CHAPTER 2

LITERATURE SURVEY

2.1 STUDIES ON PROMINENT BIOMETRIC COMPONENTS

The basic physical biometrics includes fingerprints, hand or palm geometry, and retina, iris, or facial characteristics. The Behavioral characteristics include signature, voice (which also has a physical component), keystroke pattern, and gait. Although some technologies have proven to be a reliable method, it is beyond doubt that the field of access control and biometrics as a whole shows great potential for use in end user segments, such as airports, stadiums, defense installations, and the industry and corporate workplaces where security and privacy are required. Recognition of individuals based on physiological characteristics, such as the fingerprint, the face and the iris, is a reliable research domain. Biometric algorithms allow for the recognition of individuals based on physical or logical access control systems and thus provide an efficient and convenient alternative to knowledge-based or token-based security systems.

2.1.1 Finger Print Recognition

Fingerprint recognition is used to compare an input fingerprint with only one template of the claimed identity was extensively studied and promising results were reported. Several techniques were proposed to index fingerprints using minutia points. An advantage of these approaches is that they can classify fingerprints into more classes than the exclusive classification as they exploit more discriminating features, fingerprint minutiae [19]. However, these indexing approaches construct a superfluous representation of fingerprint because minutia points are used in most automated fingerprint matching algorithms. Information contained in the feature for the indexing is therefore highly correlated with or just a subset of
that used in the subsequent finer matching. As a result, these methods can barely improve the accuracy deterioration caused by the extension of the one-to-one matching to the one-to- \( N \) matching although they can speed up the recognition process.

Advantages:

- High accuracy
- The most economical biometric PC user authentication technique
- Easy to use
- Small storage space required for the biometric template, reducing the size of the database memory required.
- It is standardized.

Disadvantages:

- For some people it is very intrusive, because is still related to criminal identification.
- It can make mistakes with the dryness or dirty of the finger’s skin, as well as with the age (is not appropriate with children, because the size of their fingerprint changes quickly).

2.1.2 Face Recognition

There are three main approaches to the problem of face recognition: \(^{[20]} \) i) based on appearance, ii) based on invariant characteristics, and iii) based on models. In the first approach the objective is to extract similar characteristics present in all faces. Normally statistical or machine learning techniques are used, and dimensionality reduction tools are very important for improving efficiency. One of the most used unsupervised tools in this respect is principal component analysis (PCA) \(^{[21]} \). This method linearly projects the high-
dimensional input space onto a lower-dimensional subspace containing all the related image information. This procedure is applied over all the face images –training set– used for the construction of the identification system. This projection space is known as *eigenfaces space*. To recognize a new face the image is transformed to the projection space, and the differences between that projection and those of the training faces are evaluated. The smallest of these differences, which in turn is smaller than a certain threshold, gives the identity of the required face.

Advantages:

- Non Intrusive
- Cheap Technology
- A person's face can be captured from a distance by a camera many casinos, stores and urban areas use facial recognition to identify potential crimes.

Fig 2.1 Sample Human Face Images with Different Posture.
Disadvantages:

- 2D recognition is affected by changes in lighting, the person’s hair, the age and if the person wears glasses.
- Requires camera equipment for user’s identification; thus it is not likely to become popular until most PCs include cameras as standard equipment.
- Facial recognition systems may require a stationary or posed user in order to capture the image, though many systems use a real-time process to detect a person's head and locate the face automatically.

### 2.1.3 Hand Geometry

The brains and opposing thumbs are the main reason for humankind to survive and evolve as suggested by anthropologists. Grasping, throwing and making tools are allowed by the versatile human hand. A media to verify identity is another use of the human hand today. The Ancient Egyptians used to classify and identify people using body measurements \(^{[22]}\). Today, precisely recording and comparing hand dimensions are done by hand geometry scanners using infrared optics and microprocessor technology. The present century has seen several hand geometry verification technologies which ranges from electro-mechanical devices to the solid state electronic scanners which has been manufactured today. Robert P. Miller \(^{[22]}\) was issued patents by the U.S. Patent office in the late 1960’s and early 1970’s for a device that measures hand characteristics, and records unique features for comparison and ID verification. “Identimation” was the name given to Miller’s machines which were highly mechanical.
Hand geometry is a biometric component where it identifies users by the shape of their hands and length of their fingers. Hand geometry readers measure a user's hand along many dimensions and compare those measurements to measurements stored in a file. Normal hand geometry devices have been manufactured since the early 1980s, making hand geometry the first biometric to find extensive computerized use.

Since hand geometry is not thought to be as unique for every human as fingerprints or irises, fingerprinting and iris recognition remain the reliable technology for high-security applications. Hand geometry is very reliable when combined with other forms of identification, such as identification cards or personal identification numbers. In large populations, hand geometry is not suitable for so-called one-to-many applications, in which a user is identified from his biometric without any other identification.

Advantages:

- Used extensively for physical-access control and time.
- Fast and Easy.
- Non intrusive.
- Suitable for outdoor installation

Disadvantages:

- Devices are bulky—not generally suitable for desktop use.
- Those with arthritis may find it difficult to use.
- Not accurate as other biometrics, but suitable for one-to-one verification.
2.1.4 Ear Recognition

The Front view of an individual’s face is exclusively employed by traditional image-based approaches to a person’s identification. More uniform distribution of color is present in the hand and so all the information is saved while original image is converted to gray scales. In case of ear the shape and appearance is fixed but when face or lip as a biometric is employed, then changing of their appearance with the expression of the subject makes problem. Ear image features are extracted by measuring various geometric relations between the fixed points which are lengths and angles. The geometrical structure observed from pixel value distances are used for the successful recognition of objects.

Identification of objects is based on these geometrical structures. Experimental result reveals that this identification method can achieve almost 89% accuracy \(^{[23]}\). The complexity of this process is less since the method does not require considering the inner boundary of
ear. Finding objects from images is one of the most essential tasks of vision systems. In addition to this, the reliability of tracking systems greatly depends on the detection of targets. A new class of biometrics based upon ear features was introduced for use in the development of passive identification systems by Alfred Iannarelli [24]. Identification by ear biometrics is promising because it is passive like face recognition, but instead of the difficulties to extract face biometrics, it uses strong and simply extracted biometrics like those in fingerprinting. The ear is a unique feature of human beings. Even the ears of “identical twins” differ in some respects. There are persons in crime laboratories that assume that the human external ear characteristics are unique to each individual and unchanging during the lifetime of an adult[25].

Over the years, suggestions have been made in the occasional literature that the shapes and characteristics of human ear are widely different and may be in fact sufficiently variant such that it is possible to differentiate between the ears of all individuals. Unfortunately, this “individuality” has apparently been taken for granted but has never been empirically established. Techniques that allow computers to understand the shape of a human ear in images and video sequences can be used in a wide range of applications. In certain domains it suffices to recognize a few different shapes, observed always from the same viewpoint.

2.1.5 Iris Recognition

Iris recognition starts with finding an iris in an image, finding its inner and outer boundaries at the pupil and sclera, detecting the upper and lower eyelid boundaries if they occlude, and detecting and excluding any superimposed eyelashes or reflections from the cornea or eyeglasses. These processes may collectively be called segmentation. Precision in
assigning the true inner and outer iris boundaries, even if they are partly invisible, is important because the mapping of the iris in a dimensionless (i.e., size invariant and pupil dilation invariant) coordinate system is critically dependent on this inaccuracy in the detection, modeling, and representation of these boundaries can cause different mappings of the iris pattern in its extracted description, and such differences could cause failures to match. It is natural to start by thinking of the iris as an annulus [26]. Soon, one discovers that the inner and outer boundaries are usually not concentric.

A simple solution is then to create a non concentric pseudo-polar coordinate system for mapping the iris, assuming that the iris and pupil share a common center and requiring only that the pupil is fully contained within the iris. The different iris recognition systems are discussed in the next section of this chapter.

Advantages:

- Verification time is generally less than 5 seconds.
- The eye from a dead person would deteriorate too fast to be useful, so no extra precautions have to been taken with retinal scans to be sure the user is a living human being.
- It is practically flat and uniform under most conditions and it has a texture that is unique even to genetically identical twins.
- Iris patterns are randomly formed, thus a person's left and right eye are different.
- The iris is very distinctive and robust, which provides quick results for both identification and verification purposes.
Disadvantages:

- Intrusive
- A lot of memory for the data to be stored, Very expensive

2.1.6 Keystroke

It is a behavioral biometric that relies on the way how a human type on a typical keyboard/keypad type device. As a person types, certain attributes are extracted and used to authenticate or identify the typist. This technology focuses on extracting quantitative information from the interactions between a user and a keypad/keyboard device and using this information to automate user authentication and/or identification. The basic idea is that each user interacts with a keyboard in a particular way, producing a unique pattern that can be associated with that particular user. A huge portion of the research in this domain has focused on identifying which aspects of our interactions with such devices are sufficiently distinctive as to provide a unique reference signature. Again, there are two principal options\[27\]: text-dependent and text-independent versions. The most common form of text-dependent systems requires users to enter their login ID and password (or commonly just their password). In the text-independent version, users are allowed to enter any text string they wish. In some implementations, a third option is used where a user is requested to enter a long text string on the order of 500–1500 characters.

