Fusion of Face and Iris Biometrics in Security Verification Systems

VALENTINE AZOM
(215041656)

A thesis submitted in fulfilment of the requirements
for the degree of Master of Science in Computer Science

in the

School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal
Durban, South Africa

January, 2016

UNIVERSITY OF KWAZULU-NATAL

COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE

DECLARATION

The research described in this thesis was performed at the University of KwaZulu-Natal under the supervision of Prof. A.O. Adewumi. I hereby declare that all materials incorporated in this thesis are my own original work except where acknowledgement is made by name or in the form of a reference. The work contained herein has not been submitted in part or whole for a degree at any other university.

Signed:

Valentine Azom
Date: January 2016

As the candidate's supervisor I have approved/disapproved this dissertation for submission.

Signed:

Date: January 2016

UNIVERSITY OF KWAZULU-NATAL

COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE

DECLARATION 1: Plagiarism

I, Valentine Azom declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    - Their words have been re-written but, the general information attributed to them has been referenced.
    - Where their exact words have been used, their writing has been placed in italics and in quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables which are copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in References section.

V.S. Azom

January 2016

Dedication

This research work is dedicated to the Almighty God, who saw me through the thick and thin of this research.

Acknowledgements

I am ever grateful to God for His love, mercies, and faithfulness towards me, most especially for the gift of life, good health, inspiration, wisdom and understanding that he granted throughout my academic programme.

I would also want to convey my heartfelt thanks to my lovely parents Mr & Mrs Azom for being there for me all through this years and most especially through this post-graduate journey.

To my supervisor Prof. A.O. Adewumi, I say, "A Big Thank you" for the motivation, drive and provision of a conducive atmosphere for my academic pursuit. His profound insights and attention to detail inspired me to keep pushing. Gratitude must also be extended to Prof. Jules Raymond Tapamo for his contributions which added so much value and depth to this research. Same goes for Dr. Micheal Olusanya and Dr. Martins Akugbe for their care, concern and advice.

I would also like to thank my siblings Prince, Promise, Adeeze and Kingsley for their love,care, words of wisdom and financial support. I also want to salute my wonderful colleagues, members of the optimization and artificial intelligence group Mr & Mrs Peter Olukanmi, Alex Alochukwu, Peter Popoola, Ebenezer Popoola and Stephen Akandwanaho for their care, support, company, contribution and encouragement through out this journey. To my friends Jerry Oguniyi, Nerson pillay, Nati, Jide Enigbokan, Ayobami Akinyelu and Mafunda thank you for the care and support.

I will also cease this opportunity to thank the entire members of Deeper Life Bible Church, KwaZulu-Natal province for their continious mentorship both in the physical and spiritual sense. They have been there even at my darkest hour, giving me a shoulder to lean on. I pray that God will continue bless their endeavors.

Finally, I want to say thank you to the University of KwaZulu-Natal for giving me the golden opportunity to be part of this great research community; I say Thank you!

Abstract

Over the years, Single biometrics has been the most preferred authentication method used in enhancing security of real-world applications over traditional methods. This is because a biometric trait cannot be stolen or forgotten by the user. Regardless of the advantages that this method presents, it also has its limitations. The performance of single biometric systems is usually affected by the environment, user mode as well as physiological defects. Given these forseen defects, multi-biometric systems have been introduced in order to reduce the effect by combining more than one modality for recognition. When considered as a classification problem, the performance of both multi-biometric and single biometric systems are impaired by the large class imbalance between the genuine and impostor scores obtained from multiple matchers. This is because, the number of genuine scores available in the training data is proportional to number of users in the database, while the number of impostor scores is proportional to the square of the number of users in the database. Resultantly therefore, classification is highly likely to favor the impostor class as the genuine scores are under-represented in the training data.

This thesis builds on the aforementioned gaps and focuses on fusion schemes inorder to solve issues encountered with single biometrics and the large class imbalance problem in biometrics. This research priviledged face and iris modalities because face templates are non-intrusive during acquisition and iris templates are distinct, accurate, stable over time and located at the face region. This means that the cost acquisition is reduced as a single sensor can be used for this purpose.

Inorder to achieve the research objective, local and global feature extraction algorithms were employed on both face and iris images to extract feature vectors. Local Binary patterns, sub-pattern Principal Component Analysis and modular Principal Component Analysis were used as local methods, Principal Component Analysis and Linear Discriminant Analysis were used as global methods. Experimental results obtained for individual face and iris sub-system shows that local methods perform better on face images, while global methods perform better for iris images.

