7

Wavelet–Based Face Recognition Schemes

Sabah A. Jassim
University of Buckingham, Buckingham MK18 1EG
United Kingdom

Abstract. The growth in direct threats to people's safety in recent years and the rapid increase in fraud and identity theft has increased the awareness of security requirements in society and added urgency to the task of developing biometric-based person identification as a reliable alternative to conventional authentication methods. In this Chapter we describe various approaches to face recognition with focus on wavelet-based schemes and present their performance using a number of benchmark databases of face images and videos. These schemes include single-stream (i.e. those using single-subband representations of face) as well as multi-stream schemes (i.e. those based on fusing a number of wavelet subband representations of face). We shall also discuss the various factors and quality measures that influence the performance of face recognition schemes including extreme variation in lighting conditions and facial expressions together with measures to reduce the adverse impact of such variations. These discussions will lead to the introduction of new innovative adaptive face recognition schemes. We shall present arguments in support of the suitability of such schemes for implementation on mobile phones and PDA's.

1. Introduction

The early part of the 21st century has ushered the shaping of a new global communication infrastructure that is increasingly dominated by new generations of mobile phones/devices including 3G and beyond devices resulting in the emergence of pervasive computing environment with less reliance on presence in specific locations or at specific times. The characteristics of such a ubiquitous environment create new security threats and the various mobile devices/nodes are expected to provide additional layers of security for online transactions and real-time surveillance. Cryptography can provide confidentiality protection mechanisms for online and mobile transactions, but authenticating/identifying the principal(s) in such virtual transactions is of utmost importance to fight crime and fraud and to establish trust between parties taking part in such transactions. Traditional authentication mechanisms are based on "something you know" (e.g. a password/PIN) or "something you own/hold" (e.g. a token/smartcard). Such authentication schemes have shown to be prone to serious threats that could have detrimental effects on global economic activities. In recent years, biometric-based authentication has provided a new approach of access control that is aimed at establishing "who you are", and research in the field of biometrics has grown rapidly. The scope of active research into biometrics has gone beyond the traditional list of

single traits of fingerprint, retina, iris, voice, and face into newly proposed traits such as handwritten signature, gait, hand geometry, and scent. Moreover, the need for improved performance has lead to active research into multimodal biometrics based on fusing a number of biometrics traits at different levels of fusion including feature level, score level, and decision level. Over the past two decades significant progress has been made in developing robust biometrics that helped realising large-scale automated identification systems.

Advances in mobile communication systems and the availability of cheap cameras and other sensors on mobile devices (3G smart phones) further motivate the need to develop reliable, and unobtrusive biometrics that are suitable for implementation on mobile and constrained devices. Non-intrusive biometrics, such as face and voice are more naturally acceptable as the person's public identity. Unfortunately the performance of known face and voice biometric schemes are lower than those of the Iris or the fingerprint schemes. The processing and analysis of face image suffer from the curse of dimension problem, and various dimension reduction schemes have been proposed including PCA (principal Component analysis). In recent years a number of wavelet-based face verification schemes have been proposed as an efficient alternative to traditional dimension reduction procedures.

The Wavelet Transform is a technique for analyzing finite-energy signals at multi-resolutions. It provides an alternative tool for short time analysis of quasi-stationary signals, such as speech and image signals, in contrast to the traditional short-time Fourier transform. A wavelet-transformed image analyses the signal into a set of subbands at different resolutions each represented by a different frequency band. Each wavelet subband encapsulates a representation of the transformed images object(s), which differ from the others in scale and/or frequency content. Each wavelet subband of transformed face images can be used as a face biometric template for a face recognition scheme, and the fusion of a multiple of such schemes associated with different wavelet subbands will be termed as multi-stream face recognition scheme.

2. Face Recognition - A brief review

Automatic face based human Identification is a particularly tough challenge in comparison to identification based on other biometric features such as iris, fingerprints, or palm prints. Yet, due to its unobtrusive nature, together with voice, the face is naturally the most suitable method of identification for security related applications, ([1], [2], [3]). Recent growth in identity theft and the rise of international terrorism on one hand and the availability of high-resolution digital video-cameras at a relatively low cost is a major driving force in the surge of interest for efficient and accurate enrolment, verification schemes of face-based authentication. Moreover, there are now new opportunities, as well as tough challenges, for mass deployments of biometric-based authentications in a range of civilian and military applications. In the rest of this section we shall briefly review the main approaches to face recognition with focus on 2D schemes directly related to this chapter's aim.

2.1 Dimension reduction approach

An important part of a face recognition process is the feature extraction of a given facial image. Two current approaches to feature extraction are the geometry feature-based methods and the more common template-based methods. In the latter, sets of face images

are statistically analysed to obtain a set of feature vectors that best describe face image. A typical face images is represented by a high dimensional array (e.g. 12000=120×100 pixels), the processing/analysis of which is a computationally demanding task, referred to in the literature as the "curse of dimensionality", well beyond most commercially available mobile devices. It is therefore essential to apply dimension reduction procedures that reduce redundant data without losing significant features. A common feature of dimension reducing procedures is a linear transformation of the face image into a "significantly" lower dimensional subspace from which a feature vector is extracted. The first and by far the most commonly used dimension reduction method is the Principal Component Analysis (PCA), also known as Karhunen-Love (KL) transform, [4]. In [5], M. Turk and Pentland used the PCA technique to develop the first successful and well known Eigenface scheme for face recognition. PCA requires the use of a sufficiently large training set of multiple face images of the enrolled persons, and attempts to model their significant variation from their average image, by taking a number of unit eigenvectors corresponding to the "most significant" eigenvalues (i.e. of largest absolute values). Essentially, the selected eigenvectors are used as the basis for a linear transformation that maps the original training set of face images around their mean in order to align with the directions the first few principal components which maximizes the variance as much of the as possible. The values in the remaining dimensions (corresponding to the non-significant eigenvalues), tend to be highly correlated and dropped with minimal loss of information.

Despite its success in reducing false acceptances, the PCA/Eigenface scheme is known to retain within-class variations due to many factors including illumination and pose. Moghaddam et al. [6] have demonstrated that the largest three eigen coefficients of each class overlap each other. While this shows that PCA has poor discriminatory power, it has been demonstrated that leaving out the first 3 eigenfaces (corresponding to the 3 largest eigenvalues) could reduce the effect of variations in illumination [6]. But this may also lead to loss of information that is useful for accurate identification.

An alternative approach to PCA based linear projection is Fisher's Linear Discriminant (FLD), or the Linear Discriminant Analysis (LDA) which is used to maximize the ratio of the determinant of the between class scatter to that of within-class scatter [7], [8]. The downside of these approaches is that a number of training samples from different conditions are required in order to identify faces in uncontrolled environments.

Other schemes that deal with the curse of dimension include Independent Component Analysis (ICA), or a combination of ICA and LDA/FLD, (see [1], [7], and [9]). Lack of within-class (variations in appearance of the same individual due to expression and/or lighting) information is known to hinder the performance of both PCA and ICA based face recognition schemes. Cappelli et al., [9], proposed a multi-space generalization of KL-transformation (MKL) for face recognition, in which a PCA-subspace is created for each enrolled classes. The downside of this approach is that a large number of images are required to create a subspace for each class.

All the statistical approaches above require a large number of training images to create a subspace, which in turn requires extra storage space (for the subspace and enrolled template/features), [10]. Current mobile devices (3G smart phones) and smartcards, which are widely used in commercial and military applications, have limited computing resources and it is difficult to implement complex algorithms, especially for face verification. Bicego et al. presented a face verification scheme based on Hidden Markov Models (HMM). Statistical

features such as the mean and variance are obtained by overlapping sub images (of a given original face image). These features are used to compose the HMM sequence and results show that the HMM-based face verification scheme, proposed by Bicego et al., outperforms other published results, [11].

2.2 Frequency transforms based approaches

Frequency transforms provide valuable tools for signal processing and analysis. Frequency information content conveys richer knowledge about features in signals/images that should be exploited to complement the spatial information. Fourier and wavelet transforms are two examples that have been used with significant success in image processing and analysis tasks including face recognition. To some extent, such transforms reduce dimension with no or little loss of information.

The work of John Daugman, ([12], [13]) and others on generalisation of Gabor functions has led to a general compact image representation in terms of Gabor wavelets. The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells, exhibit strong characteristics of spatial locality and orientation selectivity, and are optimally localized in the space and frequency domains. The Gabor wavelet model was eventually exploited to develop new approaches to face recognition. Taking into account the link between the Gabor wavelet kernels and the receptive field profiles of the mammalian cortical simple cells, it is not unreasonable to argue that Gabor wavelet based face recognition schemes mimics the way humans recognise each others.

Lades et al. [14] demonstrated the use of Gabor wavelets for face recognition using the Dynamic Link Architecture (DLA) framework. The DLA starts by computing the Gabor jets, and then it performs a flexible template comparison between the resulting image decompositions using graph-matching. L Wiskott et al, [15], have expanded on the DLA, and developed the Elastic Bunch Graph Matching (EBGM) face recognition system, whereby individual faces were represented by a graph, each node labelled with a set of complex Gabor wavelet coefficients, called a jet. The magnitudes of the coefficients label the nodes for matching and recognition, the phase of the complex Gabor wavelet coefficients is used for location of the nodes. The nodes refer to specific facial landmarks, called fiducial points. A data structure, called the bunch graph, to represent faces by combining jets of a small set of individual faces. Originally many steps (e.g. selecting the Fiducial points) were carried out manually, but gradually these would have been replaced with a automated procedures.