Users enroll into the system by entering their text either multiple times if the short text-independent system (i.e. password) is employed, or typically once if the system employs a long text string. From this enrollment process, the user’s typing style is acquired and stored for subsequent authentication purposes. This approach is well suited for remote access
scenarios: no specialized hardware is required and users are used to provide their login credentials. Some of the attributes that are extracted when a person types are the duration of a key press[^28] (dwell time) and the time between striking successive keys (digraph if the time is recorded between successive keys). These features, along with several others, are used to build a model of the way a person types. The security enhancement provided by this technology becomes evident if you leave your password written on a sticky notepad tucked inside your desk, which someone happens to find. Without this level of protection, possession of the password is that that is required for a user to access your account. With the addition of a keystroke dynamics-based biometric, it is not sufficient that the password is acquired: the password has to be entered exactly (or at least within certain tolerance limits) the way the enrolled user entered it during enrollment. If not, the login attempt is rejected.

2.1.7 Signature

The Signature verification system makes users to present a required handwritten text for authentication. This is probably the most familiar of all biometrics – though currently not the most prevalent – due to the start of computer-based passwords. There are two essentially distinct forms of signature-based biometrics: online and off-line[^29]. With an online signature verification system, the signature characteristics are extracted as the user writes, and these features are used to immediately authenticate the user. Typically, specialized hardware is required, such as a pressure-sensitive pen (a digital pen) and/or a special writing tablet. These hardware elements are designed to capture the dynamical aspects of writing, such the pen pressure, pen angle, and related information. In a remote access approach, where specialized hardware may not be possible, then the online approach is most suitable from a small portable device such as a PDA, where the stylus can be used for writing. The off-line
approach uses the static features of the signature, such as the length and height of the text, and certain specialized features such as loops (not unlike a fingerprint approach).

Typically, the data are acquired through an image of the signature, which may be photocopied or scanned into a computer for subsequent analysis. As in all behavioral biometric approaches, a writing sample must be stored in the authentication database, and the writing sample is compared to the correct reference sample before the acceptance/rejection decision to be made. Again, there is the possibility of having text-dependent or text-independent signature verification. The same conditions that apply to voice also apply here – and voice and signature are really very similar technologies – only the mode of communication has changed, which results in a different set of features that can be extracted. Signature-based authentication systems tend to be more stable than keystroke/mouse dynamics – and so may be more similar to physiological biometrics in this respect.

Fig 2.3: A person giving signature in to the system.
Signature feature space can also be classified based on whether the feature attributes are generated in a static or dynamic fashion. Generally speaking, dynamic attributes capture temporal features, and static attributes capture the spatial features of a signature. Dynamic features of a signature provide information regarding how the signature changes as it is written, examining at discrete (or possibly continuous) time points various features of the signature. These include position (x, y coordinates), curvature, acceleration, distance between successive sampling times, and related information (see the case studies for more details). In general, these attributes are made available by the enabling technologies – digital pens and pressure-sensitive tablets provide the means to capture this data with sub-millisecond temporal resolution in many cases. The static features include the degree of slant, the bounding box (horizontal and vertical dimensions that contain the signature), and pen width.

Another classification that is relevant in this context is global versus local feature space. Global features describe how the signature is produced and includes techniques such as the discrete wavelet transform, the Hough transform, and projections. Local features are typically extracted from at the stroke level (a stroke is a portion of the signature that is produced while the writing implement is on the writing surface) and include ballistic movement, local shape features, and slant/orientation information. Aspects of this rich feature space have been exploited in order to enhance the classification efficiency of online and off-line signature verification systems, as discussed in the case studies presented below.

Signature-based identity verification is a venerable technique for presenting it in an official capacity, with a legal implication of authenticity. Prior to automated biometrics, signatures were checked by visual inspection, often by people with little or no formal training. Exactly what is it about a signature that allows us to differentiate the real from a
cleverly (or even not so cleverly) crafted forgery? One answer might be that, not really looked for cleverly crafted forgeries – those produced by skilled attempts. We simply do not have the time to examine a signature at the level required to attempt to differentiate a valid from a forged version. After all, if someone has access to somebody else’s signature and practices long enough, it will probably be able to reproduce it to a degree that is indiscernible from the authentic version. If this is the case, then humans probably rely on other cues in addition to the signature for verification purposes.

2.1.8 Voice recognition

The deployment of speech as a biometric from a remote access approach is principally based on speech recognition mapping between the user’s request for authentication and their speech pattern. Generally speaking, the user will be required to enunciate a fixed speech pattern, which could be a password or a fixed authentication string, consisting of a sequence of words spoken in a particular order\[34\]. This paradigm is typically referred to as a text-dependent approach, in contrast to a text-independent approach, where the speaker is allowed to enunciate a random speech pattern. Text-dependent speaker verification is generally considered more appropriate – and more effective in a remote access approach, as the amount of data available for authentication is at a premium. This is a reflection of the potentially unbounded number of potential users of the system as the number of user increases, the computational complexity necessarily rises.

The data generated from voice signals are captured by a microphone attached to a digital device, typically a computer of some sort (though this includes mobile phones and PDAs). The signals generated by speech are analogue signals, which must be digitized at a certain frequency. Typically, most moderate grade microphones employed have a sampling
rate of approximately 32 kHz\cite{37}. The typical dynamic range of the human vocal cord is on
the order of 8 kHz, though the absolute dynamic range is approximately 1–20 kHz). The
Nyquist sampling theorem states that a signal must be sampled at least twice per cycle, so a
32 kHz sampling rate is generally more than sufficient for human voice patterns. If the signal
is sampled less than the Nyquist sampling rate, then aliasing will occur, which will corrupt
the frequency aspect of the signal (see Figure 2.1 for an example). In addition to capturing
the frequency aspect of voice signals, the amplitude of the signal must be faithfully captured,
otherwise the pitch (reflected in the amplitude) will be truncated, resulting in information
loss at higher frequencies.

Therefore, for reliable signal acquisition of voice data, the frequency and amplitude
of the signal must be acquired with high fidelity. This is an issue with speakers with a large
high frequency component, such as women and children. Typically, most modern recording
devices are capable of digitizing voice data at 16 bits or more, providing more than sufficient
dynamic range to cover human speech patterns. In addition, if the data are to be collected
over a telephone type device, the data are truncated into a small dynamic range of typically 4
kHz.

It compares live speech with a previously created speech model in which users are
requested to enunciate text as a means of identifying themselves. Voice can be employed for
either speaker identification or speaker authentication. With respect to speaker identification,
a person enunciates text, and the speech patterns are analyzed to determine the identity of the
speaker. In the literature, this is referred to as speaker-independent recognition. This mode
poses several interesting issues, such as what happens if the speaker is not contained within
the database of speakers? As in all major forms of biometrics, any individual wishing to
utilize the biometric device must, at some stage, introduce themselves to the system, typically in the form of an enrollment process. One of the principal tasks of the enrollment process is to register the person as a potential user of the biometric system. In a speaker-independent system, the user’s voice pattern is analyzed and compared to all other voice samples in the user database. There are a number of ways this comparison is made.

The closest match to the particular voice data presented for identification becomes the presumed identity of the speaker. There are three possible outcomes: i) The speaker is correctly identified; ii) the speaker is incorrectly identified as another speaker; or iii) the speaker is not identified as being a member of the system. When speakers attempt an authentication task, the speakers have provided some evidence of their identity, and the purpose of the voice recognition process is to verify that these persons have a legitimate claim to that identity. The result of this approach is a binary decision: either the claimed identity is verified or not.

Advantages:
- Socially acceptable and nonintrusive.
- Standard components and audio channels can be used.
- Language independent.
- Interoperable with passphrases or challenge/response mechanisms.

Disadvantages:
- Background noise can interfere with capture.
- Illness and stress can impede effectiveness.
- Relatively long enrollment times.
- Large data record generated.
2.2 BIOMETRIC RECOGNITION FOR IRIS

Iris recognition produces the correct result by extracting features of the input images and matching these features with known patterns in the feature database. Such a process can be divided into four main stages: image acquisition, preprocessing, feature extraction, and matching. The first stage solves the problem of how to choose a clear and well-focused iris image from an image sequence for recognition. Preprocessing provides an effective iris region in a selected image for subsequent feature extraction and matching. The extraction of feature from an iris image and matching can be done using several algorithms, each with its own pros and cons. This chapter discusses some renowned algorithms further.