To show the effectiveness of multi-biometric systems in this research, a hybrid fusion scheme that combines three classifiers based on feature and score level with a decision level fusion rule is proposed. The first two classifiers were built by performing fusion at feature level with all feature extraction algorithms, while the third classifier was built by performing fusion at score level using a local and global fea-

ture extraction algorithm. Experimental outcome revealed that the hybrid fusion scheme outperformed its unimodal systems and comparable to other fusion schemes in literature, attaining recognition accuracy of 96.34% and EER of 1.7%.

Furthermore, to proffer solutions to the issue of large class imbalance in biometric data, a serial fusion methodology using Binary Particle Swarm Optimization (BPSO) and Incremental Relevance Vector Machines (iRVM) is proposed. Face recognition is first performed using optimal features obtained from BPSO algorithm, then iris images corresponding to top-k matchers of the face images are selected and used to generate the genuine and impostor scores for classification with iRVM. The serial fusion scheme is used to lower the number of impostor scores, thereby lessening the effect of large class imbalance on the biometric data, While iRVM provides the capability to train data in batches. The results obtained shows that the proposed scheme produced improved performance over its unimodal systems with recognition accuracy of 99.06% and EER of 0.47%.

Contents

Abbreviations

ACO	**A**nt **C**olony **O**ptimization
ANN	**A**rtificial **N**eutral **N**etwork
ATM	**A**utomated **T**eller **M**achine
BLPOC	**B**and **L**imited **P**hase **O**nly **C**orrelation
BPSO	**B**inary **P**article **S**warm **O**ptimization
BSA	**B**ack-tracking **S**earch **A**lgorithm
CCD	**C**entroid **C**ontour **D**istance
CGJD	**C**omplex **G**abor **J**et **D**escriptor
DCT	**D**iscrete **C**osine **T**ransform
2DLDA	**D**imensional **L**inear **D**iscriminant **A**nalysis
DNA	**D**eoxyribo**N**ucleic **A**cid
2DPCA	**D**imensional **P**rincipal **C**omponent **A**nalysis
DTCWT	**D**ual **T**ree **C**omplex **W**avelet **T**ransform
DWT	**D**iscrete **W**avelet **T**ransform
EA	**E**volutionary **A**lgorithm
EER	**E**qual **E**rror **R**ate
FAR	**F**alse **A**cceptance **R**ate
FES	**F**uzzy **E**xpert **S**ystems
FLD	**F**isher **L**inear **D**iscriminant
FRR	**F**alse **R**ejection **R**ate
GA	**G**enetic **A**lgorithm
GAR	**G**enuine **A**cceptance **R**ate
GMM	**G**aussian **M**ixture **M**odel

GOH **G**radient **O**riented **H**istogram

HTER **H**alf **T**otal **E**rror **R**ate

ID **I**dentification

***i*GRVM** **i**ncreamental **G**ranular **R**elevance **V**ector **M**achines

IOM **I**ris **O**n the **M**ove

IR **I**mbalance **R**atio

***i*RVM** **i**ncreamental **R**elevance **V**ector **M**achines

KNN **K**-**N**earest **N**eighbor

LDA **L**inear **D**iscriminant **A**nalysis

LBPH **L**ocal **B**inary **P**attern **H**istogram

mPCA **m**odular **P**rincipal **C**omponent **A**nalysis

NI **N**ature **I**nsipered

n_ gs **n**umber of **g**enuine **s**cores

n_ ims **n**umber of **im**poster **s**cores

NNC **N**earest **N**eighbor **C**lassifier

PCA **P**rincipal **C**omponent **A**nalysis

PDM **P**ropability **D**eformation **M**odel

ROC **R**eceiver **O**perating **C**haracteristic

ROI **R**egion **O**f **I**nterest

RR **R**ecognition **R**ate

RVM **R**elevance **V**ector **M**achines

SIFT **S**cale **I**nvariant **F**eature **T**ransformation

spPCA **sp**attern **P**rincipal **C**omponent **A**nalysis

SVM **S**upport **V**ector **M**achines

TER **T**otal **E**rror **R**ate

List of Included Articles

Articles in Peer-reviewed Conference Proceedings

1. V. Azom; A.O. Adewumi; J.R. Tapamo, "Face and Iris biometrics person identification using hybrid fusion at feature and score-level," in *Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*, pp.207-212, 26-27 Nov. 2015

   doi: 10.1109/RoboMech.2015.7359524 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7359524&isnumber=7359478

Articles under review for peer review journal

1. V. Azom; A.O. Adewumi; J.R. Tapamo, Jules-Raymond, "Face and Iris biometrics person identification using hybrid fusion at feature and score-level", submited to *Journal of Computer vision and image understanding*, Elsevier.

