/HMM and DWT/PCA with an 8% increase in accuracy.**

I. Introduction

The demand for sophisticated security systems is growing quickly. Both governments and businesses recognize that the threat to individuals and property is significant. To try to address this threat, there has been an increasing focus on biometric person recognition systems.

The potential for face recognition systems in this regard is obvious. Unlike other biometrics, including iris and fingerprint recognition, face recognition can be performed with minimal compliance on behalf of the person being identified [1]. As it is the primary method used by humans to identify each other, and because an individual's face images is already stored in numerous locations, it is seen as a more acceptable method of automatic recognition [2].

A multitude of techniques have been applied to face recognition. Geometric feature matching involves extraction of significant facial features, such as eyes, nose, etc and the calculation of descriptive information about these features [3].

Many techniques fall into the *template matching* category, where faces are treated as two-dimensional arrays of intensity values, to be compared with other facial arrays. Techniques of this type include principal component analysis (PCA), where the variance among a set of face images is represented by a number of eigenfaces [4], [5]. In independent component analysis (ICA), faces are assumed to be linear mixtures of some unknown latent variables. The latent variables are assumed non-gaussian and mutually independent, and they are called the independent components of the observed data [6]. In neural network models (NNMs), the system is supplied with a set of training images along with correct classification, thus allowing the neural network to ascertain a weighting system to determine which areas of an image are deemed most important [7].

Hidden Markov models (HMMs) [8], which have been used successfully in speech recognition for a number of decades, are now being applied to face recognition. Samaria and Young used image pixel values to build a top-down model of a face using HMMs. Nefian and Hayes [9] modified the approach by using discrete cosine transform (DCT) coefficients to form observation vectors, whereas Bicego and Murino [10] used discrete wavelet transform (DWT) [11] coefficients of sub-windows generated by a raster-scan of the image.

As HMMs are one-dimensional in nature, a variety of approaches have been adopted to try to represent the two-dimensional structure of face images. These include the 1D discrete HMM (1D-DHMM) approach [12], which models a face image using two standard HMMs, one for observations in the vertical direction and one for the horizontal direction. Another approach is the pseudo-2D HMM (2D-PHMM) [13], which is a 1D HMM, composed of super states to model the sequence of columns in the image, in which each super state is a 1D-HMM, itself modelling the blocks within the columns.

Although HMMs are effective in modeling statistical information, they are not suited to model the structural information that is often inherent in a pattern. They are therefore not optimal for handling structural patterns such as the human face.

One recently developed model for pattern recognition is the structural hidden Markov model (SHMM) [14], [15]. This approach allows both the structure and the statistical properties of a pattern to be represented within the same model. SHMMs have been shown to outperform HMMs in a number of applications including handwriting recognition, but have yet to be applied to face recognition.

Hybrid multiresolution approaches have received much attention in recent years. As well as DWT/HMM, wavelet technology has been used along with PCA[16], [17], ICA[18] and support vector machines (SVMs)[19]. In face recognition, DWT provides the advantage of extracting data that is useful for recognition and discarding that which is not. It has been shown to outperform techniques based on DCT and discrete Fourier transform (DFT)[20].

The objective of the work presented in this paper is to

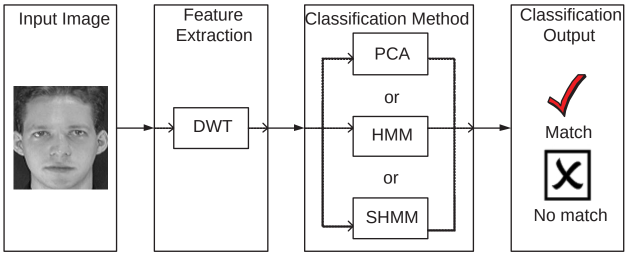

Fig. 1. Proposed System.

develop a framework for multiresolution hybrid face recognition. Techniques developed within the framework include DWT/PCA, DWT/HMM and DWT/SHMM. A comparison of these techniques will be presented and the effect of wavelet filter analyzed. In addition, a solution using computational grid technology to address the unique problems presented by face recognition using very large, distributed databases will be proposed.

The rest of this paper is organized as follows: Section 2 gives an overview of the proposed recognition framework. Section 3 explains the mathematical background. Section 4 presents the face recognition approach adopted. Section 5 contains results and analysis and section 6 contains concluding remarks.