2.2.1 Pattern Recognition Algorithms

**Discriminant Analysis Algorithms:**

**LDA** – Linear discriminant analysis (LDA) and the related Fisher's linear discriminant[^39] are methods used in statistics, pattern recognition and machine learning to find a linear combination of features which characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification.

LDA is closely related to ANOVA[^40] (analysis of variance) and regression analysis, which also attempt to express one dependent variable as a linear combination of other features or measurements. In the other two methods however, the dependent variable is a numerical quantity, while for LDA it is a categorical variable (i.e. the class label). Logistic regression and prohibit regression are more similar to LDA, as they also explain a categorical variable. These other methods are preferable in applications where it is not reasonable to

[^39]: Refer to the source for details.
[^40]: Refer to the source for details.
assume that the independent variables are normally distributed, which is a fundamental assumption of the LDA method. LDA is also closely related to principal component analysis (PCA)\[41\] and factor analysis in that they both look for linear combinations of variables which best explain the data.

LDA explicitly attempts to model the difference between the classes of data. PCA on the other hand does not take into account for any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made.

LDA works when the measurements made on independent variables for each observation are continuous quantities. When dealing with categorical independent variables, the equivalent technique is discriminant correspondence analysis.

*Merits*: They are robust against noises and cost effective.

*Demerits*: Classification accuracy is less.

### Principal Component Analysis:

**PCA** – PCA is the pattern recognition algorithm that discusses the exploration of natural symmetries (mirror images) in a well defined family of human faces. This results in an extension of data and imposes even and odd mirror images on the eigen functions of covariance matrix. Hence it is called as Principal Component Analysis.

Suppose the goal of the approach is to represent a picture of face in terms of an optimal coordinate system. The set of basis vectors which make up the coordinate system
will be referred to as Eigen pictures. They are simply the Eigen values of the covariance matrix. The average of the original sample face is determined. This method can be viewed as one that is useful for extracting the facial features.

In Principal Component Analysis the steps used to implement it are: i) Get some data and subtract the mean from the data. ii) Find out the covariance matrix. iii) Calculate the eigenvectors and Eigen values of covariance matrix. iv) Then choose the principal component to form the feature vectors. These principal components are the Eigen vectors with highest Eigen values.

**Merits:** Some of the advantages of this technique are reduced parameterization and consequent data reduction. Less cost is required as the size of the matrix is not doubled in eigenvector calculation.

**Demerits:** Some of the shortcomings are it failed to capture the minor variance unless they are explicitly accounted in the training data. It can only separate pair wise linear dependencies between pixels.

**Independent Component Analysis Algorithm:**

ICA - Independent component analysis (ICA)\(^{[42]}\) is a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals. It is a special case of blind source separation.

When the independence assumption is correct, blind ICA separation of a mixed signal gives very good results. It is also used for signals that are not supposed to be generated by a
mixing for analysis purposes. A simple application of ICA is the "cocktail party problem", where the underlying speech signals are separated from a sample data consisting of people talking simultaneously in a room. Usually the problem is simplified by assuming no time delays or echoes. An important note to consider is that if N sources are present, at least N observations (e.g. microphones) are needed to get the original signals. ICA finds the independent components (aka factors, latent variables or sources) by maximizing the statistical independence of the estimated components. One of many ways can be chosen to define independence, and this choice governs the form of the ICA algorithms.

The two broadest definitions of independence for ICA are Minimization of Mutual Information and Maximization of non-Gaussianity\cite{43}. The Minimization-of-Mutual Information (MMI) family of ICA algorithms uses measures like Kullback-Leibler Divergence\cite{44} and maximum-entropy. The Non-Gaussianity family of ICA algorithms, motivated by the central limit theorem, uses Kurtosis and Negentropy\cite{45}. Typical algorithms for ICA use centring, whitening (usually with the eigenvalue decomposition), and dimensionality reduction as preprocessing steps in order to simplify and reduce the complexity of the problem for the actual iterative algorithm. Whitening and dimension reduction can be achieved with principal component analysis or singular value decomposition. Whitening ensures that all dimensions are treated equally a priori before the algorithm is run.

Merits: ICA is important to blind signal separation and has many practical applications. It is closely related to (or even a special case of) the search for a factorial code of the data, i.e., a new vector-valued representation of each data vector such that it gets uniquely encoded by
the resulting code vector (loss-free coding), but the code components are statistically independent.

Demerits: In general, ICA cannot identify the actual number of source signals, a uniquely correct ordering of the source signals, nor the proper scaling (including sign) of the source signals.

2.2.2 Texture Analysis

The iris has a particularly interesting structure and provides abundant texture information. So, it is desirable to explore representation methods which can capture local underlying information in an iris. From the viewpoint of texture analysis, local spatial patterns in an iris mainly involve frequency and orientation information[46]. Generally, the iris details spread along the radial direction in the original image corresponding to the vertical direction in the normalized image. As a result, the differences of orientation information among irises seem to be not significant. That is, frequency information should account for the major differences of irises from different people.

Hence, such discriminating frequency information which reflects the local structure of the iris is captured. In general, the majority of useful information of the iris is in a frequency band of about three octaves. Therefore, a bank of filters is constructed to reliably acquire such information in the spatial domain. The coefficients of a filtered image effectively indicate the frequency distribution of an image. Statistic values are thus extracted from each small region in the filtered image to represent local texture information of the iris. A feature vector is an ordered collection of all features from the local regions. Li Ma et al have developed a recognition system based on this methodology.
2.2.3 Hough Transform

The Hough transform\textsuperscript{[47]} is a standard computer vision algorithm that can be used to determine the parameters of simple geometric objects, such as lines and circles, present in an image. The circular Hough transform can be employed to deduce the radius and centre coordinates of the pupil and iris regions. An automatic segmentation algorithm based on the circular Hough transform is employed by Wildes et al\textsuperscript{[48]}.

2.2.4 2-D Gabor Wavelet Transform

In 1993, Daugman developed a successful system by using the 2-D Gabor wavelet transform\textsuperscript{[49]}. In this system, the visible texture of a person's iris in a real-time video image is encoded into a compact sequence of multi-scale quadrature 2-D Gabor wavelet coefficients, whose most significant bits consist of a 256-byte “iris code.” This texture analysis method on iris recognition uses the bank of Gabor filter to capture both local and global iris characteristics to form a fixed length feature vector. Iris matching is based on the Euclidean distance between the two corresponding iris vectors. The method described is based on edge detection approach and Wildes uses a histogram based filling method. The prototype of Wildes relied on image registration and image matching is computationally very demanding.

In 1996, Wildes et al\textsuperscript{[50]} developed a prototype system based on an automated iris recognition that uses a very computationally demanding image registration technique. This system exploits normalized correlation over small tiles within the Laplacian pyramid bands as a goodness of match measure. Boles and Boashash\textsuperscript{[51]} proposed an iris identification system in which zero-crossing of the wavelet transform at various resolution levels is calculated over concentric circles on the iris, and the resulting 1-D signals are compared with
the model features using different dissimilarity functions. Ma et al.\textsuperscript{[52]} also adopted wavelet multi-resolution analysis based on Gabor filtering for iris feature extraction.

2.2.5 Phase Based Method

The use of multi-resolution two-dimensional wavelets in image analysis and computer vision has attracted much interest in recent years. A novel and efficient approach to iris recognition based on iris feature extraction using phase based analysis was introduced by Daugman\textsuperscript{[53]}. This presents an iris recognition algorithm using phase-based image matching - an image matching technique using phase components in 2D Discrete Fourier Transforms (DFTs) of given images. One of the difficult problems in feature-based iris recognition is that the matching performance is significantly influenced by many parameters in feature extraction process (e.g., spatial position, center frequencies and size parameters for 2D Gabor filter kernel), which may vary depending on environmental factors of iris image acquisition. Addressing the problem, an efficient iris recognition techniques using phase-based image matching is proposed that uses phase components in 2D Discrete Fourier Transforms (DFTs) of given images. Phase-based method chooses the phase as an iris feature.

Gabor wavelets were used to extract the phase measures as the iris feature. The phase is coarsely quantized to four values and the iris code is 256 bytes long. Then the dissimilarity between the input iris image and the registered template can be easily determined. Wavelets can be used to decompose the data in the iris region into components that appear at different resolutions. Wavelets have the advantage over traditional Fourier transform in that the frequency data is localized, allowing features which occur at the same position and resolution to be matched up\textsuperscript{[54]}. A number of wavelet filters, also called a bank of wavelets, is applied to the 2D iris region, one for each resolution with each wavelet a scaled version of some basis
function. The output of applying the wavelets is then encoded in order to provide a compact and discriminating representation of the iris pattern. Gabor filters are able to provide optimum conjoint representation of a signal in space and spatial frequency.