2. V. Azom; A.O. Adewumi; J.R. Tapamo, "Score fusion of face and iris biometrics : A case of a serial fusion using BPSO and iRVM", submitted to *Transactions on Pattern Analysis and Machine Intelligence*, IEEE.

List of Figures

List of Tables

Chapter 1

Introduction and background

1.1 Introduction

Personal identification refers to a process of associating a set of attributes to an individual for recognition purposes. In designing a system that manages the identity of an individual population, care needs to be taken in defining and creating a set of specific attributes that are distinct to each individual. Therefore, the ways of identifying a person can be divided into three categories (Nandakumar, 2008):

- *What you know* for example passords, identity number.
- *What you have* for example token key and ID card.
- *Who you are* that is face, voice and finger-print patterns.

While both the first and second category are widely used in most security applications, they still fall short when it comes to "*high risk*" security applications. This is because passwords, token keys and identification numbers can be stolen or even forgotten by a user. This limitation has prompted experts to shift their preference towards the third category which uses physical attributes of a person for authentication purposes. This is because a person's physical trait cannot be easily stolen or misplaced by the user. It is upon this premise that studies on biometric systems have been pursued.

Biometrics is the measurement of using human physiological and behavioural traits such as face, Iris, retina, finger-print, palm-print, hand-geometry, DNA, signature, odour, gait and handwriting amongst others for authentication purposes (Jain et al., 2007). A single biometric system stores distinct features of each individual which differentiates a genuine user from an impostor. With this system, one's physical presence is a pre-requisite before authentication process can be completed. These characteristics make single biometrics preferable to traditional methods such as passwords and token keys because hardly can a person's physical attributes be misplaced, stolen or forgotten.

In practical applications, biometric systems have been widely used for the following purposes (Griaule, 2014):

- Forensic applications: used for collecting evidence such as fingerprints in crime scene for identification of criminals. In surveillance, it is used for monitoring very busy places for any abnormal and suspicious behaviour.
- Government applications: digital biometric information can be included in identification documents issued by the government which include a driver's license, voter's Identity card and national identity card.
- Commercial applications: financial services represent the high risk areas, therefore biometric information can be used in ATMs, online banking systems and mobile banking systems to provide more security to customer assets.
- Immigration: biometrics is used to identify individuals for travel documents like border crossing and international passports.

1.2 Structure of Biometric Systems

Every biometric system requires a physical attribute of an individual, which it pre-processes and then extracts a set of features known as *templates* that are reserved in a database during the *enrollment* phase. In the *authentication* phase, a user presents his/her claimed identity known as the *probe*, which is compared with templates in the database to determine if a user is genuine or an impostor. In the following list below modules that make a biometric system are described as shown in Figure 1.1.

- **Acquisition module**

  This module is made up of sensors that capture the image of the biometric trait. This module plays a crucial role in terms of system performance, as capturing images with devices of poor quality or under poor environmental conditions will affect the accuracy of the system.

- **Pre-processing module**

  In this module, images captured are either enhanced or the region of interest is extracted through, a process known as *segmentation* for example performing histogram equalization on face image to improve its contrast and extracting a iris and pupil region in an eye image (Jain et al., 2007).