Z. Zhang et al, [16], compared the performance of a Geometry-based and a Gabor wavlet-based facial expression recognition using a two-layer perceptron. The first uses the geometric positions of a set of fiducial points on a face, while the second type is a set of multi-scale and multi-orientation Gabor wavelet coefficients extracted from the face image at the fiducial points. For the comparison they used a database of 213 images of female facial expressions and their results show that the Gabor wavelet –based scheme outperforms the geometric based system.

C. Lui and H. Wechsler, [17], developed and tested an Independent Gabor Features (IGF) method for face recognition. The IGF first derives a Gabor feature vector from a set of downsampled Gabor wavelet representations of face images, then reduces the dimensionality of the vector by means of Principal Component Analysis (PCA), and finally defines the independent Gabor features based on the Independent Component Analysis (ICA). The independence property of these Gabor features facilitates the application of the

Probabilistic Reasoning Model (PRM) method for classification. The Gabor transformed face images exhibit strong characteristics of spatial locality, scale and orientation selectivity, while ICA further reduce redundancy and represent independent features explicitly.
The development of the discrete wavelet transforms (DWT), especially after the work of I. Daubechies (see e.g. [18]), and their multi-resolution properties have naturally led to increased interest in their use for image analysis as an efficient alternative to the use of Fourier transforms. DWT's have been successfully used in a variety of face recognition schemes (e.g. [10], [19], [20], [21], [22]). However, in many cases, only the approximation components (i.e. the low frequency subbands) at different scales are used either as a feature vector representation of the faces perhaps after some normalisation procedures or to be fed into traditional face recognition schemes such as the PCA as replacement of the original images in the spatial domain.
J. H. Lai et al, [23], developed a holistic face representation, called spectroface, that is based on an elaborate combination of the (DWT) wavelet transform and the Fourier transform. To make the spectroface invariant to translation, scale and on-the-plane rotation, the LL wavelet subband of the face image is subjected to two rounds of transformations. The LL wavelet subband is less sensitive to the facial expression variations while the first FFT coefficients are invariant to the spatial translation. The second round of FFT is applied after the centralised FFT in the first round is represented by polar coordinates. Based on the spectroface representation, their proposed face recognition system is tested on the Yale and Olivetti face databases. They report recognition accuracy of over 94% for rank1 matching, and over 98% for rank 3 matching.
Another wavelet-based approach for face recognition has been investigated in terms of dual-tree complex wavelets (DT-CW) techniques developed by N. G. Kingsbury, (see e.g. [24]). Y. Peng et al, [25], propose face recognition algorithm that is based on the use of an anisotropic dual-tree complex wavelet packets (ADT-CWP) for face representation. The ADT-CWP differs from the traditional DT-CW in that the decomposition structure is determined first by an average face, which is then applied to extracting feature of each face image. The performance of their scheme is compared with the traditional Gabor-based methods using a number of different benchmark databases. The AD-CWP method seems to outperform the Gabor-based schemes and it is computationally more efficient.
The rest of the chapter is devoted to DWT-based face recognition tasks. We shall first give a short description of the DWT as a signal processing and analysis tool. We then describe the most common approaches to wavelet-based multi-stream face recognition.

3. Wavelet Transforms

The Wavelet Transform is a technique for analyzing finite-energy signals at multi-resolutions. It provides an alternative tool for short time analysis of quasi-stationary signals, such as speech and image signals, in contrast to the traditional short-time Fourier transform. The one dimensional Continuous Wavelet Transform CWT of f(x) with respect to the wavelet Ψ(x) is defined as follows:

$$\psi_f(j,k) = \langle f, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx$$

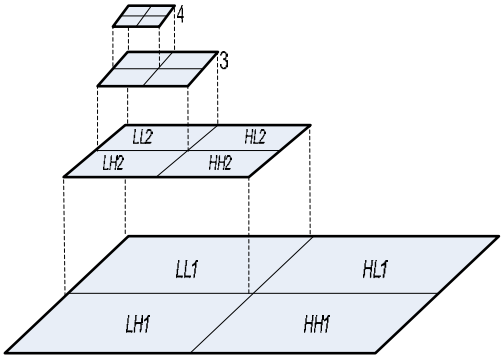
i.e. wavelet transform coefficients are defined as inner products of the function being transformed with each of the base functions $\Psi_{j,k}$. The base functions are all obtained from a single wavelet function $\Psi(x)$, called the mother wavelet, through an iterative process of scaling and shifting, i.e.

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k).$$

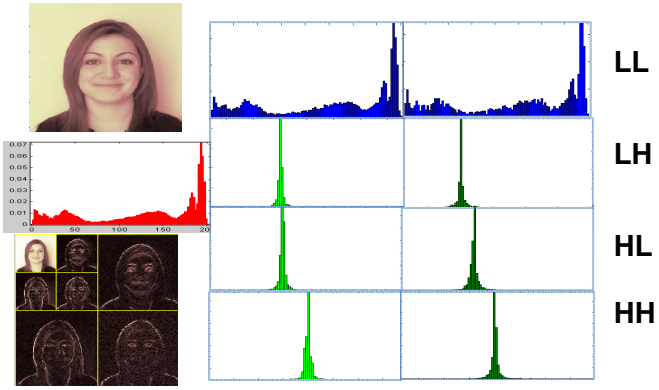
A wavelet function is a wave function that has a finite support and rapidly diminishes outside a small interval, i.e. its energy is concentrated in time. The computation of the DWT coefficients of a signal k does not require the use of the wavelet function, but by applying two Finite Impulse Response (FIR) filters, a high-pass filter h, and a low-pass filter g. This is known as the Mallat's Algorithm. The output will be in two parts, the first of which is the detail coefficients (from the high-pass filter), and the second part is the approximation coefficients (from the low-pass filter). For more details see [26].

The Discrete Wavelet Transform (DWT) is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed very efficiently. The DWT is used to decompose a signal into frequency subbands at different scales. The signal can be perfectly reconstructed from these subband coefficients. Just as in the case of continuous wavelets, the DWT can be shown to be equivalent to filtering the input image with a bank of bandpass filters whose impulse responses are approximated by different scales of the same ***mother wavelet***. It allows the decomposition of a signal by successive highpass and lowpass filtering of the time domain signal respectively, after sub-sampling by 2. Consequently, a wavelet-transformed image is decomposed into a set of subbands with different resolutions each represented by a different frequency band. There are a number of different ways of doing that (i.e. applying a 2D-wavelet transform to an image). The most commonly used decomposition scheme is the *pyramid* scheme. At a resolution depth of k, the pyramidal scheme decomposes an image I into 3k +1 subbands, $\{LL_k, LH_k, HL_k, HH_k, LH_{k-1}, HL_{k-1}, \ldots, LH_1, HL_1\}$, with LL_k, being the lowest-pass subband, (see figure 3.1(a)).

There are ample of wavelet filters that have been designed and used in the literature for various signal and image processing/analysis. However, for any wavelet filter, the LL subband is a smoothed version of original image and the best approximation to the original image with lower-dimensional space. It also contains highest-energy content within the four subbands. The subbands LH_1, HL_1, and HH_1, contain finest scale wavelet coefficients, and the coefficients LL_k get coarser as k increases. In fact, the histogram of the LL_1-subband coefficients approximates the histogram of the original image in the spatial domain, while the wavelet coefficients in every other subband has a Laplace (also known as generalised Gaussian) distribution with $\cong 0$ mean, see Figure 3.1(b). This property remains valid at all decomposition depth. Moreover, the furthest away a non-LL coefficient is from the mean in that subband, the more probable the corresponding position(s) in the original image have a significant feature, [27]. In fact the statistical properties of DWT non-LL subbands can be exploited for many image processing applications, including image/video compression, watermarking, content-based video indexing, and feature extraction.



(a)



(b)

Fig. 3.1. (a) The Multi-resolution Pyramid (b) An image, its WT to level 2 and subbands histograms.

4. DWT – based Face recognition

The LL subband of a wavelet transformed image corresponds to the low frequency components in both vertical and horizontal directions of the original image. Therefore, it is the low frequency subband of the original image. The subband LH corresponds to the low-frequency component in the horizontal direction and high-frequency components in vertical direction. Therefore it holds the vertical edge details. Similar interpretation is made on the subbands HL and HH. These remarks together with our knowledge of structure of facial features provide a strong motivation and justification for investigating wavelet-based approaches to face recognition. In fact the variety of wavelet decomposition schemes and filter banks provide a very rich and a complex source of information that could be exploited to deal with the tough challenges and difficulties associated with face recognition in the presence of expression and extreme variations in lighting conditions.