II. Proposed System for Face Recognition

A. Local Aspect

A framework for face recognition has been developed. As can be seen in Figure 1, a face image is input to the system for recognition. DWT is used to extract features, which then undergo classification. For this stage, a variety of methods may be employed, including PCA, HMM and SHMM. The results of recognition, including the identities of the closest matches and confidence scores, are output to the user.

B. Distributed Aspect

As part of this ongoing research, the use of computational grid technology will be investigated as a means for developing a distributed face recognition system that is capable of *retaining robustness and reliability* when deployed in a large-scale environment. This approach to face recognition would allow a switch from the tendency for face recognition systems to be seen as isolated installations, to a more collaborative approach. For example, geographically distributed law enforcement agencies might have separate criminal 'mugshot' databases but a shared interest in identifying a suspect who migrates across territories. It would therefore seem logical to allow inter-organizational access to these face databases. Currently, however, issues such as security concerns incite reluctance to share access in this way. Hence, the use of grid technology as a facilitator to such sharing will be investigated. Issues such as speed, accuracy and security of coordinated sharing will be examined.

It is envisaged that a recognition client will interact with face databases on resources at a number of distributed locations. The system will make use of Globus Toolkit [21] middleware to interact with the remote databases.

When a user provides the client application with a face image for recognition, the client will extract a feature vector from the image. This feature vector will then be supplied to *matching agents* at each distributed database location. The matching agents will employ one of the recognition algorithms described in this paper to search for possible matches for the supplied face image. All possible matches will then be returned to the client where they will be displayed to the user.

III. Mathematical Background

A. Discrete Wavelet Transform

DWT is increasingly being used in image processing and feature extraction where previously other transforms would have been used. The advantage of DWT over DFT and DCT is that DWT performs a multiresolution analysis of a signal with localization in both time *and* frequency [11], [22]. In addition to this, functions with discontinuities and functions with sharp spikes require fewer wavelet basis vectors in the wavelet domain than sine-cosine basis vectors to achieve a comparable approximation.

DWT can be implemented as a set of filter banks, comprising a high-pass and low-pass filter. In standard wavelet decomposition, the output from the low-pass filter can then be decomposed further, with the process continuing recursively in this manner. DWT can be mathematically expressed by Equation 1:

$$\mathrm{DWT}_{\mathrm{x(n)}} = \begin{cases} \mathrm{d_{j,k}} = \sum x(n) h_j^*(n - 2^j k) \\ \mathrm{a_{j,k}} = \sum x(n) g_j^*(n - 2^j k) \end{cases} \tag{1}$$

The coefficients $d_{j,k}$ refer to the detail components in signal $x(n)$ and correspond to the wavelet function, whereas $a_{j,k}$ refer to the approximation components in the signal. The functions $h(n)$ and $g(n)$ in the equation represent the coefficients of the high-pass and low-pass filters respectively.

For the case of images, the one dimensional DWT can be readily extended to two dimensions and produces four sets of coefficients. The 'low-low' (LL) quadrant contains low frequency information in both directions. The 'high-high' (HH) quadrant contains high frequency in both directions. Both the 'high-low' (HL) and 'low-high' (LH) quadrants contain low frequency information along one direction and high frequency information along the other.

Figure 2 illustrates the effect of applying the wavelet transform to an image. Figure 2a shows an image taken from the AT&T Face Database [23]. Figure 2b shows the effect of applying a one-level wavelet decomposition to this image (using the Haar wavelet) and Figure 2c shows the complete decomposition.

B. Principal Component Analysis

PCA is a dimensional-reduction technique. It works by finding a new coordinate system for a set of data, where the

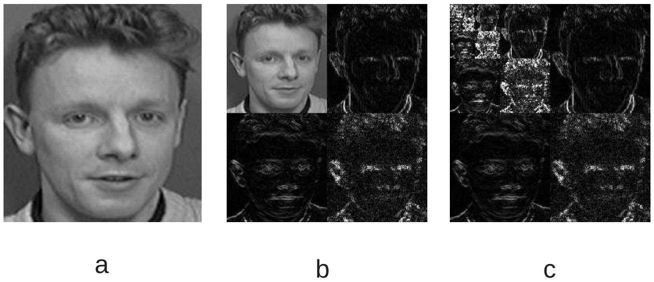

Fig. 2. Wavelet transform of image (a) Original Image (b) 1-level Haar decomposition (c) Complete Haar decomposition

axes (or *principal components*) are ordered by the variance contained within the training data. A set of faces can be represented as points in this new coordinate system.