A Gabor filter can be constructed by modulating a sine/cosine wave with a Gaussian. This is able to provide the optimum conjoint localization in both space and frequency, since a sine wave is perfectly localized in frequency, but not localized in space. Modulation of the sine with a Gaussian provides localization in space, though with loss of localization in frequency. Decomposition of a signal is accomplished using a quadrature pair of Gabor filters, with a real part specified by a cosine modulated by a Gaussian, and an imaginary part specified by a sine modulated by a Gaussian\[55\]. The real and imaginary filters are also known as the even symmetric and odd symmetric components respectively. The centre frequency of the filter is specified by the frequency of the sine/cosine wave, and the bandwidth of the filter is specified by the width of the Gaussian.

2.2.6 Rotation compensated iris matching

The use of multiple comparisons in many systems leads to higher storage requirements and increased time to enroll and verify. To address this problem, it is known that correlation based methods could be used to determine the shift or delay of one function with respect to another\[56\]. Applications of correlation in object detection have been extended to biometric authentication in face, fingerprint and iris recognition. Kumar et al\[57\] investigated the performance of advanced correlation filters for such applications since they are known to offer good matching performance in the presence of image variability. Aoki et al\[58\] used phase-based image matching to achieve good results in fingerprint and iris recognition. In all such work the use of 2D cross-correlation requires the storage of the entire
database of images along with their iris codes. Apart from a dramatic increase in storage requirements, two dimensional operations were computationally more intensive and slower the matching process significantly. Reducing this process to one dimension for iris matching has two major advantages. Apart from the obvious gain in both time and storage efficiency, there was improved accuracy due to elimination of eyelid and eyelash affected regions during the peak and displacement calculations. The correlation computation is carried out on a circular sequence generated by averaging a band of gray values from an annular region of the normalized iris image near the pupil. This has the advantage for Fourier domain calculation of being naturally periodic.

Biometric systems in general work in two stages: enrolment and matching. In this system proposed by Donald M. Monro\[^{[59]}\] for the human iris, during enrolment an eye image was acquired, the iris was normalized and a binary feature vector was generated using patch-based zero-crossings of FFT\[^{[l]}\] amplitudes. For each enrolled iris, a periodic sequence was extracted from the 512x80 normalized irises by averaging rows 5-9 of the image, counted from the pupil boundary. This avoided outer regions which might be obscured by eyelashes or eyelids, and was far enough from the pupil boundary to avoid irregularities. The conjugate of the 1D FFT of this was then stored along with its feature vector.

During authentication, a newly acquired candidate image was normalized and the FFT of the same band is calculated for use in rotation compensation with the enrolled iris FFT. For the verification process, the rotation calculation is done only once, with the claimed subject’s stored FFT. For identification, the calculation can be carried out with each enrolled FFT. In either case the product of the 1D FFT from the candidate iris with the stored conjugate FFT is formed, and the inverse FFT of the product is calculated to give the cross-
correlation between the two sequences. After normalization, the peak sharpness was measured and its location is noted. For similar irises a sharp peak is expected while a more flat curve would correspond to a non-match. The degree of iris-rotation is indicated by the position of the peak. If the peak is sufficiently sharp, the normalized image from the candidate iris is shifted into alignment with the registered iris and the iris code is finally calculated for matching. Since the pre-selection is based on the correlation peak, it is necessary to have a metric for independent discrimination. To make such a decision robust to image variability, it should be based on a larger region of the correlation output.

From observations of various correlation plots it was concluded that a good metric might be based on the variance in a restricted region about the peak. A number of metrics had been tried, and good discrimination was obtained with a Peak-to-Sidelobe Ratio (PSR). For this, the sidelobe value $S$ was taken as the mean of a 333 value region centered on the peak, i.e. 166 correlation values either side of the peak, excluding 13 values centered on the peak. The narrow peak and wide sidelobe exploit the global dominance of the matching peak while not allowing local maxima of non matches to bias the discrimination.

Donald M. Monro has also proposed a patch code technique for the iris matching. The iris image is divided into rectangular patches with particular size and orientation to apply this iris coding technique. First a 1D intensity signal is obtained by averaging the patch across its width to reduce noise. Using broad patches also makes iris image registration easier, which is important for rotation invariant iris recognition. The FFT is then applied to this 1D signal to obtain spectral coefficients. In order to reduce the spectral leakage during the FFT, a window is employed before the FFT. The Frequency Magnitude Differences between adjacent patches are calculated and a short binary code is generated from the zero crossings of each
difference. These constitute the feature vectors of the iris code. A nearest neighbor classifier is used in matching. The distance between different iris feature vectors is measured by the weighted Hamming Distance.

A novel method for iris matching using zero crossings of a one dimensional Discrete Cosine Transform (DCT) as a means of feature extraction for later classification was also proposed by Donald M.Monro\textsuperscript{[60]}. The DCT of a series of averaged overlapping angular patches were taken from normalized iris images and a small subset of coefficients was used to form sub feature vectors. Iris codes were generated as a sequence of many such sub features, and classification is carried out using a weighted Hamming distance metric. He have also demonstrated the use of novel patch encoding methods in capturing iris texture information, proposed the worst-case (nearest nonmatch) EER [Equal Error Rate] as a new practical metric for evaluating systems, and investigated better classifier designs for wider interclass seperability.

2.2.7 Evaluation and Comparison

- Phase based method is relatively simple and efficient compared to existing methods. Experimental results show that this approach has good recognition performance and speed. Daugman’s approach has the ability of comparing millions of iris codes within a minute.

- In a zero crossing representation of a one-dimensional wavelet transform was calculated to characterize the texture of the iris. In, a texture analysis approach was proposed. Multi-channel Gabor filtering was used to capture global and local details in an iris image. Additionally, independent component analysis (ICA), and wavelet packets for iris pattern analysis have been used.
Table 2.1 Comparative study of Existing Systems

<table>
<thead>
<tr>
<th>Features</th>
<th>Phase based Method</th>
<th>Zero crossing Method</th>
<th>Texture based Method</th>
<th>Intensity variation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature vector</td>
<td>Gabor wavelet used</td>
<td>Zero crossings of wavelet</td>
<td>Bank of spatial filters</td>
<td>Based on ICA</td>
</tr>
<tr>
<td>Matching</td>
<td>Phase components in 2DDFT</td>
<td>Fast matching scheme using XOR operation</td>
<td>Euclidean distance between two iris vectors</td>
<td>Hamming Distance</td>
</tr>
<tr>
<td>Accuracy</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Security</td>
<td>Excellent</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Demerit</td>
<td>Nill</td>
<td>Time consuming</td>
<td>User acceptance is minimum</td>
<td>Lack in accuracy</td>
</tr>
</tbody>
</table>

- One of the most interesting aspects of texture synthesis method is that it can be considered to be made up of patterns. A pattern is essentially an arrangement. It is characterized by the order of the elements of which it is made, rather than by the intrinsic nature of these elements.

- It is required that iris image be effectively transformed into the distinctive feature vectors in ICA for recognition is used in intensity variation method. So, the size of the iris code should be compact for enrollment or recognition, and it must be accomplished at very high rates of speed to extract iris feature.
• One of the most interesting aspects of texture synthesis method is that it can be considered to be made up of patterns. A pattern is essentially an arrangement. It is characterized by the order of the elements of which it is made, rather than by the intrinsic nature of these elements.

• A pattern is essentially an arrangement. It is characterized by the order of the elements of which it is made, rather than by the intrinsic nature of these elements.

2.3 VARIOUS FEATURE ENCODING ALGORITHMS

2.3.1 Wavelets

Wavelets are used to decompose the data in the iris region into components that appear at different resolutions. Wavelets have the advantage over traditional Fourier transform, where the frequency data is localized, matching features which occur at the same position and resolution. A number of wavelet filters, also called a bank of wavelets, can be applied to the 2D iris region, one for each resolution with every wavelet a scaled version of some basis function. The result of applying the wavelets is then encoded in order to provide a compact and discriminating representation of the iris pattern.

2.3.2 Gabor Filter

Gabor filters are able to provide optimum conjoint representation of a signal in space and spatial frequency. A Gabor filter is constructed by modulating a sine/cosine wave with a Gaussian. This is able to provide the optimum conjoint localization in both space and frequency, since a sine wave is perfectly localized in frequency, but not localized in space. Modulation of the sine with a Gaussian provides localization in space, though with loss of
localization in frequency. Decomposition of a signal is accomplished using a quadrature pair of Gabor filters, with a real part specified by a cosine modulated by a Gaussian, and an imaginary part specified by a sine modulated by a Gaussian. The real and imaginary filters are also known as the even symmetric and odd symmetric components respectively.