With appropriate pixel value scaling the low LL subband, displayed as an image, looks like a smoothing of the original image in the spatial domain (see Figure 3.1(b)). For efficiency

purposes and for reason of normalising image sizes, non-DWT based face recognition schemes such as PCA pre-process face images first by resizing/downsampling the images. In such cases, matching accuracy may suffer as a result of loss of information. The LL subbands of the face image, does provide a natural alternative to these pre-processing procedures and this has been the motivation for the earlier work on wavelet-based face recognition schemes that have mostly combined with LDA and PCA schemes (e.g. [10], [28], [29], [30], [31]). Below, we shall describe face recognition schemes, developed by our team, that are based on the PCA in a single wavelet subband and summarise the results of performance tests by such schemes for some benchmark face databases. We will also demonstrate that the use of the LL-subband itself as the face feature vector results in comparable or even higher accuracy rate. These investigations together with the success of biometric systems that are based on fusing multiple biometrics (otherwise known as multi-modal biometrics) have motivated our work on multi-stream face recognition. This will be discussed in section 5.

4.1 PCA in the Wavelet Domain

Given the training set Γ of images, applying a wavelet transform on all images results in a set $W_k(\Gamma)$ of multi-resolution decomposed images. Let $L_k(\Gamma)$ be the set of all level-k low subbands corresponding to the set $W_k(\Gamma)$. Apply PCA on the set $L_k(\Gamma)$ whose elements are the training vectors in the wavelet domain (i.e. LL_k subbands). Note that each wavelet coefficient in the LLk subband is a function of $2^k x 2^k$ pixels in the original image representing a scaled total energy in the block. Figure 3.2, below, shows the first 4 eigenfaces obtained from a dataset of images in the spatial domain as well as in the low subbands at levels 1 and 2 using the Haar wavelet filter. There are no apparent differences between the eigenfaces in the spatial domain and those in the wavelet domain.

Fig. 3.2. Eigenfaces in (a) spatial domain, (b) LL1, and (c) LL2

Diagram1, below, illustrates the enrolment and matching steps which will cover face recognition in the wavelet domain with and without the application of PCA. The diagram applies equally to any wavelet subband including the high frequency ones.
There are many different wavelet filters to use in the transformation stage, and the choice of the filter would have some influence on the accuracy rate of the PCA in the wavelet domain. The experiments are designed to test the effect of the choice of using PCA or not, the choice of wavelet filter, and the depth of decomposition. The performance of the various possible schemes have been tested for a number of benchmark databases including ORL (also known as AT&T see http://www.uk.research.att.com/facedatabase.html), and the controlled

section of the BANCA, [32]. These datasets of face images do not involve significant variation in illumination. The problem of image quality is investigated in section 6. Next we present a small, but representative, sample of the experimental results for few wavelet filters.

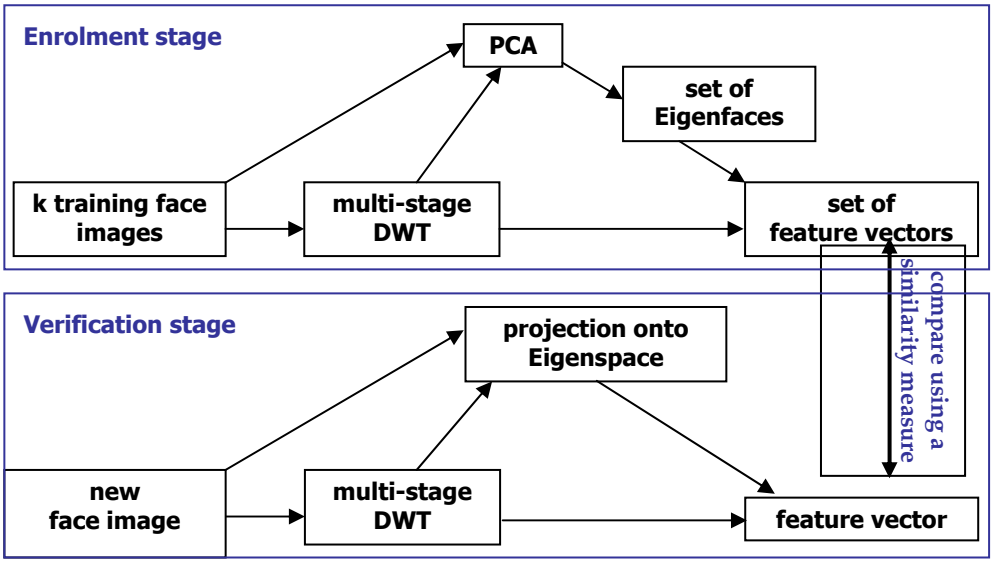


Diagram 3.1 Verification scheme.

The ORL Experiments.

- **Enrolment/Training module** There are 40 subjects in ORL. In the first instance, we split the subjects into two equal groups: Group1 and Group2. For each group we trained the system with 4 different sets of five images for each subject. These sets were, respectively, frames 1-5, frames 6-10, even indexed frames and odd indexed frames. In total, we conducted 8 different training sessions for these groups.
- ***The Testing Module.*** In each of the training sessions that consisted of 20 subjects, the remaining 100 images of the trained subjects as well as 100 impostor images (5 images per subject, selected in the same way as in the training scheme) were used to test the many-to-one identification schemes.

Chart 3.1, below, contains the test results of experiments that were designed to test the verification accuracy when the Haar wavelet filter is used to different level of decompositions. It shows the average accuracy rates for the various identification schemes measured using different number of eigenfaces (20, 30, 40 and 50) for the schemes that involve the use of PCA. The results indicate that regardless of the number of eigenvalues chosen, PCA in the LL subbands outperform the PCA in the spatial domain, and LL3 is the best performing subband in this case. Moreover, the wavelet LL3 scheme without the PCA achieves the best performance. Another interesting observation, not explicit in this chart, is that the accuracy rates for the various training schemes vary widely around the stated

averages, indicating that accuracy can be improved further by making a *careful* choice of the training images for the enrolled subjects.

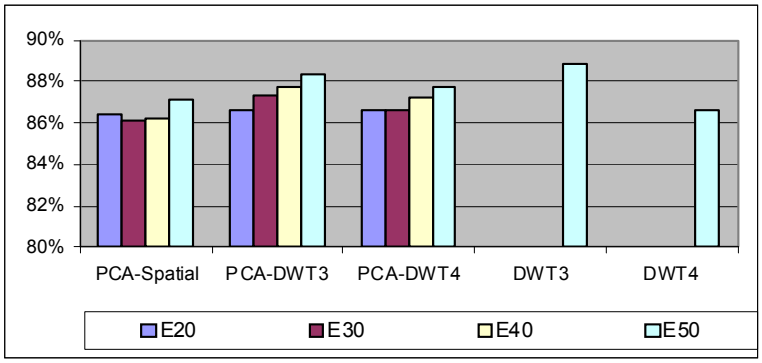


Chart 3.1. Identification accuracy for Spatial PCA, Wavelet PCA, and Wavelet-only features

The superior performance of the wavelet-only scheme compared to the other schemes, has desirable implication beyond the computational efficiency. While most conventional face recognition schemes require model/subspace training, wavelet-based recognition schemes can be developed without the need for training, i.e. adding/removing classes do not require rebuilding the model from scratch.

Jen-Tzung Chien etal ([10]) who used all the 40 subjects of ORL to test the performance of a number of recognition schemes including some of the wavelet-based ones investigated here. In those experiments, there were no impostors, i.e. untrained subjects. Thus we conducted experiments where all the 40 subjects were used for training. We trained the system 4 times each with a set of 5 different frames and in each case the remaining 200 images (5 frames for each subject) in the database were used for testing.. On average, all schemes have more or less achieved similar accuracy rate of approximately 89%. Similar experiments with 35 trained subjects, the rest being impostors, have been conducted but in all cases the results were similar to those shown above.

Chart 3.2 contains the results of verifications rather identifications. The experiments were carried out to test the performance of wavelet-based verification schemes, again with and without PCA. Here, two filters were used, the Haar as well as the Daubechies 4 wavelets, and in the case of Daubechies 4 we used two versions whereby the coefficients are scaled for normalisation in the so called Scaled D3/d4. The results confirmed again the superiority of PCA in the wavelet domain over PCA in the spatial, and the best performance was obtained when no PCA was applied. The choice of filter does not seem to make much difference at level 3, but Haar outperforms both versions of Daubechies 4.

The superiority of the PCA in the wavelet domain over the PCA in the spatial domain can be explained in terms of the poor within class variation of PCA and the properties of the linear transform defined by the low-pass wavelet filter. The low-pass filter defines a contraction mapping of the linear space of the spatial domain into the space where LL subbands resides (i.e. for any two images the distance between the LL-subbands of two images is less than that between the original images). This can easily be proven for the Haar filter. This will help reduce the within class variation.