To achieve this, a set of face images $\{x_i\}$ is represented as a matrix X, where:

$$\mathrm{X} = [x_1 x_2 x_3 ... x_M] \tag{2}$$

and X is of dimension N x M, with N being the number of pixels in an image. The 'average' face is calculated and subtracted from each face in X, giving X':

$$X' = [(x_1 - \overline{x})(x_2 - \overline{x})(x_3 - \overline{x})...(x_M - \overline{x})] \tag{3}$$

The principal components of this set are found by calculating the eigenvectors of the covariance matrix C, where:

$$C = \sum_{i=1}^{M} X' X'^T \tag{4}$$

The calculated eigenvectors are used as an orthogonal basis to represent the training set faces. The k eigenvectors with the highest associated eigenvalues represent the greatest variance within the face set and are retained for representation. Each training set face is then projected onto these eigenvectors, producing a k-dimensional feature vector. New images are also projected onto these eigenvectors and the resulting k-dimensional vectors are compared with those of known individuals.

C. Hidden Markov Models

HMMs are used to characterize the statistical properties of a signal [8]. They have been used in speech recognition applications for many years and are now being applied to face recognition. An HMM consists of a number of non-observable states and an observable sequence, generated by the individual hidden states. Figure 3 illustrates the structure of a simple HMM.

HMMs are defined by the following elements:

- N, the number of hidden states in the model
- M, the number of different observation symbols
- $S = \{S_1, S_2, ..., S_N\}$, the finite set of possible hidden states. The state of the model at time t is given by $q_t \in S, 1 \leq t \leq T$, where T is the length of the observation sequence.

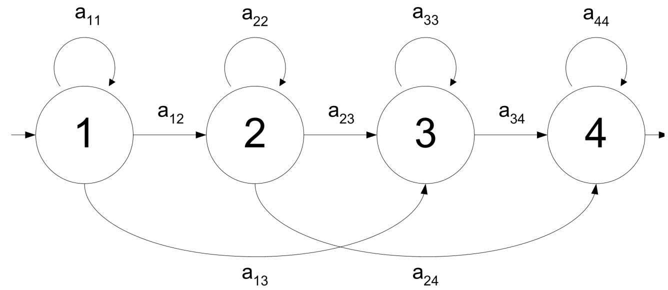

Fig. 3. A simple left-right HMM

- $A = \{a_{ij}\}$, the state transition probability matrix, where
$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i] \quad 1 \leq i, j \leq N \tag{5}$$
with
$$0 \leq a_{i,j} \leq 1$$
and
$$\sum_{j=1}^{N} a_{ij} = 1, 1 \leq i \leq N$$
- $B = \{b_j(k)\}$, the emission probability matrix, indicating the probability of a specified symbol being emitted given that the system is in a particular state, i.e.
$$b_j(k) = P[O_t = k | q_t = S_j] \tag{6}$$
with
$$1 \leq j \leq N$$
and O_t is the observation symbol at time t.
- $\Pi = \{\pi_i\}$, the initial state probability distribution, i.e.
$$\pi_i = P[q_1 = S_i] 1 \leq i \leq N \tag{7}$$
with
$$\pi_i \geq 0$$
and
$$\sum_{i=1}^{N} \pi_i = 1$$

An HMM can therefore be succinctly defined by the triplet

$$\lambda = (A, B, \Pi).$$

HMMs are typically used to address three unique problems [8]:

- **Evaluation:** Given a model λ and a sequence of observations O, what is the probability that O was generated by model λ, i.e. $P(O|\lambda)$.
This problem is typically used for pattern recognition tasks; the probability of an observation sequence being generated is evaluated against a number of distinct HMMs, each of which correspond to a class of pattern.

- **Decoding:** Given a model λ and a sequence of observations O, what is the hidden state sequence q^* most likely to have produced O, i.e. $q^* = argmax_q[P(q|\lambda, O)]$. This problem can be used to learn about the structure of a model, or to find the optimal state sequence in some application, such as speech recognition.
- **Parameter Estimation:** Given an observation sequence O, what model λ is most likely to have produced O. This problem is referred to as training, as the model's parameters are adjusted until some convergence criterion is reached. Typically, a number of observation sequences are used to train a model [8].