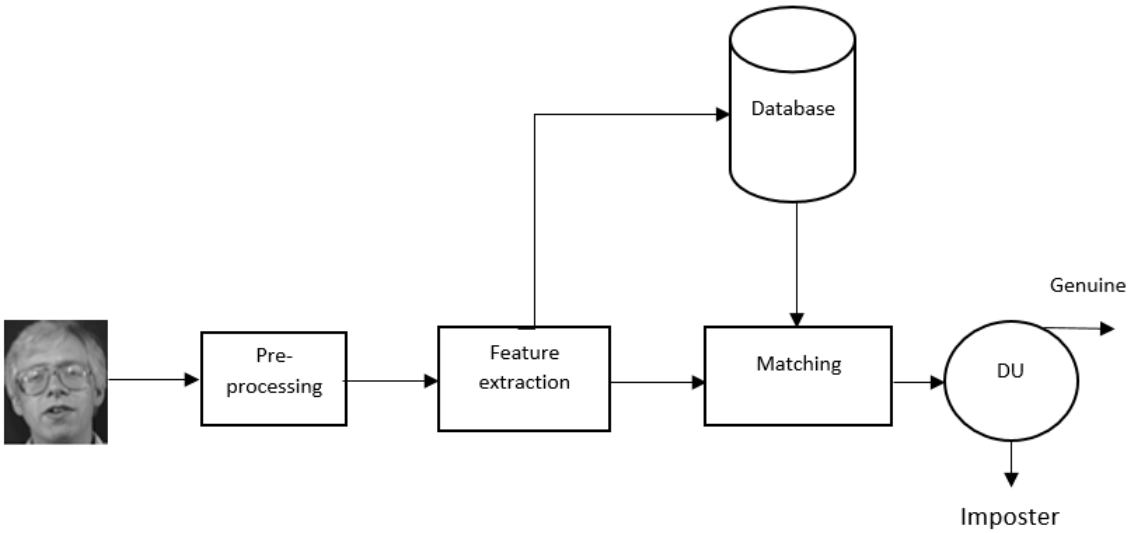


Figure 1.1: Structure of a biometric system

- **Feature extraction module**

  After the image has been pre-processed, significant details associated to the image are extracted either in form of one-dimension feature vector or set of key-points descriptors that best describe the image. These features are stored as *templates* in the biometric database as they are expected to be distinct and invariant to the probe image.

- **Matching module**

  In this module, the difference between the features obtained from the probe and stored templates is conducted to find the best match amongst the stored templates that is most similar the probe image.

There are two states in which a biometric system can operate namely; *verification* and *identification* state as displayed in Figure 1.2 (Ramadan et al., 2015). With verfication mode, a user's probe image is matched with only his/her template reserved in the biometric database and if the match score is less than the system threshold, the user's request is accepted. Otherwise, it is declined. In the identification state, the user's identity is found by matching the his/her probe image against all templates in the database. The template with the lowest match score that is less than the specified system threshold is termed as the claimed identity.