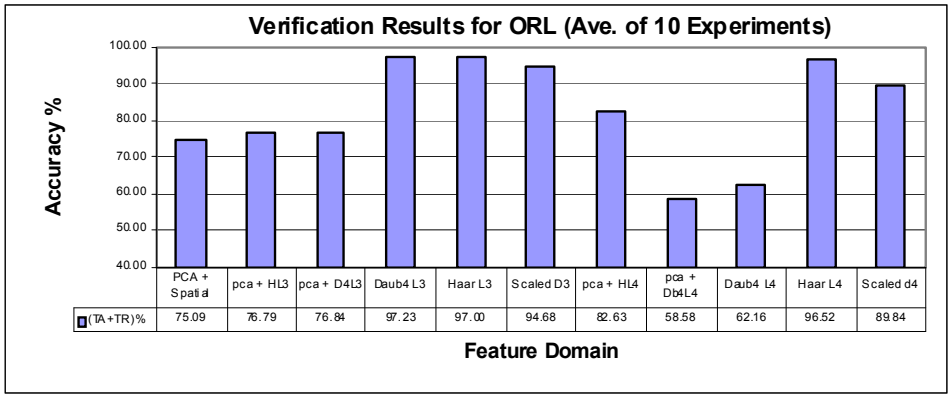


Chart 3.2.Verification accuracy for Spatial PCA, Wavelet PCA, and Wavelet-only features

The trend in, and the conclusions from these experiments are confirmed by other published data. For example, C.G.Feng et al, [33] have tested and compared the performance of PCA in the spatial domain and in wavelet subbands at different levels for the Yale database. Table 3.2, below, reports the recognition accuracy for the Daubechies 4 filter and confirms our conclusions. Note that the inter-class separation experiment in [33] can be seen to demonstrate that the contraction mapping nature of the low-pass filter transformation does not have adverse impact on the inter-class separation.

Method	PCA on original image	PCA on WT subband 1 Image	PCA on WT subband 2 Image	PCA on WT subband 3 Image	Proposed Method (PCA on WT subband 4 image)
Accuracy	78.78%	75.75%	83.03%	81.18%	85.45%

Table 3.2 [¥]. Performance comparison using Yale database

5. Multi-stream face Recognition

A wavelet transformed image is decomposed into a set of frequency subbands with different resolutions. Each frequency subband gives rise to a different feature vector representation of the face and has the potential to be used for recognition. The performances of such schemes vary significantly depending on many factors including the chosen wavelet filter, the depth of decomposition, the similarity measure, the sort of processing that the corresponding coefficients are subjected to, and the properties of subband as described at the end of section 3. The fact that identification schemes that are based on the fusion of different biometric modalities have shown to significantly outperform the best performing single modality scheme, raises the possibility of fusing different signal representing the same modality. Moreover, different subbands of wavelet-transformed face images, each representing the

[¥] Reproduced from [CGFeng].

face in different way, makes them perfect candidates for fusion without costly procedures. Diagram 2, below, depicts the stages of the wavelet based multi-stream face recognition for 3 subbands at level 1, but this could be adopted for any set of subbands at any level of decomposition.

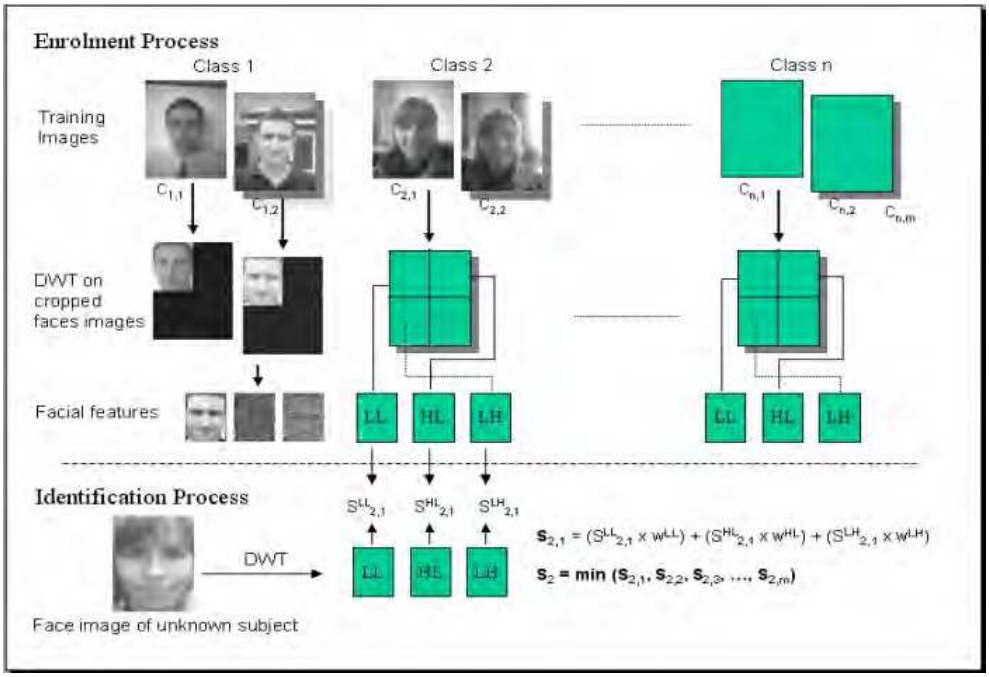


Diagram 2. Multi-stream Wavelet face recognition scheme

In this section we shall investigate the viability of fusing these streams as a way of improving the accuracy of wavelet-based face recognition. We shall establish that the fusion of multiple streams of wavelet-based face schemes does indeed help significantly improve single stream face recognition. We have mainly experimented with the score fusion of wavelet subbands at one decomposition depth. Limited experiments with other level of fusion did not achieve encouraging results.

The experiments reported here are based on the performance of the multi-stream face wavelet recognition for databases that involve face images/videos captured under varying recording conditions and by cameras of different qualities. These databases are the Yale database, and the BANCA audio-visual database. More extensive experiments have been conducted on the PDAtabase audio-visual database of videos recorded on a PDA within the SecurePhone EU-funded project (www.secure-phone.info).

5.1 The Yale database experiments

Identification experiments reported in table 5.1., below, are based on the "leave one out" protocol. The table contain the performance of 3 single wavelet–based down to level 3 (the LL3, LH3 and HH3 subbands schemes), the fusion of the three subband streams for a

selection of fixed weight combinations and for comparison we include results from some of the best performing face recognition schemes reported in Yang, [34], and Belhuemer et al, [35]. These results demonstrate that among the single subband streams, the LH3 is the best performing one. The multi-stream fusion of the three subbands for all but one weight configuration outperform the best single stream scheme, illustrating the conclusion that the multi-stream approach yields improved performance. Comparing the results with those from the state of the art schemes reported in [14] and [26] shows that the multi-steam fusion of the two single streams LH3 and HL3 subbands outperform all but 3 of the SOTA schemes. One can predict with confidence that the multi-stream fusing of several subbands at different level of decomposition would result in significantly improved performance.

Method	Features/Weights			Error Rate(%)
	LL_3	HL_3	LH_3	
Single-stream	1	0	0	23.03
	0	1	0	14.55
	0	0	1	12.73
Multi-stream	0	0.4	0.6	9.70
	0	0.3	0.7	9.09
	0	0.2	0.8	9.70
	0.1	0.3	0.6	10.91
	0.1	0.25	0.65	10.91
	0.2	0.2	0.6	12.73
	0.2	0.3	0.5	12.12
	0.2	0.4	0.4	13.33
Yang[26]	Eigenface(EF_{30})			28.48
	Fisherface (FF_{14})			8.48
	ICA			28.48
	SVM			18.18
	K.Eigenface (EF_{60})			24.24
	k.Fisherface (FF_{14})			6.06
Belhumeur et al.[14] (Full Face)	Eigenface (EF_{30})			19.40
	Eigenface (EF_{30}, w/o 1st 3 EF)			10.8
	Correlation			20.00
	Linear Subspace			15.60
	Fisherface			0.60

Table 5.1. Fusion Experiments – Yale database

5.2 BANCA database experiments

Experiments reported here are only a small, but representative, sample conducted on the BANCA database and is limited to the use of the G evaluation protocol, [32]. The experiments are conducted on the English section of the database which include recording for 52 subjects. Each subject participated in 12 sessions, each consisting of two short clips

uttering a true-client text while in the second clip he/she acts as an impostor uttering a text belonging to another subject. The 12 sessions are divided into 3 groups:

- the controlled group – sessions 1-4 (high quality camera, controlled environment and a uniform background)
- the degraded group – sessions 5-8 (in an office using a low quality web camera in uncontrolled environment).
- the adverse group – sessions 9-12 (high quality camera, uncontrolled environment)

For the G evaluation protocol, the true client recordings from session 1, 5, and 9 were used for enrolment and from each clip 7 random frames were selected to generate the client templates. True-client recordings from sessions 2, 3, 4, 6, 7, 8, 10, 11, and 12 (9 videos) were used for testing the identification accuracy. From each test video, we selected 3 frames and the minimum score for these frames in each stream was taken as the score of the tested video in the respective stream. In total, 468 tests were conducted. Identification accuracies of single streams (first 3 rows) and multi-stream approaches for the G protocol are shown in Table 5.2. Across all ranks the LH-subband scheme significantly outperformed all other single streams. The multi-stream fusion of the 3 streams outperformed the best single stream (i.e. the LH subband) by a noticeable percentage. The best performing multi-stream schemes are mainly the ones that give >0.5 weight to the LH subband and lower weight to the LL-subband. Again these experiments confirm the success of the multi-stream approach.