D. Structural Hidden Markov Models

One of the major problems of HMMs is that they have great difficulty in learning to capture the long range dependencies in a sequence that often constitute the structural information contained within. Therefore, in this section, the mathematical expression of SHMMs is introduced. The entire description of the SHMM can be found in [14], [15].

Let $O = (O_1, O_2, \cdots, O_S)$ be the time series sequence (the entire pattern) made of S subsequences (also called subpatterns). The entire pattern can be expressed as: $O = (o_{1_1} o_{1_2} \cdots o_{1_{r_1}}, \cdots, o_{s_1}, o_{s_2}, \cdots, o_{s_{r_s}})$, (where r_1 is the number of observations in subsequence O_1 and r_2 is the number of observations in subsequence O_2, etc., such that $\sum_{i=1}^{i=s} r_i = T$). A local structure C_j is assigned to each subsequence O_i. Therefore a sequence of local structures $C = (C_1, C_2, \ldots, C_S)$ is generated from the entire pattern O. The probability of a complex pattern O given a model λ can be written as:

$$P(O \mid \lambda) = \sum_C P(O, C \mid \lambda). \tag{8}$$

Therefore, there is a need to evaluate $P(O, C \mid \lambda)$. The model λ is implicitly present during the evaluation of this joint probability, so it is omitted. Therefore:

$$P(O, C) = P(C, O) = P(C \mid O) \times P(O) \tag{9a}$$
$$= P(C_S \mid C_{S-1} \ldots C_2 C_1 O_S \ldots O_1) \tag{9b}$$
$$\times P(C_{S-1} \ldots C_2 C_1 \mid O_S \ldots O_1) \times P(O). \tag{9c}$$

It is assumed that C_i depends only on O_i and C_{i-1}, and the structure probability distribution is a Markov chain of order 1. Figure 4 illustrates an SHMM of order 1. It has been proven in [14] that the likelihood function of the observation sequence can be expressed as:

$$P(O \mid \lambda) \approx \sum_C \left[\prod_{i=1}^{S} \frac{P(C_i \mid O_i) P(C_i \mid C_{i-1})}{P(C_i)} \times P(O) \right] \tag{10}$$

The organization (or syntax) of the symbols o_i is introduced mainly through the term $P(C_i \mid O_i)$ since the transition probability $P(C_i \mid C_{i-1})$ does not involve the inter-relationship of the symbols o_i. Besides, the term $P(O)$ of Equation 10 is viewed as a traditional HMM.

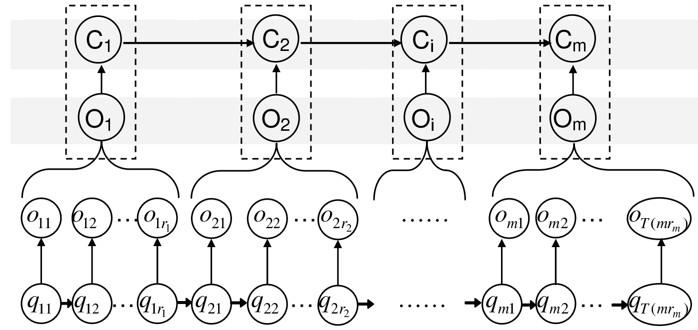

Fig. 4. A graphical representation of a first-order structural hidden Markov model.

Therefore, an SHMM can be defined as follows:

Definition 3.1: A structural hidden Markov model is a quintuple $\lambda = [\pi, \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}]$, where:

- π is the initial state probability vector,
- $\mathcal{A}$ is the state transition probability matrix,
- $\mathcal{B}$ is the state conditional probability matrix of the visible observations,
- $\mathcal{C}$ is the posterior probability matrix of a structure given a sequence of observations,
- $\mathcal{D}$ is the structure transition probability matrix.

An SHMM is characterized by the following elements:

- **N**, the number of hidden states in the model. The individual states are labeled as 1, 2, …, N, and denote the state at time t as q_t.
- **M**, the number of distinct observations o_i
- π, the initial state distribution, where $\pi_i = P(q_1 = i)$ and $1 \leq i \leq N$, $\sum_i \pi_i = 1$.
- $\mathcal{A}$, the state transition probability distribution matrix: $\mathcal{A} = \{a_{ij}\}$, where: $a_{ij} = P(q_{t+1} = j \mid q_t = i)$ and $1 \leq i, j \leq N$, $\sum_j a_{ij} = 1$.
- $\mathcal{B}$, the state conditional probability matrix of the observations, $\mathcal{B} = \{b_j(k)\}$, in which: $b_j(k) = P(o_k \mid q_j)$, $1 \leq k \leq M$ and $1 \leq j \leq N$, $\sum_k b_j(k) = 1$. In the continuous case, this probability is a density function expressed as a finite weighted sum of Gaussian distributions (mixtures).
- **F**, the number of distinct local structures.
- $\mathcal{C}$ is the posterior probability matrix of a structure given its corresponding observation sequence: $\mathcal{C} = c_i(j)$, where: $c_i(j) = P(C_j \mid O_i)$. For each particular input string O_i, we have: $\sum_j c_i(j) = 1$.
- $\mathcal{D}$, the structure transition probability matrix: $\mathcal{D} = \{d_{ij}\}$, where $d_{ij} = P(C_{t+1} = j \mid C_t = i)$, $\sum_j d_{ij} = 1$, $1 \leq i, j \leq F$.

Figure 4 depicts a graphical representation of an SHMM.

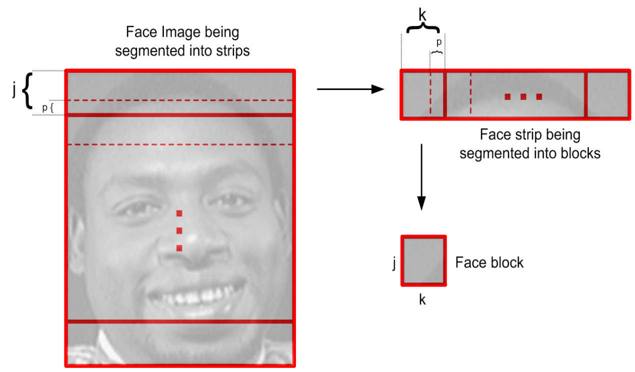

Fig. 5. An illustration showing the creation of the block sequence.

E. Problems assigned to an SHMM

There are four problems that are assigned to a SHMM:

- **Evaluation:** Given a model λ and an observation sequence $O = (O_1, \cdots, O_s)$, the goal is to evaluate how well does the model λ match O.
- **Statistical decoding:** In this problem, an attempt is made to find the best state sequence. This problem is similar to problem 2 of the traditional HMM and can be solved using Viterbi algorithm as well.
- **Local structure decoding:** This is the most important problem, involving an attempt to determine the "optimal local structures of the model". An example of an optimal sequence of local structures is: <round, curved, straight,..., slanted, dark, light,...,>. This sequence of structures helps describing any objects. For example, autonomous robots based on this learning machine can be trained to recognize the components of a human face described as a sequence of shapes such as: <round (human head), vertical line in the middle of the face (nose), round (eyes), ellipse (mouth),...,>.
- **Parameter estimation (Training):** This problem consists of attempting to optimize the model parameters $\lambda = [\pi, \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}]$ to maximize $P(O \mid \lambda)$.

IV. Recognition Approach

A. DWT/PCA

1) Feature Extraction: As part of the multiresolution framework approach to face recognition described in this paper, the same method of feature extraction is used for all recognition techniques.

Each face image is divided into overlapping horizontal strips of height j pixels where the strips overlap by p pixels. Each horizontal strip is subsequently segmented vertically into blocks of width k pixels, with overlap of p. This is illustrated in Figure 5. For an image of width w and height h, there will be approximately $((h/(j-p))+1)*(w/(k-p))+1)$ blocks.

Each block then undergoes wavelet decomposition, producing an average image and a sequence of detail images. This can be shown as $[\mathrm{a_J}, \{d_j^1, d_j^2, d_j^3\}_{j=1\ldots J}]$ where $\mathrm{a_J}$ refers to the approximation image at the Jth scale and d_j^k is the detail image at scale j and orientation k. For the work described, 4-level wavelet decomposition was employed, producing a vector with one average image and twelve detail images. The L2 norms of the wavelet detail images were subsequently calculated and it is these that were used to form the observation vector for that block. The image norms from all the image blocks are collected, in the order the blocks appear in the image, from left-to-right and from top-to-bottom – this forms the image's observation vector.

2) Face Recognition using PCA: A face's feature vector undergoes PCA, resulting in an eigenface weight vector for the face image. This vector is compared with those of previously encoded training set face images and the Euclidean distances between it and the training set images are calculated. The identity of the training set face with the lowest Euclidean distance is returned as the closest match.