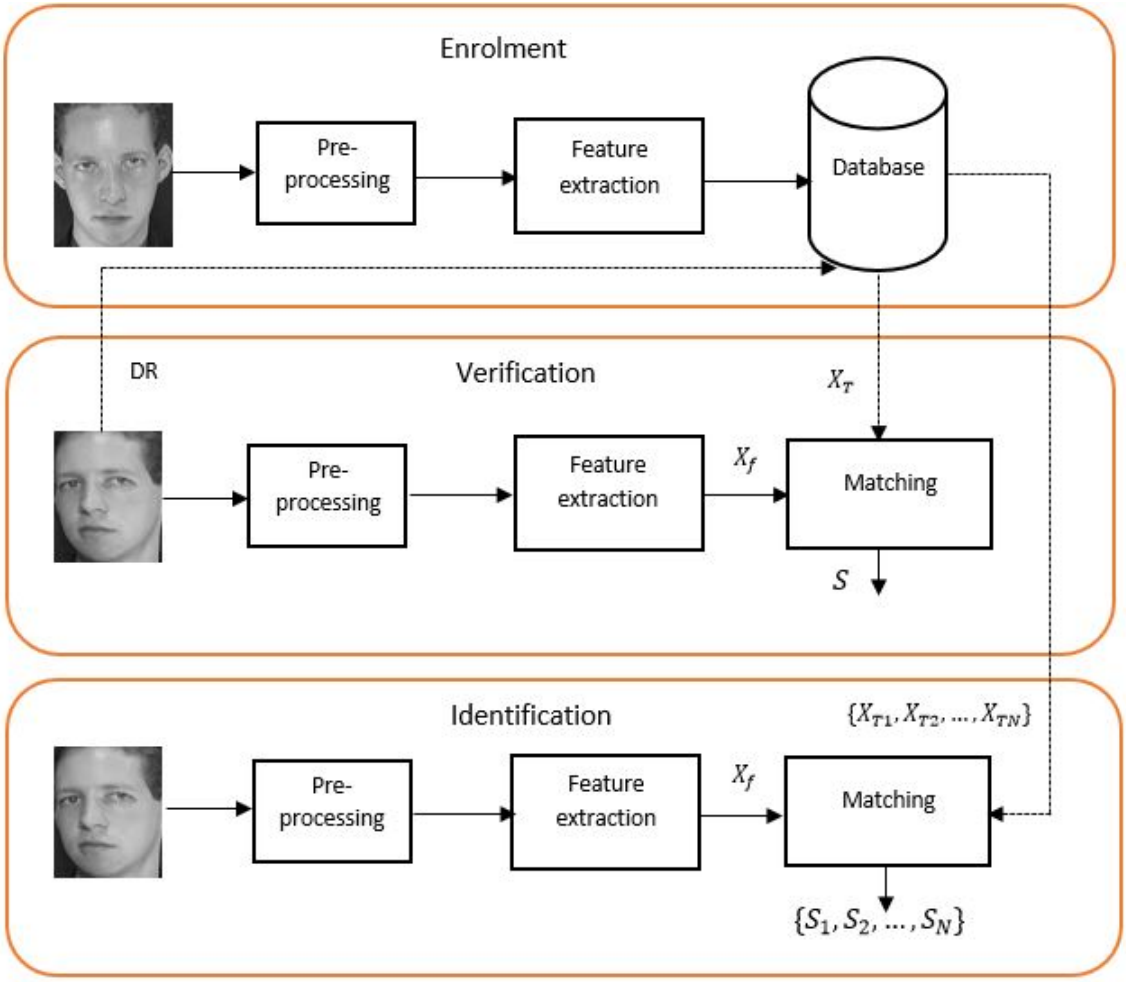


Figure 1.2: Operating modes for biometric systems (Jain et al., 2007),
Key: X_f and X_T represent the feature vectors of the probe image and stored templates and S_N represents matching score set

1.3 Multi-biometric systems

A biometric system that uses a single biological or physical trait for authentication is known as a *unimodal* system. In the design of unimodal system, the choice of biological traits is often determined based on the following criteria (Mehrotra, 2014):

- *universality* : every person has a biological trait that can be enrolled into the system.
- *permanence* : the trait does vary with aging.
- *uniqueness* : the biological trait is distinct for every individual.
- *collectability* : the ease of acquiring the individual trait.
- *performance* : refers to the accuracy achieved by the system.
- *acceptability* : preference in the user community.

In Table 1.1 shows the comparison of different modalities based on six characteristics explained above. A critical look at this table shows that, if a face biometric system is to be designed, the performance of the system will be affected by inter-class variance as not all facial templates will be distinct, this can occur when facial templates of identical twins or related individuals are captured. For example father and son, are captured during enrollment (Jain et al., 2004). More so, the shape and texture of the human face changes as one ages, making the permanence for this modality to be low. Similarly, considering the design of a fingerprint biometric system, the enrollment process can be affected by the problem of non-universality in cases where the fingerprint patterns cannot be captured for example elderly people have faded or damaged fingerprint samples that prevent the use of biometric systems. This is also true with other individuals, like manual labourers whose fingerprints are rarely usable (Nandakumar, 2008). In general, from Table 1.1 none of the single biometric systems are completely able to meet real world requirements, however when one or more of these templates are combined, they can be able to compliment the shortcomings of one another.

Therefore, a system that combines information presented by multiple sensors, algorithms, samples and traits is known as a *multi-biometric system* (Ramadan et al., 2015). This can be achieved by fusing multiple biometric templates of the same

Table 1.1: Comparsion of different modalities based on their characteristics (Ramadan et al., 2015), Key: H:High,M:Medium,L:Low

modality	universality	permanence	uniqueness	collectability	performance	acceptance
Face	H	L	M	H	L	H
Finger-print	M	H	H	M	H	H
Hand-geometry	M	M	M	H	M	M
Hand-vein	M	M	M	M	M	H
Iris	H	H	H	M	H	L
Voice-print	M	L	L	M	L	H
DNA	H	H	H	L	H	L
Gait	M	L	L	H	L	H

indivdual, or multiple samples of the same biometric, which can improve the performance of the system and prevent circumvention. Multi-biometric systems offer the following advantages as compared to single biometric systems (Jain et al., 2007):

- Multi-biometric systems provide solution to the issue of non-universality experienced in single biometric systems as it creates a certain degree of flexibility, that allows the user to be authenticated with another biometric, if one is not available. As an example, if a fingerprint pattern of a user cannot be captured, the iris or another biometric can be used to authenticate the user.

- Multi-biometric systems reduce circumvention from intruders in the sense that more than one biometric trait will have to be spoofed simultaneously before the system can be compromised. This capability makes multi-biometric systems non-susceptible to spoof attacks as compared to single biometric systems.

- Multi-biometric systems tackle the noise present in a biometric image. They provide the capability to combine a noisy image with the less noisy one, thereby enhancing the performance of the system. For example consider a face-iris multi-modal system, often times the image of the face biometric could be affected by poor illumination or bad camera quality. Combining the noisy face image with more accurate biometric such as the iris will enhance system performance.