The centre frequency of the filter is specified by the frequency of the sine/cosine wave, and the bandwidth of the filter is specified by the width of the Gaussian. Daugman makes use of a 2D version of Gabor filters in order to encode iris pattern data\textsuperscript{[49]}. A 2D Gabor filter over the image domain \((x,y)\) is represented as

\[
G(x,y) = e^{-\pi[(x-x_0)^2/\alpha^2+(y-y_0)^2/\beta^2]}e^{-2\pi i\left[\mu_0(x-x_0)+v_0(y-y_0)\right]}
\]

where \((x_o,y_o)\) specify position in the image, \((\alpha,\beta)\) specify the effective width and length, and \((u_o,v_o)\) specify modulation, which has spatial frequency \(\omega_0 = (u_o^2 + v_o^2)^{1/2}\). Daugman demodulates the output of the Gabor filters in order to compress the data. This is done by quantatizing the phase information into four levels, for each possible quadrant in the complex plane. It has been shown by Oppenheim and Lim\textsuperscript{[61]} that phase information, rather than amplitude information provides the most significant information within an image. Taking only the phase will allow encoding of discriminating information in the iris, while discarding redundant information such as illumination, which is represented by the amplitude component.

\subsection*{2.3.3 Haar Wavelet}

Lim et al\textsuperscript{[62]} also use the wavelet transform to extract features from the iris region. Both the Gabor transform and the Haar wavelet are considered as the mother wavelet. From
multi-dimensionally filtering, a feature vector with 87 dimensions is computed. Since each dimension has a real value ranging from -1.0 to +1.0, the feature vector is sign quantised so that any positive value is represented by 1, and negative value as 0. This results in a compact biometric template consisting of only 87 bits. Lim et al. compare the use of Gabor transform and Haar wavelet transform, and show that the recognition rate of Haar wavelet transform is slightly better than Gabor transform by 0.9%.

2.4 STUDIES ON FEATURE EXTRACTION USING VARIOUS METHODS

2.4.1 Li Ma et al Method

From the viewpoint of texture analysis, local spatial patterns in an iris mainly involve frequency and orientation information. Generally, the iris details spread along the radial direction in the original image corresponding to the vertical direction in the normalized image. As a result, the differences of orientation information among irises seem to be not significant. Hence, frequency information should account for the major differences of irises from different people. Thus Li Ma et al have proposed a scheme\textsuperscript{[63]} to capture such discriminating frequency information which reflects the local structure of the iris. The majority of useful information of the iris is in a frequency band of about three octaves. Therefore, a bank of filters is constructed to reliably acquire such information in the spatial domain. The coefficients of the filtered image effectively indicate the frequency distribution of an image. Two statistic values are extracted from each small region in the filtered image to provide local texture information of the iris. Thus a feature vector is said to be an ordered collection of all features from the local regions.

In a Iris recognition method designed by Li Ma et al\textsuperscript{[64]}, they record the position of local sharp variation points as features instead of locating and recognizing those small
blocks. First, the background in the iris image is removed by localizing the iris. In order to achieve invariance to translation and scale, the annular iris region is normalized to a rectangular block of a fixed size. After lighting correction and image enhancement, a set of 1-D intensity signals containing the main intensity variations of the original iris for subsequent feature extraction is constructed. Using wavelet analysis, they record the position of local sharp variation points in each intensity signal as features. Directly matching a pair of position sequences is also very time-consuming. So they have adopted a fast matching scheme based on the exclusive OR operation to solve this problem. Local sharp variations are generally used to characterize the important structures of transient signals.

Hence a set of 1-D intensity signals which are capable of retaining most sharp variations in the original iris image is constructed. Wavelet transform is a particularly popular approach to signal analysis and has been widely used in image processing. In this design, a special class of 1-D wavelets (the wavelet function is a quadratic spline of a finite support) is adopted to represent the resulting 1-D intensity signals. The position of local sharp variation points is recorded as features.

Local details of the iris generally spread along the radial direction in the original image corresponding to the vertical direction in the normalized image. Therefore, information density in the angular direction corresponding to the horizontal direction in the normalized image is much higher than that in other directions i.e., it may suffice only to capture local sharp variations along the horizontal direction in the normalized image to characterize an iris. In addition, since the basic idea is to represent the randomly distributed blocks of the iris by characterizing local sharp variations of the iris, it is unnecessary to capture local sharp variation points in every line of the iris image for recognition.
Bearing these two aspects in mind, the 2-D normalized image into a set of 1-D intensity signals $S$ is decomposed. $I$ is the normalized image of $K \times L$. $I_x$ denotes gray values of the $x$th row in the image $I$. $M$ is the total number of rows used to form a signal $S_i$, $N$ is the total number of 1-D signals. In essence, each intensity signal is a combination of $M$ successive horizontal scan lines which reflect local variations of an object along the horizontal direction. A set of such signals contains the majority of the local sharp variations of the iris. Moreover, such processing reduces the computational cost required for subsequent feature representation.

In experiments, they found that the iris regions close to the sclera contain few texture characteristics and are easy to be occluded by eyelids and eyelashes. Therefore, they have extracted features only in the top-most 78% section (corresponding to the regions closer to the pupil) of the normalized image. The relation between the total row number $K$ of the normalized image, the total number $N$ of 1-D signals and the number $M$ of rows used to form a 1-D signal is denoted as $K \times 78\% = N \times M$. Since the total row number $K$ of the normalized image is fixed, the product of the total number $N$ of 1-D signals and the number $M$ of rows used to form a 1-D signal is a constant in experiments. The recognition rate of the proposed algorithm can be regulated by changing the parameter $M$. A small $M$ leads to a large set of signals which results in characterizing the iris details more completely, and thus increases recognition accuracy. A large $M$, however, implies a lower recognition rate with a higher computational efficiency. This way, the tradeoff between speed and accuracy is done. In experiments, $M=5$ and $N=10$ were chosen.
2.4.2 Bamberger Method\textsuperscript{[65]}

The Limitations of using the Gabor filters lies in the selection of filters, which is dependent on the frequency properties of the image. This attribute representation should have enough information to separate various irises and be less sensitive to noise. The Directional Filter Bank [DFB] proposed by Bamberger is based on the analysis bank and a synthesis bank. The analysis bank of the DFB divides the original image into eight sub band images while the synthesis bank combines the sub band images into one image. DFB assures separable one-dimensional filtering and reconstruction of the original input image, yet it has the disadvantage of frequency scrambling, when a low frequency area is misplaced in the sub band images resulting in distortions in the decomposed directional sub band images. However, the DFB can be modified by eliminating frequency scrambling through re-sampling and back-sampling matrices.

The main property of the DFB is its ability to extract 2-D directional information of an image, which is important in image analysis and other applications. It has been used in texture classification, image de-noising, image enhancement and recognition etc. The DFB is fully decimated and perfect reconstruction (PR). This means that the total number of sub band’s coefficients is the same as that of the original image, and used to reconstruct the original image without error. It can be implemented by a tree structure consisting of three levels of two-band systems. Each level can be implemented by using separable poly phase filters, which make the structure extremely computationally efficient. The DFB is improved in, where visible sub bands are constructed. One major drawback of the conventional DFB is the way the low-frequency band is split. It is known that for natural images, the energy is concentrated at DC and its neighboring bands. Hence, a small perturbation occurring at this
region can have a significant impact on the directional information in the sub bands. All the sub bands in the conventional DFB meet at DC and, thus, require a very sharp frequency partition in order to have accurate estimation. In practice, such ideal brick wall filters are not feasible and therefore, the low-frequency component is usually removed before the DFB is applied. Hence, the resulting decomposition is no longer maximally decimated, which explains why the DFB has found limited applications.

2.4.3 Gupta Et al Method

P. Gupta et al\textsuperscript{[66]} have devised a recognition system using corner detection of Iris. Corners in the normalized iris image can be used to extract attributes for distinguishing two iris images. Firstly, the normalized iris image is used to detect corners using covariance matrix. Secondly, the detected corners between the database and query image are used to find cross correlation coefficient and at last if the number of correlation coefficients between the detected corners of the two images is greater than a threshold value then the candidate is accepted by the system.

In a methodology proposed by Donald M. Monro,\textsuperscript{[67]} for each enrolled iris a periodic sequence is extracted from the 512x80 normalized irises by averaging rows 5-9 of the image, counted from the pupil boundary. This avoids outer regions which may be concealed by eyelashes or eyelids, and is far enough from the pupil boundary to evade irregularities. The conjugate of the 1D FFT of this is then stored along with its feature vector. During authentication, a newly acquired candidate image is normalized and the FFT of the same band is calculated for use in rotation benefits with the enrolled iris FFT. For verification the rotation calculation will be done only once, with the claimed subject’s stored FFT. For identification, the calculation will be carried out with each enrolled FFT. In either case the
product of the 1D FFT from the candidate iris with the stored conjugate FFT is formed, and the reverse FFT of the product is calculated to give the cross-correlation between the two sequences. After normalization, the peak sharpness is measured and its location is noted. For similar irises a sharp peak is expected while a more flat curve would correspond to a non-match. The degree of iris-rotation is indicated by the position of the peak. If the peak is adequately sharp, the normalized image from the candidate iris is shifted into alignment with the registered iris and the iris code is finally calculated for matching.