B. DWT/HMM

1) Feature Extraction: Feature extraction is conducted in the same manner to that used for PCA face recognition. The image is decomposed into blocks, with each block undergoing wavelet decomposition. The L2 norms of the detail images are then used to form an observation vector.

This vector, along with the observation vectors from all other training images of the same individual, are used to train the HMM for this individual using maximum likelihood (ML) estimation. As the detail image norms are real values, a continuous observation HMM is employed. One HMM is trained for each identity in the database.

2) Face Recognition using HMM: A number of images are used to test the accuracy of the face recognition system. In order to ascertain the identity of an image, a feature vector for that image is created in the same way as for the images used to train the system. For each trained HMM, the likelihood of that HMM producing the observation vector is calculated. The image is classified as the identity of the HMM that produces the highest likelihood value.

C. DWT/SHMM

1) Feature Extraction: SHMM modeling of the human face has never been undertaken by any researchers or practitioners in the biometric community. Our approach of adapting the SHMMs machine learning to recognize human faces is novel. The SHMM approach to face recognition consists of viewing a face as a sequence of blocks of information O_i that belong to some predefined facial regions as depicted in Figure 6. This phase involves extracting observation vector sequences from subimages of the entire face image. As with recognition using standard HMMs, DWT is used for this purpose.

The observation vectors are obtained by scanning the image from left to right and top to bottom using a fixed-size two dimensional window (called block) and performing DWT analysis at each subimage. The subimage is decomposed to a certain level and the energies of the subbands are selected to form the observation sequence O_i for the SHMM.

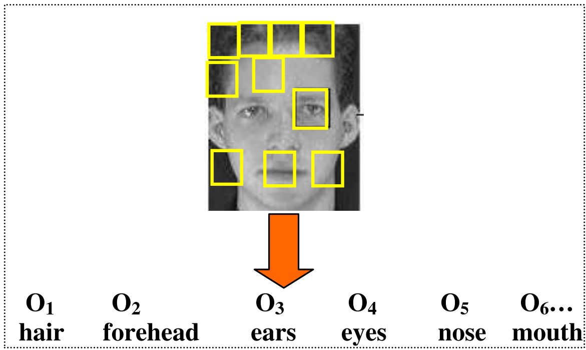

Fig. 6. A face O is viewed as an ordered sequence of observations O_i. Each O_i captures a significant facial region such as: "hair", "forehead", "eyes", "nose", "mouth", etc. These regions come in a natural order from top to bottom and left to right.

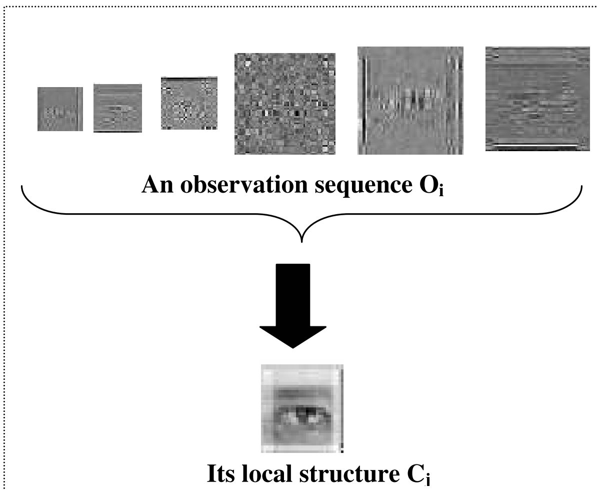

Fig. 7. A block O_i of the whole face O is a time-series of norms assigned to the multi-resolution detail images. This block belongs to the local structure "eyes".

The local structures of the SHMM correspond to the facial regions of the face. These regions are: hair, forehead, eyes, nose, mouth, etc. However, the observation sequence O_i corresponds to the different resolutions of the block images of the face. *The sequence of norms of the detail images $d_{i,k}$ represents the observation sequence O_i.* Each block is assigned one and only one facial region. Formally, a local structure C_j is simply an equivalence class that gathers all "similar" O_i. *Two $O_i^{\prime}s$ (two set of detail images) are equivalent if they share the same facial region of the human face.* Figure 7 depicts an example of a local structure and a sequence of observations. This modeling enables the SHMM to be trained efficiently since several sets of detail images are assigned to the same facial region.