Weights			Identification Accuracy for Rank n for G test configuration									
LL	HL	LH	1	2	3	4	5	6	7	8	9	10
1.00	0.00	0.00	58.55	67.95	72.65	77.14	80.34	82.69	84.62	85.47	86.54	87.61
0.00	1.00	0.00	56.84	65.60	72.22	76.28	79.49	82.26	84.83	85.68	87.18	87.82
0.00	0.00	1.00	69.23	80.77	85.68	89.10	91.45	92.52	93.38	94.44	95.09	95.94
0.20	0.20	0.60	76.28	85.47	88.89	90.81	92.74	93.38	94.44	95.73	96.37	96.79
0.20	0.30	0.50	76.07	83.12	88.46	91.45	93.38	93.80	95.09	95.30	95.73	96.15
0.25	0.35	0.40	74.15	81.62	87.61	89.74	91.24	92.31	92.95	94.66	95.30	95.30
0.10	0.30	0.60	76.71	85.90	89.32	92.09	93.16	93.80	94.87	95.51	96.37	96.79
0.10	0.20	0.70	75.00	85.68	88.89	92.74	94.02	94.23	94.44	95.09	96.15	96.58

Table 5.2 Rank based results for single and multi-stream identification using test protocol G

6. Quality-based Adaptive face Recognition

The performance of most face recognition schemes including those mentioned earlier deteriorates when tested in uncontrolled conditions when compared to their performance under normal recording conditions. These effects are often the result of external factors such as extreme variations in illumination, expression, and occlusion. To deal with varying recording conditions, most existing schemes adopt normalization procedures that are applied irrespective of the recording conditions at the time of recording. Such strategies are known to improve accuracy in adverse conditions at the expense of deteriorated performance in somewhat normal recording conditions that generate well/reasonably lit images, and thereby yielding little or no improved overall accuracy. This section focuses on

the development of adaptive approaches to deal with such variations, whereby the application of normalisation procedures will be based on certain criteria on image quality that are detected automatically at the time of recording. We shall describe some quantitative quality measures that have been incorporated in adaptive face recognition systems in the presence of extreme variation in illumination. We shall present experimental results in support of using these measures to control the application of light normalisation procedures as well as dynamic fusion of multi-stream wavelet face recognition whereby the fusion weighting become dependent on quality measures.

6.1. QUALITY ASESSMENT MEASURES

Quality measures play an important role in improving the performance of biometric systems. There has been increasing interest by researchers in using quality information to make more robust and reliable recognition systems (e.g. [36], [37], [38], [39], [40]). Quality measures can be classified as modality-dependent and modality-independent. Modality dependent measures (such as pose or expression) can be used for face biometric only, while modality-independent measures such as (contrast and sharpness) can be used for any modality because they do not need any knowledge about the specific biometric. For multi-modal and multi-streams biometrics, there is a need to combine the various trait/stream quality measures to build adaptive weighting associated with the matching scores produced by their individual matchers, ([41], [42]). Quality measures can also be classified in terms of the availability of reference information: full reference, reduced reference, and no reference quality assessment approaches, ([43]).

Face image quality measures must reflect some or all aspects variation from a "norm" in terms of lighting, expression, pose, contrast, eyes location, mouth location, ears location, blur and so forth, ([44], [45]). New quality measures based on wavelets have been developed for different biometrics, [46]. Here, we will focus on image quality measures as a result of variation in lighting conditions and its use for improving performance of face recognition and dynamic fusion of multi-streams.

UNIVERSAL IMAGE QUALITY INDEX

Illumination image quality measures must either reflect luminance distortion of any image in comparison to a known reference image, or regional variation within the image itself. The *universal image quality index* (Q) proposed by Wand and Bovik,[47] incorporates a number of image quality components from which one can extract the necessary ingredients an illumination image quality measure that fits the above requirements. For two signals/vectors $X=\{x_i \mid i=1,2,.....N\}$ *and* $Y=\{y_i \mid i=1,2,...N\}$, Q(X,Y) is defined as:

$$Q(X,Y) = \frac{4\,\sigma_{xy}\;\bar{X}\;\bar{Y}}{(\sigma^2 x+\sigma^2 y)[(\bar{x})^2+(\bar{y})^2]} \tag{1}$$

where,

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} X_i \quad , \quad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \, , \quad \sigma_x{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \, ,$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

It models any distortion as a combination of three components: loss of correlation, luminance distortion, and contrast distortion. In fact, Q is the product of three quality measures reflecting these three components:

$$Q(X,Y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \bullet . \frac{2\,\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \bullet \frac{2\,\sigma_x \sigma_y}{\sigma^2{}_x + \sigma^2{}_y} . \tag{2}$$

The luminance quality index is defined as the distortion component:

$$\text{LQI} = \frac{2\,\bar{x}\,\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \tag{3}$$

In practice, the LQI of an image with respect to another reference image is calculated for each window of size 8x8 pixels in the two images, and the average of the calculated values defines the LQI of the entire image. The LQI is also referred to as the Global LQI as opposed to regional LQI, when the image is divided into regions and the LQI is calculated for each region separately, [48].

The distribution of LQI values for the images in the different subsets of the Extended Yale B database reveal an interesting, though not surprising, pattern. There is a clear separation between the images in sets 1 and 2, where all images have LQI values > 0.84, and those in sets 4 and 5 where all LQI vales < 0.78. Images in set 3 of the database have LQI values in the range 0.5 to 0.95.

The use of LQI with a fixed reference image that has a perceptually good illumination quality investigated as a pre-processing procedure prior to single-stream and multi-streams wavelet-based face recognition schemes, for adaptive face recognition schemes with improved performance over the non-adaptive schemes.

In the case of multi-streams schemes, a regional version of LQI index is used to adapt the fusion weights, [48]. A. Aboud et al, [37], have further developed this approach and designed adaptive illumination normalization without a reference image. We shall now discuss these approaches in more details and present experimental evidences on their success.

In order to test the performance of the developed adaptive schemes, the relevant experiments were conducted on the Extended Yale database, [49], which incorporates extreme variations in illumination recording condition. The cropped frontal face images of the extended Yale B database provide a perfect testing platform and framework for illumination based image quality analysis and for testing the viability of adaptive face recognition scheme. The database includes 38 subjects each having 64 images, in frontal pose, captured under different illumination conditions. In total number there are 2414 images. The images in the database are divided into five subsets according to the direction of the light-source from the camera axis as shown in Table 6.1.

Subsets	Angles	Image Numbers
1	$\theta < 12$	263
2	$20 < \theta < 25$	456
3	$35 < \theta < 50$	455
4	$60 < \theta < 77$	526
5	$85 < \theta < 130$	714

Table 6.1 Different illumination sets in the extended Yale B database

Samples of images for the same subject taken from different subsets of the Extended Yale B database are shown in Figure 6.1. LQI values are respectively 1, 0.9838, 0.8090, 0.4306, and 0.2213.

(a) Subset 1 (b) Subset 2 (c) Subset 3

(d) Subset 4 (e) Subset 5

Fig. 6.1. Sample images from different subsets in the Extended Yale B.

6.2. LQI–based Adaptive illumination normalization for face recognition.

Histogram Equalisation (HE) has been used as a mean to improve face recognition when the sample image suffers from poor illumination. In extreme cases when the presented sample is poorly illuminated HE improves the chance of recognition, but there are side effects and there are evidences that HE does reduce image quality and recognition accuracy in the cases of well lit images. An analysis of the effect of HE on the recognition accuracy of the various single-subband wavelet face recognition schemes for the different subsets of images in the Extended Yale B database has confirmed these conclusions, ([36], [50]). For the three level 2 wavelet subbands (LL2, LH2, and HL2), applying HE yields a reasonable-to-significant improvement in accuracy for sets 4 and 5; while the accuracy dropped as a result of HE application for sets 1, 2 and 3. What is also interesting, in this regard, is that as a result of the application of HE the values of LQI improved significantly for images in sets 4 and 5 but to a much lesser extent in set 3, while the LQI values in sets 1 and 2 has deteriorated greatly. The LQI of all images in sets 4 and 5 after HE became > 0.78.

These observation and the remarks, at the end of the section 6.1 provide the perfect threshold adopted by Sellahewa et al, [36], for the first Image quality based adaptive

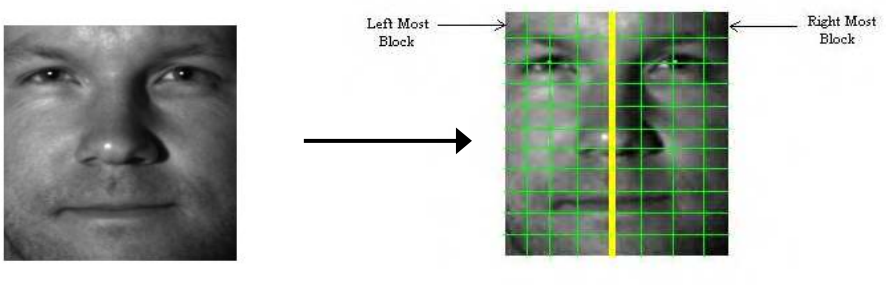
illumination normalisation procedure and the adaptive face recognition. The use of the threshold of 0.8 for LQI below which HE is applied, has led to improved face recognition in the different single subband streams as well as in the multi-stream cases. The improvement was across all subsets but to varying degrees and more significantly in sets 4 and 5, (for more details see [36]). The identification error rates for some multi-stream wavelet schemes will be presented and discussed in the last subsection. AHE refers to this LQI-based adaptive use of HE.