- Multi-biometric systems allow indexing of large databases in order to reduce the search time. For example, the face images can be used to find the first top matchers which can then be used to make the final decision against an iris database.

Multi-biometric systems depend on templates obtained from multiple single biometric systems. Based on the type of biometric, multi-biometric systems can be divided into six classes namely:

- **Multi-sensor system**

  This system makes use of more than one sensor to capture same biometric trait of an individual. For example, different positions of a 2D face image can be captured and then combined into a 3D model for recognition (Kisku et al., 2010).

- **Multi-algorithm system**

  In this system, more than one feature extraction algorithms are employed to obtain the salient information from the same biometric of an individual. For example using minutiae [1] and texture based feature extraction algorithm for finger-print recognition (Jain et al., 2007).

- **Multi-instance system**

  This system combines different sample of the same biometric for example combining the right and left iris samples of an individual for recognition or combining the left and right thumb prints for recognition. This is often referred as multi-unit system in literature (Mehrotra et al., 2012).

- **Multi-sample**

  A system that captures multiple samples of the same biometric of an individual. For example capturing different facial expressions and positions for a face recognition system (Jain et al., 2007).

- **Multi-modal**

  This combines biometric evidences of one or more biological trait of an individual, for example combining face and iris templates of the same individual, combining fingerprint and iris templates (Azom et al., 2015).

[1] minutiae refers to significant features that describe a fingerprint image for example the ridge patterns of a fingerprint image

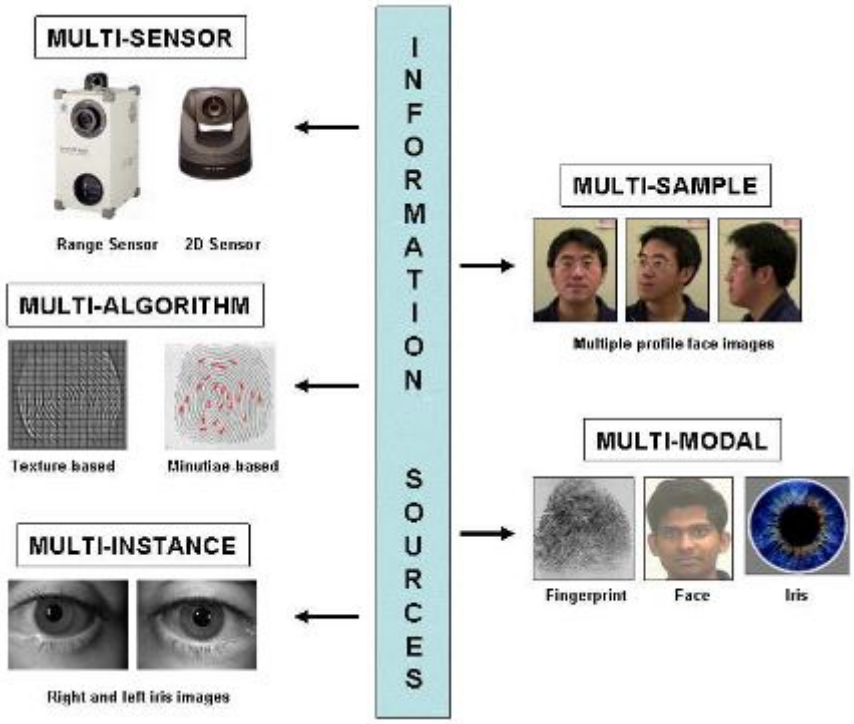


Figure 1.3: Types of Multi-biometric systems (Jain et al., 2004)

- **Hybrid system** The hybrid system incorporates more than one of the five systems described above. That is a system can both be multi-algorithmic and multi-modal at the same time (Jain et al., 2007; Azom et al., 2015).

Figure 1.3 shows a diagrammatic representation of the different multi-biometric systems just described above.