2) Face Recognition using SHMM: The training phase of the SHMM consists of building a model $\lambda = [\pi, \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}]$ for each human face during a training phase. Each parameter of this model will be trained through the wavelet multi-resolution analysis applied to each face image of a person. The testing phase consists of decomposing each test image into blocks and assigning a facial region to each one of them. This phase is conducted via the k-means clustering algorithm. The value of k corresponds to the number of facial regions (or local structures). Each face is expressed as a sequence of blocks O_i with their facial regions C_i. The recognition phase will be performed by computing the model λ^* in the training set (database) that maximizes the likelihood of a test face image.

TABLE I
COMPARISON OF FACE RECOGNITION ACCURACY WHEN PERFORMED ON THE AT&T DATABASE (%)

	DWT/PCA	DWT/HMM	DWT/SHMM
Haar	94.3	95.75	97.5
Bio9/7	94.4	93.5	95.1
Coiflet(3)	96.8	96.5	97.8

TABLE II
COMPARISON OF FACE RECOGNITION ACCURACY WHEN PERFORMED ON THE ESSEX95 DATABASE (%)

	DWT/PCA	DWT/HMM	DWT/SHMM
Haar	86.1	84.2	89.4
Bio9/7	81.2	78.0	84.6
Coiflet(3)	87.1	85.6	90.7

V. RESULTS AND ANALYSIS

Experiments were carried out using two different face databases. The AT&T (formerly ORL) Database of Faces [13] contains ten grayscale images each of forty individuals. The images contain variation in lighting, expresssion and facial details (for example, glasses / no glasses). The Essex Faces95 database [24] contains twenty colour images each of seventy-two individuals. These images contain variation in lighting, expression and position. For the purposes of the experiments carried out, the Essex faces were converted to grayscale prior to training.

The experiments were carried out using five-fold cross validation. This involved splitting the set of training images for each person into five equally-sized sets and using four of the sets for system training with the other being used for testing. The experiments were repeated five times with a different set being used for testing each time, to provide a more accurate recognition figure.

It was assumed that all test images were from known individuals. Accuracy of an individual run is thus defined as the ratio of correct matches to the total number of face images tested, with final accuracy equalling the average accuracy figures from each of the five cross validation runs. For the case of DWT/PCA, the quoted accuracy figure is the highest that was achieved during experimentation (as the number of eigenfaces used for encoding influences the recognition rate).

Accuracy figures for the experiments are shown in Tables I and II.

TABLE III
COMPARITIVE RESULTS ON AT&T DATABASE

Method	Accuracy (%)	Ref
DCT/HMM	84	[9]
PCA	91	[25]
ICA	85	[6]
Gabor Filters & Rank Correlation	91.5	[26]
2D-PHMM	94.5	[13]
DWT/SHMM	97	(Proposed)

As can be seen from Table I, the best recognition results were achieved for the SHMM approach to face recognition. Indeed, the incorrect match rate for Haar / SHMM is 40% lower than for the equivalent HMM test. It is obvious that the rich structural information that the SHMM can represent contributes to higher recognition performance. PCA results were next highest and outperformed HMM recognition in all cases. Results for the Essex95 database showed a similar pattern, although the results overall were lower due to the larger size of the database.

The Coiflet(3) wavelet produced better results than the other wavelet filters. This is not surprising as the Coiflet wavelet family was designed to overcome some of the representation shortfalls exhibited by the Daubechies wavelets. The biorthogonal 9/7 wavelet also performs well – this is consistent with previous results which have shown the biorthogonal 9/7 wavelet to be particularly effective for image applications such as compression and segmentation. The Haar wavelet performed least well – its simplistic structure means that it is unable to capture more detailed image information.

For comparison purposes, an experiment was performed to find the accuracy for DWT/SHMM when using five images from the AT&T database for training and five images for testing. As Table III shows, the DWT/SHMM approach to face recognition compares well with other techniques from the literature that have used this training set.

VI. CONCLUSION

We have presented a framework for multiresolution face recognition. This framework demonstrates the impact of DWT on three classifiers: PCA, HMM and SHMM. Furthermore, we have introduced a novel approach DWT/SHMM that optimally combines the multiresolution effect with the local structure concept into one single probabilistic model. Experiments have revealed that DWT/SHMM is a promising approach since it has outperformed the other combinations. Finally, we have provided a computational grids-based solution to address the problem of distributed and large databases.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, 2003.