6.3 No-Reference Quality Index

The choice of reference image for image quality assessment is a factor that may affect the adaptive normalisation procedures and it may not be a simple task. Defining image quality measures without a reference image is a desirable task and more so for illumination assessment in relation to face images. The frontal pose of a human face is more or less symmetric; hence it is easy to design a symmetric version of the LQI without the need for a reference image. A without a reference luminance quality measure, can be defined in a two steps process, where in the first step the LQI for the left half of a face image is measured with reference to its right half and the second step uses a form of histogram partitioning that aims to measure some aspects of distortion from normal face histogram.

Step 1. The symmetric adaptive local quality index (SALQI). For a face image (I), SALQI is defined as follows:

1. Divide I into left and right half sub-images, I_L and I_R respectively, and let I_{FR} be the horizontal flipping of I_R.
2. Starting from the top left corner, use equation (3), above, to compute LQI of the 8x8 windows in I_{FR} with respect to the corresponding windows in I_L, as indicated below



3. After calculating the quality map $\{m_i = LQI_i: i=1,\ldots,N\}$, a pooling strategy as indicated in equations (4) and (5) to calculate the final quality-score of the image (I) as a weighted average of the m_i's:

$$Q = \frac{\sum_{i=1}^{N} m_i * w_i}{\sum_{i=1}^{N} w_i} \tag{4}$$

$$\text{where, } w_i = g(x_i, y_i) \text{ , and } \quad g(x,y) = \sigma_x^{\,2} + \sigma_y^{\,2} + C \tag{5}$$

Here, $x_i = \mathrm{I}_{\mathrm{L,i}}$ and $\mathrm{y_i} = \mathrm{I}_{\mathrm{FR,i}}$, where $\mathrm{I}_{\mathrm{FR,i}}$ is the mirrored block of $\mathrm{I}_{\mathrm{L,i}}$ of a row. The C is a constant representing a baseline minimal weight. The value range of SALQI is [0, 1] and its equals 1 if and only if the image is perfectly symmetrically illuminated.

Step 2. The Middle Half index (MH). The SALQI provides an indication of how symmetrical the light is distributed, but it does not distinguish between a well-lit face images from an evenly dark image. SALQI produces high quality scores for such images. To overcome this problem we use histogram partitioning**:** A good quality image normally has a dynamic range covering the full grey scale and its histogram covers well the middle part. The MH index is thus defined as:

$$MH = \frac{Middle}{Bright + Dark} \tag{6}$$

Where, **Middle =** No. of pixels in the middle range between a Lower bound LB and an Upper bound UB,

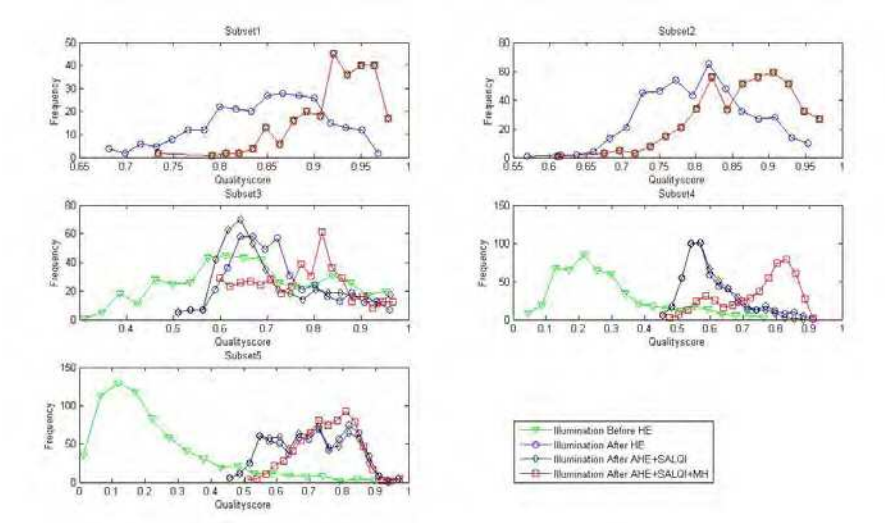
Bright = No. of pixels in the bright region of the histogram greater than UB,

Dark = No. of pixels in the dark region of the histogram less than LB,

Examining a number of so-called normal images, the LB and UB are set at 63 and 243, respectively. The MH value ranges from 0 to Max = (M/2), where M is the size of the image. The larger MH is, the better the quality is. Its maximum value depends on the image dataset.

6.4 The Symmetric Adaptive Histogram Equalisation (SAHE)

This is another adaptive scheme that uses both the SALQI and MH values to control the use of HE. Chart 6.1 displays the distribution of image LQI, SALQI and MH indices in the various subsets of the extended Yale B data base before and after the application of HE, AHE, SALQI version of AHE, and the full SAHE. For subset 1 and subset 2, we see that the application of HE results in deterioration in quality, and both AHE and MH maintain the same original quality. This confirms that for well lit images that exhibit similar illumination characteristics to those in subsets 1 and 2 (i.e. SALQI > 0.65) no normalisation is needed. The other 3 subsets benefit, to varying degrees, from pre-processing. But they benefit more from the full version of SAHE which includes the use of MH.



Charts 6.1. Distribution of for extended Yale B database before and after various normalisation.

A. Aboud et al in [37] have tested the performance of an SAHE-adaptive wavelet-based face recognition scheme in comparison with the corresponding versions with no normalization, and with the LQI-based adaptive which only used a single threshold (approx. 0.8). In the corresponding experiments, two different wavelets are used: Daubechie-1 (i.e Haar), and Daubechie-2 (also known as Daub 4), at three decomposition levels. Again the testing was based on the Extended Yale B database. The dataset are divided into two groups: training set and testing set. The training set has (38) images, one image per subject which is chosen to be (P00A+000E+00). The testing set consists of all the remaining (2394) images, i.e. 63 images per subject. Different values of SALQI and MH quality indices have been used as thresholds for SAHE approach. Recognition results, displayed in Figure 6.2, show that the LH2 subband gives the best results under varying illumination and the error rate for the SAHE with SALQI <0.6, is about 0.30% less than what was achieved by the LQI–based AHE application. However, SAHE resulted in slightly increased error rates for subset 3 images while reduced the errors of subset 4 and subset 5. The results for LL2 features are significantly better, although these error rates are much higher than the errors with LH2.

1. Calculate the quality scores for the image (I) using (SALQI) and (MH)
2. If (SALQI < Thershold1) and (MH < Threshold 2) Then
IF (MH < Thershold3) Then {Apply normalization algorithm on the whole image (I)}
Else if (MH >= Thershold3) Then
a. Apply HE on the left region of image (I) and compute SALQI
b. Apply HE on the right region of image (I) and compute SALQI
c. Apply HE on left and right regions of the image (I) and compute SALQI
Select the case that has higher SALQI value
End if
3. Else if (SALQI >= Thershold1) and (MH >= Thershold2) Then
{Do not apply histogram normalization algorithm on image (I)}
4. End if

Fig. 6.4 Symmetrical Adaptive Histogram Equalization Algorithm

No pre-process	8.89	18.20	83.30	95.82	97.20	70.71
HE, ZN	3.11	25.88	70.99	90.11	85.57	64.52
AHE, LQI < 0.80	2.67	22.81	69.01	90.11	84.03	63.05
SAHE, SALQI < 0.60	2.67	7.89	37.8	73.76	76.61	48.36
SAHE, SALQI < 0.70	2.67	7.89	38.02	73.76	76.47	48.36
SAHE, SALQI < 0.80	2.67	20.83	40	73.76	76.47	51.22
SAHE, SALQI < 0.90	2.67	7.89	38.24	75.1	76.05	48.57

(a) Wavelet Haar, suband: LL2

No pre-process	8.00	0.00	30.55	71.10	95.24	50.97
HE, ZN	7.56	0.44	17.58	26.62	14.15	14.31
AHE, LQI < 0.80	7.11	0	11.65	20.34	11.76	10.94
SAHE, SALQI < 0.60	7.11	0	12.97	18.25	11.34	10.61
SAHE, SALQI < 0.70	7.11	0	12.97	18.63	11.48	10.73
SAHE, SALQI < 0.80	7.11	0	12.53	18.44	12.32	10.86
SAHE, SALQI < 0.90	7.11	0	12.75	18.63	11.34	10.65