2.4.4 Donald M Monro Methods

The technique presented by D. M. Monro and D. Zhang[67] exploits the local frequency variation as another novel method for coding the iris image. The iris image is divided into rectangular patches with particular size and orientation to apply the iris coding technique. The ‘patch’ is the basic fragment used in this technique. To obtain optimum performance the length, width, orientation (angle) and the relative position of a series of patches is tuned. First a 1D intensity signal is obtained by averaging the patch across its width to reduce noise. Using broad patches also makes iris image registration easier, which is important for rotation invariant iris recognition. The FFT is then applied to this 1D signal to obtain spectral coefficients. In order to reduce the spectral leakage during the FFT, a window is employed before the FFT. The Frequency Magnitude Differences between adjacent patches are calculated and a short binary code is generated from the zero crossings of each difference. These constitute the feature vectors of the iris code. A nearest neighbor classifier is used in matching. The distance between different iris feature vectors is measured by the weighted Hamming Distance. In order to make the iris recognition system rotation invariant, every registered iris image is coded from several initial positions around the circumference.
Seven ‘slip’ templates are made from one registered iris image. An iris to be recognized will be matched to each slip template of the registered iris image and the minimum distance taken as the matching distance.

D. M. Monro\cite{59} has also proposed a technique using Discrete Cosine Transform (DCT).\cite{70} The DCT is a real valued transform, which calculates a truncated Chebyshev series possessing well-known mini-max properties and can be implemented using the Discrete Fourier Transform (DFT). Due to its strong energy compaction property, the DCT is widely used for data compression. In addition, the feature extraction capabilities of the DCT coupled with well-known fast computation techniques have made it a candidate for pattern recognition problems.

In particular, the DCT has been shown to produce good results on face recognition, where it has been used as a less computationally intensive replacement for the Karhunen-Loeve transform (KLT),\cite{68} which is an optimal technique according to the least squares metric for projecting a large amount of data onto a small dimensional subspace. The KLT decomposes an image into principal components ordered on the basis of spatial correlation and is statistically optimal in the sense that it minimizes the mean square error between a truncated representation and the actual data. The DCT, with its variance distribution closely resembling that of the KLT, gives optimality with much lower computational complexity. Additionally, its variance distribution decreases more rapidly compared to other deterministic transforms. Although no transform can be said to be optimal for recognition, these well-known properties stimulates to investigate the DCT for effective non semantic feature extraction from human iris images.
As like in the Fourier-based iris coding work, it is started from a general paradigm whereby the feature vectors is derived from the zero crossings of the differences between 1D DCT coefficients calculated in rectangular image patches. Averaging across the width of these patches with appropriate windowing helps to smooth the data and mitigate the effects of noise and other image artifacts. This then enables us to use a 1D DCT to code each patch along its length, giving low-computational cost. The selection of the values for the various parameters had been done by extensive experimentation over the CASIA and Bath databases to obtain the best predicted Equal Error Rate (EER). The two data sets were used in their entirety to optimize the parameters of the method. Experimentally, overlapping patches have given the best EER in combination with the other parameters.

It has been found that horizontally aligned patches worked best, and a rotation of 45 degrees was better than 0 degrees or 90 degrees. This distinctive feature of the code introduces a blend of radial and circumferential texture allowing variations in either or both directions to contribute to the iris code. To form image patches, bands of pixels along 45 degree lines through the image are selected. A practical way of doing this is to slew each successive row of the image by one pixel compared to its predecessor. Patches are then selected in 11 overlapping horizontal bands. Each patch has eight pixels vertically (overlapping by four) and 12 horizontally (overlapping six). In the horizontal direction, a weighted average under a 1/4 Hanning window is formed. In effect, the resolution in the horizontal (iris circumferential) direction is reduced. Averaging across the width of the patch helps to reduce the degrading effects of noise and the use of broad patches makes for easier iris registration. In the vertical direction (45 degrees from the iris radial), eight pixels from each patch form a 1D patch vector, which is then windowed using a similar Hanning window.
prior to application of the DCT in order to reduce spectral leakage during the transform. The differences between the DCT coefficients of adjacent patch vectors are then calculated and a binary code is generated from their zero crossings. These 8-bit code fragments (codelets) are the basis of the matching process.

For comparing two iris codes, a nearest-neighbor approach is taken, where the distance between two feature vectors is measured using the product-of-sum (POS) of individual subfeature Hamming distances (HD)\[^{69}\]. The iris code as a rectangular block of size M x N, M being the number of bits per sub feature and N the total number of sub features in a feature vector. Corresponding sub feature bits are XOR-ed and the resultant N-length vector is summed and normalized by dividing by N. This is done for all M sub feature bits and the geometric mean of these M sums gives the normalized HD lying in the range of 0 to 1. For a perfect match, where every bit from Feature 1 matches with every corresponding bit of Feature 2, all M sums are 0 and so is the HD, while, for a total opposite, where every bit from the first Feature is reversed in the second, MN/Ns are obtained with a final HD of 1. Since a total bit reversal is highly unlikely, it is expected that a random pattern difference should produce an HD of around 0.5.

The motion estimation between two normalized iris images is carried out by taking successive blocks of size BlxBl with a 50% overlap from the first image and computing the vector to the nearest corresponding block in the second image within a Nh pixel neighborhood around it. The block matching criteria is the minimum normalized Mean Absolute Difference (MAD) between the pixels in the two blocks. It is identified that the vector magnitudes are small and their orientations are highly correlated. Variation will primarily be due to localization error and non-elastic deformations of the pupil. On the other
hand, the motion vectors from non-matching irises are large and random. While the matching X and Y-vector histograms have a distinct peak and fast decreasing tails, the non-matching ones are more uniform and evenly spread out. These relatively uniform histograms are thus least affected by iris shifts which cause the X-vector histogram peak to undergo a shift equal to the pixel shift between the images compared. The Y-Vector histogram is not affected due to iris rotations. For classification, the vector information is converted into numerical scores for comparison against a cutoff threshold to make a match/non-match decision. A number of metrics which are independent of the position of the vector histogram peaks were considered as classification metrics.

2.4.5 Daubechies Methods

The Daubechies wavelet transforms\cite{70} are other methods used by several recognition systems in the same way as the Haar wavelet transform by calculating the running averages and differences via scalar products with scaling signals and wavelets. The only dissimilarity between them consists in how these scaling signals and wavelets are defined. This wavelet type has stabilized frequency responses but non-linear phase responses. Daubechies wavelets use overlapping windows, so the high frequency coefficient spectrum reflects all high frequency changes. Therefore Daubechies wavelets are useful in compression and noise removal. Daubechies wavelet is one of the db wavelet families which is an orthonormal and compactly supported wavelet. For feature extraction, Daubechies wavelet is applied four times to the unwrapped iris. LL (HH) signifies that the low pass (high pass) filter is applied to the image along two dimensions (X, Y). The 4-level wavelet decomposition detail and approximation coefficients of unwrapped iris image are tracked. Since the unwrapped image has a size of 512× 56 pixels, after 4 times decomposition, the size of last part is 6× 34.
feature vector is arranged by combining 408 (408 = 6 × 68) features in the LH and HL of level-4 (vertical and horizontal approximation coefficients $V = [LH4 \, HL4]$) in this algorithm. Then based on the sign of each entry, the phase of each component of $V$ (in radians) is computed. For the facility of matching process, +1 is assigned to positive numbers and 0 to others.

2.4.6 Mehrotra Method

In a work proposed by Hunny Mehrotra[71], iris biometric system operates on an input iris image taken from CCD or infrared cameras. All subsequent stages depend upon the quality of acquired input image. The input image is passed to pre processing module. The changed area underlying between inner pupil boundary and outer iris boundary is used for extraction of unique features. A novel grouping of texture and phase based classifier is used for feature extraction. The texture information from normalized iris image is obtained using Haar wavelet decomposition and phase data is extracted using Laplacian of Gaussian (LOG) Gabor Wavelet. The matching scores from the individual recognizers are combined using weighted sum of score technique.