[2] R. Gross, S. Baker, I. Matthews, and T. Kanade, "Face recognition across pose and illumination," in *Handbook of Face Recognition*, S. Z. Li and A. K. Jain, Eds. Springer-Verlag, June 2004.

[3] A. Amira and P. Farrell, "An automatic face recognition system based on wavelet transforms," *Proceedings of 2005 International Symposium on Circuits and Systems*, vol. 6, pp. 6252 – 6255, May 2005.

[4] H. Moon and P. Phillips, "Computational and Performance Aspects of PCA-based Face Recognition Algorithms," *Perception*, vol. 30, pp. 303 – 321, 2001.

[5] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, pp. 71 – 86, 1991.

[6] P. C. Yuen and J.-H. Lai, "Face representation using independent component analysis." *Pattern Recognition*, vol. 35, no. 6, pp. 1247–1257, 2002.

[7] E. Kussul, T. Baidyk, and M. Kussul, "Neural network system for face recognition." in *ISCAS (5)*, 2004, pp. 768–771.

[8] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," pp. 267–296, 1990.

[9] A. Nefian and M. Hayes, "Hidden markov models for face recognition," in *ICASSP98*, 1998, pp. 2721–2724.

[10] U. M.Bicego and V.Murino, "Using hidden markov models and wavelets for face recognition," in *ICIAP03*, 2003, pp. 52–56.

[11] I. Daubechies, "Wavelet transforms and orthonormal wavelet bases," in *Different perspectives on wavelets (San Antonio, TX, 1993)*, ser. Proc. Sympos. Appl. Math. Providence, RI: Amer. Math. Soc., 1993, vol. 47, pp. 1–33.

[12] H.-S. Le and H. Li, "Recognizing frontal face images using hidden markov models with one training image per person." in *ICPR (1)*, 2004, pp. 318–321.

[13] F. Samaria, "Face recognition using hidden markov models," Ph.D. dissertation, Cambridge University Engineering Department, 1994.

[14] D. Bouchaffra and J. Tan, "Introduction to structural hidden markov models: Application to handwritten numeral recognition," *Intelligent Data Analysis Journal*, vol. 10:1, 2006.

[15] D.Bouchaffra and J. Tan, "Structural hidden markov models using a relation of equivalence: Application to automotive designs," *Data Mining and Knowledge Discovery Journal*, vol. 12:1, 2006.

[16] P. Y. G.C. Feng and D. Dai, "Human Face Recognition Using PCA on Wavelet Subband," *Journal of Electronic Imaging*, vol. 9, pp. 226 – 233, April 2000.

[17] M. T. Harandi, M. N. Ahmadabadi, and B. N. Araabi, "Face recognition using reinforcement learning." in *ICIP*, 2004, pp. 2709–2712.

[18] H. K. Ekenel and B. Sankur, "Multiresolution face recognition." *Image Vision Comput.*, vol. 23, no. 5, pp. 469–477, 2005.

[19] M. Safari, M. T. Harandi, and B. N. Araabi, "An svm-based method for face recognition using a wavelet pca representation of faces." in *ICIP*, 2004, pp. 853–856.

[20] C. Travieso, J. Alonso, M. Ferrer, and L. Sanson, "Facial identification using transformed domain by svm," in *IEEE International Carnahan Conference on Security Technology*, October 2004, pp. 321–324.

[21] I.Foster and C.Kesselman, "Globus: A metacomputing infrstructure toolkit," *The Internationl Journal of SupercomputerApplications and High Performance Computing*, vol. 11, no. 2, pp. 115–128, 1997.

[22] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation." *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 11, pp. 674–693, 1989.

[23] F. Samaria and A. Harter, "Parameterisation of a Stochastic Model for human Face Identification," in *IEEE Workshop on Applications of Computer Vision*, Sarasota (Florida), December 1994.

[24] D.Hond and L.Spacek, "Distinctive descriptions for face processing." in *BMVC*, 1997, pp. 320–329.

[25] P. Nicholl, A. Amira, and R. Perrott, "An Automated Grid-enabled Face Recognition System using Hybrid Approaches," *Proceedings of the 5th IEE/IEEE Postgraduate Research Conference on Electronics, Photonics, Communications and Networks (PREP'05)*, pp. 144 – 146, 2005.

[26] O. Ayinde and Y.-H. Yang, "Face recognition approach based on rank correlation of gabor-filtered images." *Pattern Recognition*, vol. 35, no. 6, pp. 1275–1289, 2002.