1.4 Performance of biometric systems

Consider a biometric system made up of N users with m images captured equally for each enrolled user I_n *where* $n = 1, 2, \cdots, N$. if $T = \{T_1, T_2, \cdots, T_i\}$ define the operating thresholds of the system and let $X = \{X_1, X_2, \cdots, X_Q\}$ represent the feature sets stored in the database. Authentication at operating threshold T_i is performed for a probe image P by first obtaining its feature set X_P and calculating the similarity S_i with every feature set in X. The system then makes its final

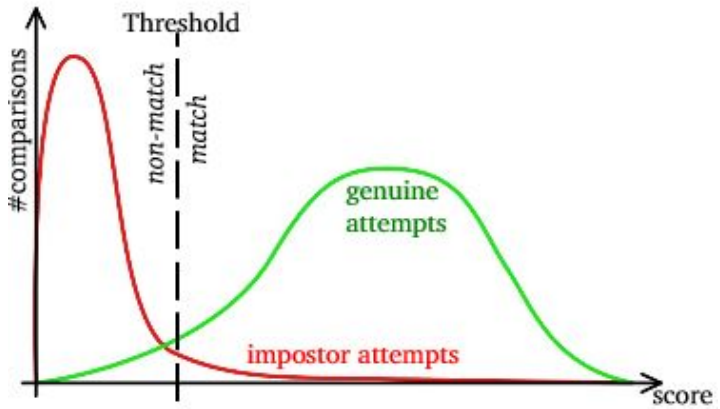


Figure 1.4: Genuine and Imposter score distribution (Jain et al., 2005)

decision using equation 1.4.1

$$\begin{aligned} &Let\ S_i = \|X_P - X_Q\|\ for\ Q = 1, 2, \cdots, N \\ &if\ n_0 = \underset{i}{\text{argmin}}\ S_i,\ then\ P = \begin{cases} genuine, & \text{if } S_{n_0} < T_i, \\ impostor, & \text{otherwise.} \end{cases} \end{aligned} \tag{1.4.1}$$

From the above equation S_i (where i is the index of a feature template in database) is said to be a *genuine score* if feature sets X_P and X_Q belongs to the same individual, otherwise it is called *impostor score*. The number of genuine scores and number of impostor scores are calculated as in equation 1.4.2, if one probe image of each user I_n is used for each attempt (Rattani and Tistarelli, 2009). Figure 1.4 shows the score distribution for both genuine and impostor attempts, as is revealed the two bell shaped curves overlap and the region of this overlap defines the FAR and FRR described below. In an ideal biometric system this two curves will not overlap, however this is difficult to achieve in real-life scenarios as there is no 100% accurate biometric system.

$$\begin{aligned} \text{number of genuine scores} &= N \times m \\ \text{number of imposter scores} &= N \times (N-1) \times m \end{aligned} \tag{1.4.2}$$

In general, the performance of a biometric system is determined by its recognition and error rates which are described as follows:

- **Recognition Rate (RR)** This is the percentage of correctly identified probe images fed into the biometric system and also known as the rank-1 recognition rate. The rank-k recognition rate refers to the number of correctly matched

templates that are contained in the top-k matchers selected from the database. The formula is defined below:

$$RR = \frac{R_k}{N} \tag{1.4.3}$$

where k represent the number of top-matchers selected and N represent the number of images in the database.

- **False Rejection Rate (FRR)** It is defined as the proportion of genuine users that were wrongly classified as impostors by the system. it generally calculated as the proportion of the genuine scores that are greater than the operation threshold T_i.
- **False Acceptance Rate (FAR)** It is defined proportion of impostors who were wrongly classified as genuine users. It is often calculated as the proportion of impostor scores that are less than the operating threshold T_i.
- **Genuine Acceptance Rate (GAR)** The proportion of genuine users correctly identified by the biometric system and it is calculated as $1 - FRR$.
- **Equal Error Rate (EER)** is the point where FRR is equal to the FAR. In general, the lower the EER the more accurate the system becomes.
- **Total Error Rate (TER)** is the sum of the FRR and FAR. The minimium value of this error is expressed as $2 \times EER$. The lower the minimum TER the more accurate the system becomes.
- **Receiver Operating Characteristics Curve** This is the curve that shows the performance of the system at different thresholds. it is a plot of the FRR or GAR against the FAR.

The accuracy of a biometric system deployed to production environment can also be expressed in the form "FRR 1%@FAR 1/10000" or "GAR 99%@FAR 1/10000" . *This means that at the operating threshold the system,* 1%*(*99%*) of the genuine users are rejected(accepted), considering one out of ten thousand imposter attempts is accepted as a match.*

1.5 Problem Statement

Although biometrics are generally preferred over traditional systems and have been deployed in real-life applications, there are still a number of challenges associated with it and these need to be addressed as no ideal biometric system has been achieved. These issues include:

- *Non-universality*: If a biometric system is able to capture the trait of every individual in its population, then it is said to be universal. However, not all biometric systems are universal as it is very difficult to capture fingerprints of the population made up of elderly people and manual workers (Eskandari et al., 2013). More so, it will be difficult to obtain good quality of iris images from people who have eye ailments such as glaucoma and cataract among others. This leads to high TER of the biometric system.