(b) Wavelet Haar, suband: LH2

	Set1	Set2	Set3	Set4	Set5	All
No pre-process	8.44	14.25	80.66	95.63	97.20	69.36
HE, ZN	1.78	20.83	67.47	90.30	85.71	62.84
AHE, LQI < 0.80	0.89	17.54	64.84	90.87	84.45	61.36
SAHE, SALQI < 0.60	0.89	4.61	30.99	72.05	77.03	46
SAHE, SALQI < 0.70	0.89	4.61	31.21	71.86	76.89	45.96
SAHE, SALQI < 0.80	0.89	15.79	33.19	72.05	77.03	48.57
SAHE, SALQI < 0.90	0.89	4.61	31.43	73.38	76.47	46.21

(c) Wavelet Daub 4, suband: LL2

	Set1	Set2	Set3	Set4	Set5	All
No pre-process	14.67	0	35.60	66.35	89.64	49.83
HE, ZN	13.33	0	24.84	28.33	18.35	17.80
AHE, LQI < 0.80	13.33	0	21.76	22.24	16.11	15.19
SAHE, SALQI < 0.60	13.33	0	20.22	21.48	15.83	14.65
SAHE, SALQI < 0.70	13.33	0	20.22	21.48	15.83	14.65
SAHE, SALQI < 0.80	13.33	0	20.66	21.48	16.39	14.90
SAHE, SALQI < 0.90	13.33	0	20.22	21.29	15.69	14.56

(d) Wavelet Daub4, suband: LH2

Table 6.2 Identification error rates of wavelet-based face recognition system

6.5 Regional LQI and Adaptive fusion of multi stream face recognition

The previous parts of this section demonstrated the suitability of using the AHE and SAHE as a mean of controlling the application of illumination normalisation procedure (HE) and the benefits that this yields for single and multi-stream face recognition schemes. However, in real-life scenarios, variations in illumination between enrolled and test images could be confined to a region, rather than the whole, of the face image due to the changes in the direction of the light source or pose. Therefore, it is sensible to measure the illumination quality on a region-by-region basis. Sellahewa et al, [48], has experimented with a rather simple regional modification of the LQI, whereby we split the image into 2x2 regions of equal size, and tested the performance of the Regional AHE based adaptive multi-stream face recognition. Figure 6.5 and Figure 6.6 present that Identification error rate for the RLQI-based fusion of (LL2, LH2) and (LH2, HL2), respectively, using 10 different weighting configurations.

LL2 + LH2		Identification Error Rates (%)					
WLL	WLH	Set 1	Set 2	Set 3	Set 4	Set 5	Total
1.0	**0.0**	3.56	17.54	38.24	73.57	75.91	50.13
0.9	**0.1**	2.22	7.68	30.11	65.78	70.73	43.27
0.8	**0.2**	2.22	3.29	22.64	57.03	63.31	36.83
0.7	**0.3**	1.33	0.44	18.46	46.2	51.26	29.38
0.6	**0.4**	2.22	0.00	15.16	36.12	38.94	22.81
0.5	**0,5**	3.56	0.00	14.73	27.95	27.17	17.51
0.4	**0.6**	4.89	0.00	13.41	19.96	18.91	13.13
0.3	**0.7**	5.78	0.00	12.97	17.87	14.57	11.36
0.2	**0.8**	8.80	0.00	14.07	15.59	11.62	10.4
0.1	**0.9**	8.89	0.00	13.85	13.5	10.36	9.6
0.0	**1.0**	10.67	0.00	14.95	14.45	10.36	10.19

Fig. 6.5 Non-Adaptive Fusion (LL, LH)

LH2 + HL2		Identification Error Rates (%)					
WLH	WHL	Set 1	Set 2	Set 3	Set 4	Set 5	Total
1.0	**0.0**	10.67	0.00	14.95	14.45	10.36	10.19
0.9	**0.1**	8.44	0.00	14.73	12.55	9.52	9.26
0.8	**0.2**	7.56	0.00	12.75	12.74	9.38	8.8
0.7	**0.3**	5.78	0.00	11.65	11.41	9.8	8.25
0.6	**0.4**	5.33	0.00	9.23	10.65	11.76	8.16
0.5	**0,5**	4.44	0.00	7.47	10.65	13.17	8.16
0.4	**0.6**	3.11	0.00	7.47	11.98	18.49	9.93
0.3	**0.7**	2.22	0.00	8.13	15.78	24.37	12.58
0.2	**0.8**	5.33	0.00	9.23	19.96	36.97	17.8
0.1	**0.9**	7.11	0.00	12.75	29.09	51.68	25.08
0.0	**1.0**	9.33	0.88	15.16	39.35	61.62	31.19

Fig. 6.6 Non-Adaptive Fusion (LH, HL)

The use of RLQI has obviously resulted in further improvement in accuracy of multi-stream recognition schemes. With best overall error rate of 9.6 for the (LL2, LH2) fused scheme achieved when LL2" was given a small weight of 0.1, while best error rate for the (LH2, HL2) fused scheme is 8.16 achieved when have nearly equal weights. What is more interesting is that the best performance over the different sets is achieved with different weighting configurations in both cases. This shows that the wavelet-based multi-stream recognition scheme, developed previously, has no objective means of selecting fusion parameters and that it performed differently for face images captured with different lighting conditions has led to developing of a new adaptive approach to face recognition. This suggests a dynamic scheme of weighting that depends on image quality. Figure 6.7, below, presents the results obtained for using quality-based adaptive fusion of two or 3 subbands. In this case if the LQI of the image is>0.9 then the score for LL2 will be given a 0.7 weighting otherwise it is given a 0 weighting. The LH2 and HL2 subbands get equal proportion from the left over.

Feature	Subband	Identification Error rate %					
		Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	All
subband	LL2	3.56	17.54	38.24	73.57	75.91	50.13
	LH2	10.67	0.00	14.95	14.45	10.36	10.19
	HL2	9.33	0.88	15.16	39.35	61.62	31.19
Adaptive Furion	LL2+LH2	2.22	0.00	14. 73	14.45	10.36	9.34
	LL2+LH2+HL2	7.47	0.22	1.78	10.65	13.17	7.95

Fig. 6.7 Adaptive Fusion

It is clear that, this dynamic choice of weighting of the scores has led to further improvement over the non-adaptive static selection of weighting.

7. CONCLUSIONS AND FUTURE WORK

In this chapter we have reviewed face recognition schemes, and in particular we advocated the use of wavelet-based face recognition. The fact that a wavelet-transform of face image into a number of different subbands representing the face at different frequency range and different scales, has been exploited to develope several single-stream face recognition schemes one for each wavelet subband. The performances of several of these were tested over a number of benchmark databases, which revealed different error rates, but achieving comparable/better results compared to PCA based schemes. This approach has the advantage of being very efficient and being scalable.

We have also shown that one mimicked the success of fusion approach to multi-modal biometric-based recognition by using multi-stream face recognition that is based on fusing a number of single streams. Even the fusion of a small (<4) number of single streams has led to significant improvement in performance.

Finally, we have demonstrated with a significant degree of success that the challenge of face recognition in the presence of extreme variation illumination can be dealt with using adaptive quality –based face recognition. The main advantages of using quality measures are the avoidance of excessive unnecessary enhancement procedures that may cause undesired artefacts, reduced computational complexity which is essential for real time applications, and improved performance.
The work on quality- based adaptive fusion and adaptive wavelet multi-stream wavelet face recognition will be expanded in the future to deal with other quality issues as well as efficiency challenges.

8. REFERENCES

1. W. Zhao, R. Chellappa, A. Rosefeld, and P. J. Phillips, "Face Recognition: A Literature Survey," Technical report, Computer Vision Lab, University of Maryland, 2000.
2. D. M. Etter. "The Link Between National Security and Biometrics", Proc. of SPIE Vol 5779, Biometric Technology for Human Identification II, pp 1-6, March 2005
3. T. Sim, R. Sukthankar, M. D. Mullin, and S. Baluja. "High-Performance Memory-based Face Recognition for Visitor Identification," *ICCV-99, Paper No. 374.*
4. M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 103–108, 1990.
5. M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience vol. 3, no. 1, pp. 71–86, 1991.
6. B Moghaddam, W Wahid and A pentland, Beyond eigenfaces: Probabilistic matching for face recognition, Proc. of face and gesture recognition, pp. 30 –35, 1998.
7. J. Yi, J. Kim, J. Choi, J. Han, and E. Lee, "Face Recognition Based on ICA Combined with FLD," in Biometric Authentication, M. Tistarelli and J. B. Eds, eds., Proc. Int'l ECCV Workshop, pp. 10–18, June 2002.
8. G. L. Marcialis and F. Roli, "Fushion of LDA and PCA for Face Verification," in Biometric Authentication, M. Tistarelli and J. B. Eds, eds., Proc. Int'l ECCV Workshop, pp. 30–37, June 2002.
9. R. Cappelli, D. Maio, and D. Maltoni, "Subspace Classification for Face Recognition," in Biometric Authentication, M. Tistarelli and J. B. Eds, eds., Proc. Int'l ECCV Workshop, pp. 133–141, June 2002.
10. J.-T. Chien and C.-C. Wu, "Discriminant Wavelet faces and Nearest Feature Classifiers for Face Recognition," in IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 12, pp. 1644–1649, December 2002.
11. M. Bicego, E. Grosso, and M. Tistarelli, "Probabilistic face authentication using Hidden Markov Models" in Biometric Technology for Human Identification II, Proc. SPIE vol. 5779, pp. 299–306, March 2005.
12. J.G. Daugman, "Two-Dimensional Spectral Analysis of Cortical Receptive Field Profile," Vision Research, vol. 20, pp. 847-856, 1980.
13. J.G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 36, no. 7, pp. 1,169-1,179, 1988.