2.4.7 Xiaofu et al Method

In a paper by Xiaofu He and Pengfei Shi[72], an iris feature extraction method by using 2D-CWT [Continuous Wavelet Transform] is presented. The scheme of feature extraction is to use the multi-level coefficients of decomposition parts of normalized iris image via 2D-CWT, which can be implemented using a dual-tree structure. For each tree, its structure is similar to 2D DWT [Discrete Wavelet Transform] except that the different filters are applied for perfect reconstruction and the outputs of sub bands images are congregated into complex wavelet coefficients. These two trees have the same structure. In order to
realize perfect reconstruction from decomposed sub band images, a low pass filter and a high pass filter at the first level need to be specially designed and denoted as LPPF\_A (Low Pass Pre-Filter for tree A) and HPPF\_A (High Pass Pre-Filter for Tree A) for tree A, LPPF\_B and HPPF\_B for tree B respectively, which are called pre-filters. The other complex filters in the higher levels are set to LPF\_A (Low Pass Filter for Tree A) and HPF\_A (High Pass Filter for Tree A) for tree A, LPF\_B and HPF\_B for tree B respectively. Each normalized iris image was decomposed into n levels using 2D-CWT which resulted in 3(high frequency)\times2(trees)\times n high frequency components and 2 low frequency components from dual-tree structure.

The iris feature vector consists of high frequency decomposition coefficients as well as real and imaginary part from the highest level. For instance, the length of the iris feature vector L is defined as,

\[
L = \left(\frac{W}{2^n}\right) \times \left(\frac{H}{2^n}\right) \times 3(\text{high frequency}) \times 2(\text{trees})
\]

Where, W and H is the width and height of the normalized iris image and n is the levels. Since the real and imaginary coefficients of high frequency component are extracted as iris features, the phase information is preserved as a good directional selectivity.

**2.4.8 Neural Network Method**

An iris-recognition algorithm using neural networks first has to identify the approximately concentric circular outer boundaries of the iris and the pupil in a photo of an eye\textsuperscript{[73]}. The set of pixels covering only the iris is then transformed into a bit pattern that preserves the information that is essential for a statistically meaningful comparison between two iris images. The mathematical methods used resemble those of modern lossy compression algorithms for photographic images. A discrete wavelet transform is used in
order to extract the spatial frequency range that contains a good best signal-to-noise ratio considering the focus quality of available cameras. The result is a set of complex numbers that carry local amplitude and phase information for the iris image. The covariance matrix of discrete wavelet transform of Edge of iris images has been used for Iris Recognition using Probabilistic neural networks.

The output of neural network is the class of ship. This system can classify the noisy ship image very well. Covariance matrix of discrete wavelet transform of Edge of Iris image has been used as input of probabilistic neural network. This method for ship classifier design offers good class discriminancy. This method can classify noisy ship image. The canny edge detection has been used. Neural network that had been used is LVQ. LVQ, or Learning Vector Quantization, is a prototype-based supervised classification algorithm. LVQ can be understood as a special case of an artificial neural network. LVQ is one of the competitive neural networks.

The iris consists of many irregular small blocks, such as freckles, coronas, stripes, furrows, crypts, and so on. Furthermore, the distribution of these blocks in the iris is also random. Such randomly distributed and irregular blocks constitute the most distinguishing characteristics of the iris. Intuitively, if each of these blocks in the image are precisely located and the corresponding shape is recognized as well, then a high performance algorithm is obtained. But it is almost impossible to realize such an idea. Unlike fingerprint verification, where feature extraction can rely on ridge following, it is difficult to well segment and locate such small blocks in gray images. Moreover, classifying and recognizing the shape of such blocks is unpractical due to their great irregularity. From the viewpoint of signal processing, however, we can regard these irregular blocks as a kind of transient
signals. Therefore, iris recognition can be solved using some approaches to transient signal analysis. As we know, local sharp variations denote the most important properties of a signal.

2.4.9 Dyadic Wavelet Method

As a well-known multi resolution analysis approach, the dyadic wavelet transform\[74\] has been widely used in various applications, such as texture analysis, edge detection, image enhancement and data compression. It can decompose a signal into detail components appearing at different scales. The scale parameter of the dyadic wavelets varies only along the dyadic sequence. Here, the purpose is to precisely locate the position of local sharp variations which generally indicate the appearing or vanishing of an important image structure.

The dyadic wavelets satisfy such requirements as well as incur lower computational cost, and are thus adopted in this design. In this algorithm, the function is a quadratic spline which has a compact support and one vanishing moment. This means that local extremum points of the wavelet transform correspond to sharp variation points of the original signal. Therefore, using such a transform, one can easily locate the iris sharp variation points by local extremum detection. According to Mallat, the dyadic wavelet transform based on the above wavelet function could be calculated with a fast filter bank algorithm.

The wavelet transform of a signal includes a family of signals providing detail components at different scales. There is an underlying relationship between information at consecutive scales, and the signals at finer scales are easily contaminated by noise. Considering these two points (information redundancy at consecutive scales and the effect of noise on signals at finer scales), only two scales to characterize differences among 1-D intensity signals is used. It is reasonable to consider that a local minimum of the wavelet
transform described above denotes the appearing of an irregular block and a local maximum
denotes the vanishing of an irregular block. A pair of adjacent local extremum points (a
minimum point and a maximum point) indicates that a small block may exist between them.

However, there are a few adjacent local extremum points between which the
amplitude difference is very small. Such local extremum points may correspond to relatively
faint characteristics in the iris image (i.e., local slow variations in the 1-D intensity signals)
and are less stable and reliable for recognition. A threshold-based scheme is used to suppress
them. If the amplitude difference between a pair of adjacent local extrema is less than a
predetermined threshold, such two local extremum points are considered from faint iris
characteristics and not used as discriminating features. That is, only distinct iris
characteristics (hence local sharp variations) are utilized for accurate recognition.

For each intensity signal $S_i$, the position sequences at two scales are concatenated to
form the corresponding features. The feature vector generally consists of about 660
components. In addition, at each scale, each feature component denoting the actual position
of local sharp variation points is replaced by its difference with the previous component. This
will save memory since the difference (i.e., the interval between two consecutive local sharp
variation points) is less than 256 and can be represented in a byte.

Based on the analysis to the structure and the characteristic of iris texture, the merits
of adopting the wavelet transform coefficients to carry on the iris recognition are many.
Wavelet function is equal to the window function to some extent, but the width of the
window is variable, therefore it may solve the contradiction between the time resolution and
the frequency resolution well, its change rules enables the wavelet transform to have the fine
local characteristic which is effective to analyze the sudden-change signal and the strange
The small scale corresponds to the position where the signal changed suddenly; the big scale corresponds to the position where the signal changed slowly.

Wavelet transform has the superior characteristic of time-frequency analysis, the wavelet transform coefficients can reflect the character of the signal or function; It may display the parts easily where the inter-class differs mostly by the characteristic of its variable focus, thus the difference between the inter-class was magnified and it was useful for improving the classified accuracy. Multi-resolution analysis of wavelet transform decomposes the iris image to detail sub-strips of different scale and the different resolution; it may carry on processing and the observation from each scale for the image. The local characteristic of any signal may obtain through the scale and position of the wavelet transform.

The results after carrying on wavelet transform to the signal can be regarded as the linear substitution for the primary signal. Although this kind of transform did not reduce the dimension of the signal, it provided a possibility for observing the signal from another aspect actually, and made the wavelet transform coefficients describes the signal more predominantly. It can obtain a series of coefficients at different scales after wavelet transform, these coefficients described the characteristic of the signal completely, and thus they may serve as classifying the character subset.

It is observed that the upper eyelid and eyelash always stochastically covered much area of the upper iris, therefore the available texture information which located at the region of left half side in the rectangular iris image is few. Usually lower eyelid covered some pixels in the outer iris, but these pixels had few texture information. In the normalized iris image, it is only needed to use 32 row pixels which located at the right half side from top to
bottom (beginning from iris interior boundary to outside) that can satisfy the need of the iris recognition. The local detail of the iris is generally along the radial direction, namely in the normalized image along the vertical direction.

Therefore in the normalized image information intensity is higher along horizontal direction. It has been feasible to carry on wavelet transform in the normalized image along the horizontal direction to seek for the characteristic quantity. It is equal to carry on one dimension wavelet transform only, and the computation reduced greatly. The wavelet transform real coefficients of each point at different scale are computed, these coefficients can reflect the iris texture well, the shape of one block texture was adjacent to the shape of adopting wavelet when the corresponding coefficients were bigger, namely these coefficients can reflect iris texture information well.