14. M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Würtz, and W. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," IEEE Trans. Computers, vol. 42, no. 3, pp. 300–311, 1993.
15. L. Wiskott, J-M Fellous, N. Krüger, and C. Malsburg, "Face Recognition by Elastic Bunch Graph Matching", IEEE Tran. On Pattern Anal. and Mach. Intell., Vol. 19, No. 7, pp. 775-779, 1997.
16. Z. Zhang, M. Lyons, M. Schuster, and_ S. Akamatsu, "Comparison Between Geometry-Based and Gabor-Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron", Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition, Nara Japan, IEEE Computer Society, pp. 454-459 (1998).
17. Chengjun Liu and Harry Wechsler, "Independent Component Analysis of Gabor features for Face Recognition", IEEE Trans. Neural Networks, vol. 14, no. 4, pp. 919-928, 2003.
18. I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," IEEE Trans. Information Theory, vol. 36, no. 5, pp. 961-1004, 1990.
19. K. Etemad and R. Chellappa, "Face Recognition Using Discriminant Eigenvectors," Proc. IEEE Int'l Conf. Acoustic, Speech, and Signal Processing, pp. 2148-2151, 1996.
20. Dao-Qing Dai, and P. C. Yuen. "Wavelet-Based 2-Parameter Regularized Discriminant Analysis for Face Recognition," Proc. AVBPA Int'l Conf. Audio-and Video-Based Biometric Person Authentication, pp. 137-144, June, 2003.
21. D. Xi, and Seong-Whan Lee. "Face Detection and Facial Component Extraction by Wavelet Decomposition and Support Vector Machines," Proc. AVBPA Int'l Conf. Audio-and Video-Based Biometric Person Authentication, pp. 199-207, June, 2003.
22. F. Smeraldi. "A Nonparametric Approach to Face Detection Using Ranklets," Proc. AVBPA Int'l Conf. Audio-and Video-Based Biometric Person Authentication, pp. 351-359, June, 2003.
23. J.H. Lai, P. C. Yuen!, G. C. Feng," Face recognition using holistic Fourier invariant features", Pattern Recognition 34, pp. 95-109, (2001)
24. N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals", J.of Appl. And Comp. Harmonic Analysis, 01 (3), pp. 234-253, May 2001.
25. Y. Peng, X. Xie, W. Xu, and Q. Dai, "Face Recognition Using Anistropic Dual-Tree Complex Wavelet Packets", Proc IEEE Inter. Conf. on Pattern Recognition, 2008.
26. Sidney Burrus Ramesh, C, A Gopinath, and Haittao Guo. Introduction to Wavelet and Wavelet Transforms A Primer. Prentice Hall. Inc., 1998.
27. Naseer AL-Jawad, "Exploiting Statistical Properties of Wavelet Coefficients for Image/Video Processing and Analysis", DPhil Thesis, University of Buckingham, 2009.
28. D. Xi, and Seong-Whan Lee. "Face Detection and Facial Component Extraction by Wavelet Decomposition and Support Vector Machines," Proc. AVBPA Int'l Conf. Audio-and Video-Based Biometric Person Authentication, pp. 199-207, June, 2003.
29. A. Z. Kouzani, F. He, and K. Sammut. "Wavelet Packet Face Representation and Recognition," Proc IEEE Conf. Systems, Man, and Cybernetics, pp. 1614-1619, 1997.
30. Dao-Qing Dai, and P. C. Yuen. "Wavelet-Based 2-Parameter Regularized Discriminant Analysis for Face Recognition," ProcComputer vol. 33, no. 2, pp. 50–55, February 2000.
31. H. Sellahewa,"Wavelet–based Automatic Face Recognition for Constrained Devices", Ph.D. Thesis, University Of Buckingham, (2006).

32. E. Bailly-Bailli´ere, S. Bagnio, F. Bimbot, M. Hamouz, J. Kittler, J. Mari´ethoz, J. Matas, K. Messer, V. Popovici, F. Por´ee, B. Ruiz, and J. Thiran, "The BANCA Database Evaluation Protocol," in Audio-and Video-Based Biometric Person Authentication, Proc. AVBPA Int'l Conf, pp. 625–638, June 2003.
33. G C Feng a b, P C Yuen b and D Q Dai, "Human Face Recognition Using PCA on Wavelet Subband", SPIE Journal of Electronic Imaging 9 (2), pp: 226–233, (2000).
34. M-H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel methods" in Automatic Face and Gesture Recognition, Proc. IEEE Int'l Conf, pp. 215–220, 2002.
35. P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible lighting conditions?," in Proc. IEEE Conf
36. Harin Sellahewa, Sahah A. Jassim, "Illumination and Expression Invariant Face Recognition: Toward Sample Quality –based Adaptive Fusion", in Biometrics: Theory, Applications and Systems, Proc. Second IEEE Int'l Conference 10414693, 1-6 (2008).
37. A. Aboud, H.Sellahewa, and S. A. Jassim, "Quality Based Approach for Adaptive Face Recognition", in Proc. SPIE, Vol. 7351, Mobile Multimedia/image processing, Security, and Applications, Florid, April 2009.
38. Li Ma, Tieniu Tan," Personal Identification Based on Iris Texture Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(12), 20-25 (2003).
39. Oriana Yuridia Gonzalez Castillo, "Survey about Facial Image Quality", Fraunhofer Institute for Computer Graphics Research, 10-15 (2005).
40. Yi Chen, Sarat C. Dass, and Anil K. Jain, " Localized Iris Image Quality using 2-D Wavelets ", Proc. Advances in Biometrics Inter. Conf., ICB 2006, 5-7 (2006).
41. Sarat C. Dass, Anil K. Jain, "Quality-based Score Level Fusion in Multibiometric Systems", Proc. 18th Inter. Conf. on Pattern Recognition 4(1), 473 - 476 (2006).
42. Jonas Richiardi, Krzysztof Kryszczuk, Andrzej Drygajlo, "Quality Measures in Unimodal and Multimodal Biometric Verification ", Proc. 15th European Conference on Signal Processing EUSIPCO,2007.
43. Hanghang Tong, Mingjing Li, Hong-Jiang Zhang, Changshui Zhang, Jingrui He. "Learning No-Reference Quality Metric by Examples", Proc. The 11th International Multi-Media Modelling Conferene 1550, 247- 254 (2005)
44. Robert Yen, "A New Approach for Measuring Facial Image Quality", Biometric Quality Workshop II, Online Proc. National Institute of Standards and Technology, 7-8, (2007).
45. Krzysztof Kryszczuk, Andrzej, "Gradient–based Image Segmentation for Face Recognition Robust to Directional Illumination", Proc. SPIE, 2005.
46. Azeddine Beghdadi, "A New Image Distortion Measure Based on Wavelet Decomposition", Proceeding of IEEE ISSPA2003, 1-4 (2003).
47. Alan C. Bovik, Zhou Wang",A Universal Image Quality Index", IEEE Signal Processing Letters. 9(.3), 81-84 (2002).
48. Harin Sellahewa, Sahah A. Jassim, "Image Quality-based Face Recognition" To appear in IEEE Trans. On Instrumentation And Measurements, Biometrics, 2009.
49. Georghiades, A.S. and Belhumeur, P.N. and Kriegman, D. J, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose", IEEE Transactions on Pattern Analysis and Machine Intelligence"23(6), 643-660 (2001).
50. Harin Sellahewa, Sahah A. Jassim, "Image Quality-based Adaptive Illumination Normalisation for Face Recognition" in Proc. SPIE Conf. on Biometric Technology for Human Identification, Florid, April 2009.

Face Recognition

Edited by Milos Oravec

ISBN 978-953-307-060-5
Hard cover, 404 pages
Publisher InTech
Published online 01, April, 2010
Published in print edition April, 2010

This book aims to bring together selected recent advances, applications and original results in the area of biometric face recognition. They can be useful for researchers, engineers, graduate and postgraduate students, experts in this area and hopefully also for people interested generally in computer science, security, machine learning and artificial intelligence. Various methods, approaches and algorithms for recognition of human faces are used by authors of the chapters of this book, e.g. PCA, LDA, artificial neural networks, wavelets, curvelets, kernel methods, Gabor filters, active appearance models, 2D and 3D representations, optical correlation, hidden Markov models and others. Also a broad range of problems is covered: feature extraction and dimensionality reduction (chapters 1-4), 2D face recognition from the point of view of full system proposal (chapters 5-10), illumination and pose problems (chapters 11-13), eye movement (chapter 14), 3D face recognition (chapters 15-19) and hardware issues (chapters 19-20).

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Sabah A. Jassim (2010). Wavelet–Based Face Recognition Schemes, Face Recognition, Milos Oravec (Ed.), ISBN: 978-953-307-060-5, InTech, Available from: http://www.intechopen.com/books/face-recognition/wavelet-based-face-recognition-schemes

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821