For ease of the observation the wavelet transform coefficients had been carried on translation in the coordinate axis along the horizontal direction, and translated 200 units toward right, namely the coordinate “0” expressed coefficient - 200, “200” expressed coefficient 0, analogized in turn. It is discovered that the coefficient distribution under the different scale conformed to the normal distribution on the whole, when the coefficient is probably 0 the peak value was obtained, this provided the good chance for carrying on the feature extraction and the code. In order to reduce the complexity of computation, the binary code to the extraction characteristic points is used. Namely when the coefficient value is bigger than 0 or equal to 0 its corresponding position code of one scale is “1”, and marked this point as characteristic point; when the coefficient value is smaller than 0 its corresponding position code of one scale is “0”, and marked this point as non-characteristic point. In order to reduce the iris code size, a one dimensional wavelet transformation is used.
First, the iris area is converted into 1D form by dividing the iris area into 5x5 pixel blocks. Then the mean intensity of each block is calculated in a left to right, top to bottom fashion. Each signal is transformed using 2nd order Coiflet wavelet transformation with four decomposition levels. In each decomposition level, the wavelet transformation divides the signal into approximation signals and detail coefficients.

In this method, the 4th level approximation coefficients are used to represent the iris. To reduce the code size and the similarity measurement process, all coefficients were stored in integer form. The final stage is matching the iris code of the unknown and the iris stored in the system. The matching had been carried out by computing the distance between two iris codes. The Hamming distance (HD) is used to measure the distance between two iris codes.

2.4.10 Huang et al Method

The joint space-spatial frequency representations have received special attention in the fields of image processing, vision, and pattern recognition. Huang et al\cite{75} presented a multi-resolution decomposition technique, Empirical Mode Decomposition (EMD), which is adaptive and appears to be suitable for non-linear and non-stationary signal processing. The EMD method was originally proposed for the study of ocean waves, and found potential applications in geophysical exploration, underwater acoustic signals, noise removal filter and biomedicine etc. The major advantage of EMD is that the basic functions can be directly derived from the signal itself. Compared with Fourier analysis, EMD analysis is adaptive while the basis functions of Fourier analysis are linear combinations of fixed sinusoids. The principle of EMD is to decompose a signal into a sum of oscillatory functions, namely intrinsic mode functions (IMFs) that have the same numbers of extrema and zero-crossings or differ atmost by one; and they are symmetric with respect to local zero mean. The above
two conditions fulfill the physically necessary conditions to define a meaningful instantaneous frequency.

On the other hand, if blindly applied to any analytic signal, the instantaneous frequency may result in a few paradoxes; it may go beyond the band for band limited signals or it may not represent one of the frequencies in the Fourier spectrum in the global sense. So, those two conditions allow the calculation of a meaningfully instantaneous frequency. Specifically, the first condition is similar to the narrow-band requirement, whereas the second condition modifies a global requirement to a local one by using the local mean of the envelopes defined by the local maxima and the local minima. Moreover, this is to certify that the instantaneous frequency will not have unnecessary fluctuations as induced by asymmetric waveforms.

To make use of EMD for practical applications, the signal must have at least two extrema, one maximum and one minimum, to successfully decompose the signal into IMFs. And those IMF components are obtained from the signal by the means of an algorithm called shifting process. The algorithm proposed by Chien-Ping Chang et al extracts each mode locally and excludes the highest frequency oscillations out of the original signal. The EMD algorithm extracts the oscillatory mode which exhibits the highest local information from the data and leaves the remainder as a “residual” (“approximation” in wavelet analysis). According to the major merits of EMD, the process of deriving the basis functions is empirical and the basis functions are obtained dynamically from the signal itself.

By observing the sample iris images the irregular blocks of the iris are slightly darker than their surrounding areas. Therefore, it is reasonable to consider that the residual presents the basic characteristics of the iris and the detail denotes the variation of the noise
represented by the highest local information. For the ROI of each normalized iris image I, pixel sequences from the different rows are concatenated to form the 1-D vector V. After concatenation and before performing EMD, the linear re-scaling is applied to each vector to adjust the average of each data set to zero and normalize the standard deviation to unity before further using the ROI vector. After applying the EMD algorithm, the feature vector of each EMD residual from the 1-D vector VN can be obtained.

2.5 STUDY ON VARIOUS MATCHING ALGORITHMS

2.5.1 Shape Matching

The shape of an object is among its most typical features, because of its edges and curves. From the computational point of view, the main obstacle to matching or recognizing shapes is deformation. Since the possible deformations grows exponentially in the size of the shape, matching shapes globally by searching for the correct deformation is intractable in practice. The problem is further complicated by occlusion and missed edge detections. Hongzhi Wang [76] have represented a model for recognition directly, without requiring extensive learning. The global/local dilemma in shape matching by introducing image segmentation for shape matching is attacked. To address the unreliability of segmentations, averaging the overall possible segmentations weighted by their probabilities conditioned on the images was proposed. This method combines the advantages of global and local approaches where it can match shape globally yet efficiently, and is robust to local-shape variation yet remains sensitive to the detailed boundary shapes.

As this shape matching technique can be used for several objects, it cannot be used for the biometric components like fingerprint, face or Iris as the discriminant features of these components are closely coupled. Such biometric components look similar to human eye
and as the shapes are alike to its counterpart the computational view of these objects cannot be recognized by this type of matching technique.

### 2.5.2 Hamming Distance

The Hamming distance gives a measure of how many bits are the same between two bit patterns. Using the Hamming distance of two bit patterns, a decision can be made as to whether the two patterns were generated from different irises or from the same one. In comparing the bit patterns $X$ and $Y$, the Hamming distance, $HD$, is defined as the sum of disagreeing bits (sum of the exclusive-OR between $X$ and $Y$) over $N$, the total number of bits in the bit pattern.

$$HD = \frac{1}{N} \sum_{j=1}^{N} X_j \text{XOR} Y_j$$

Since an individual iris region contains features with high degrees of freedom, each iris region will produce a bit-pattern which is independent to that produced by another iris, on the other hand, two iris codes produced from the same iris will be highly correlated.

If two bit patterns are completely independent, such as iris templates generated from different irises, the Hamming distance between the two patterns should equal 0.5. This occurs because independence implies the two bit patterns will be totally random, so there is 0.5 chance of setting any bit to 1, and vice versa. Therefore, half of the bits will agree and half will disagree between the two patterns. If two patterns are derived from the same iris, the Hamming distance between them will be close to 0.0, since they are highly correlated and the bits should agree between the two iris codes. The Hamming distance is the matching metric employed by Daugman, and calculation of the Hamming distance is taken only with bits that are generated from the actual iris region.
2.5.3 Weighted Euclidean Distance

The weighted Euclidean distance (WED)\textsuperscript{[78]} can be used to compare two templates, especially if the template is composed of integer values. The weighting Euclidean distance gives a measure of how similar a collection of values are between two templates. This metric is employed by Zhu et al.

\[ WED(k) = \sum_{i=1}^{N} \frac{(f_i - f_{i}^{(k)})^2}{(\delta_{i}^{(k)})^2} \]

Where \( f_i \) is the \( i \)th feature of the unknown iris, and \( f_{i}^{(k)} \) is the \( i \)th feature of iris template, \( k \), and \( \delta_{i}^{(k)} \) is the standard deviation of the \( i \)th feature in iris template \( k \). The unknown iris template is found to match iris template \( k \), when \( WED \) is a minimum at \( k \).

2.5.4 Normalized Correlation

Wildes et al. make use of normalized correlation between the acquired and database representation for goodness of match\textsuperscript{[79]}. This is represented as

\[ \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} (p_1[i,j] - \mu_1)(p_2[i,j] - \mu_2)}{nm\sigma_1\sigma_2} \]

Where \( p_1 \) and \( p_2 \) are two images of size \( n \times m \), \( \mu_1 \) and \( \sigma_1 \) are the mean and standard deviation of \( p_1 \), and \( \mu_2 \) and \( \sigma_2 \) are the mean and standard deviation of \( p_2 \). Normalized correlation is advantageous over standard correlation, since it is able to account for local variations in image intensity that corrupt the standard correlation calculation.
2.6 SUMMARY

In the feature-based iris recognition, the matching performance is significantly influenced by many parameters in the feature extraction process which may vary, depending on the environmental factors of iris image acquisition. Given a set of test iris images, extensive parameter optimization is required to achieve a higher recognition rate. Addressing these problems, an image matching technique using only the phase components in 2D Discrete Fourier Transforms (DFTs) was proposed by Kazuyuki Miyazawa et al\cite{68}. The technique of phase based image matching has so far been successfully applied to high-accuracy image registration tasks for computer vision applications. Due to large databases for verification systems, the matching time for the recognition using the above phase based matching technique is quite high though the accuracy is good enough. To minimize the matching time in large databases, a new implementation of hierarchical and parallel phase based image matching techniques are proposed.

In this thesis, new methods and solutions are developed in order to improve the iris recognition system’s performance when using huge iris databases. An iris recognition system is a pattern recognition system which performs personal identification by using unique and rich texture information extracted from an iris image in order to establish the authenticity of a specific physiological characteristic possessed by the user. The iris recognition system applied in this thesis makes use of some of the existing techniques such as the canney edge detection algorithm, Hough transform and finally phase only correlation for template encoding as a matching metric.