- *Noisy data*: This is mainly a result of poor sensor quality or environmental conditions. The performance of biometric systems depends to a large extent, on the quality of input trait. Therefore, poor maintenance of sensors may lead to poor image quality due to the presence of dirt remains (Islam, 2014). Also the illumination conditions present at the point of capture may also affect the quality of the biometric, if the illumination conditions are poor, there is a high likelihood of capturing poor quality images.

- *Inter-class variance*: occurs when there is an overlap between templates of users within the biometric population. In biometrics, feature templates obtained from each individual are meant to be unique, however this is not always the case as two identical twins and genetically related individuals (e.g father and son) can be registered in a face recognition system (Nandakumar, 2008). This increases the FAR of the system.

- *Intra-class variance*: templates stored in databases exhibit large variation either due to improper interaction between the user and sensor or some other changes in environmental conditions. This usually occurs for a face recognition system when the facial expression present in the probe image vary from that stored in the database (Fathima et al., 2014). There are also cases where face texture changes, a good example being the presence of wrinkles, which emerge when one ages. Normally, features extracted from users are meant to be invariant to these changes, however this is not always the case.

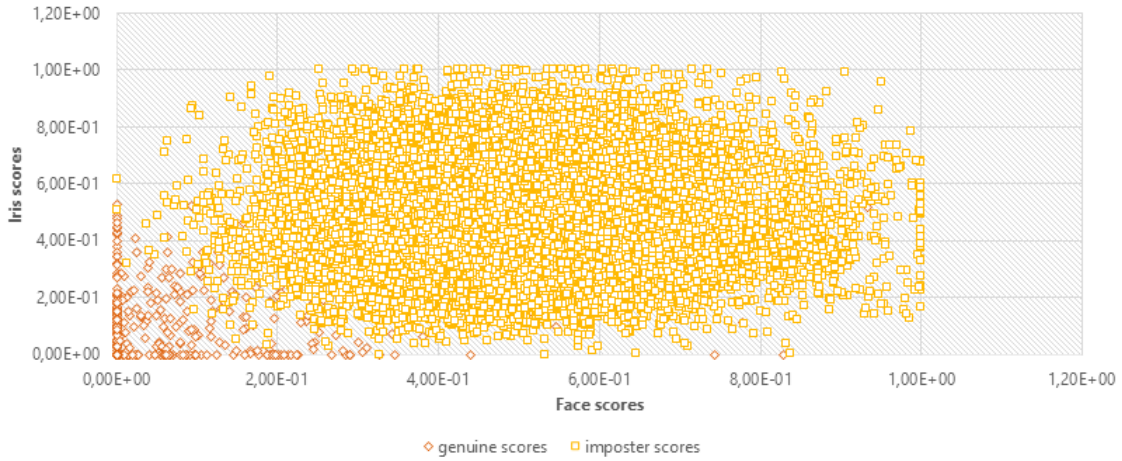


Figure 1.5: Score distribution of genuine and imposter scores from a face and iris matcher

- *Large population size*: Most biometric systems deployed in real-life environments involves large population and the FAR of the system increases as the database grows. This means the accuracy and throughput of the system is adversely affected when population size is large.

- *Large class imbalance*: From equation 1.4.2 it is seen that number imposter scores greatly encompasses the genuine scores as number of users increase (Mehrotra, 2014). This means that during classification, the impostor class becomes highly favoured as compared to the genuine class. This decreases the accuracy of prediction for the genuine class, which in turn increase the FAR of the system. Figure 1.5 shows the distribution of genuine and impostor scores for face and iris matcher showing how the genuine scores are under-represented as compared to the impostor scores.

The research questions examined in this research are outlined as:

- How can multi-biometric systems be used to reduce effects of noisy data, illumination, inter/intra-class variance that affect the accuracy of single biometric systems?

- Which viable techniques are capable of training biometric data with large class imbalance?

1.6 Research objectives

The above research challenges are addressed to enhance the performance of multi-biometric systems. The research objectives are:

- Develop a fusion technique that will reduce the effects of non-universality, inter and intra-class variance experienced with unimodal systems.
- Develop a fusion technique that reduces the effect of large class imbalance between the genuine and impostor scores in the biometric data.

1.7 Thesis contribution

This thesis seeks to tackle two issues from the ones outlined in the previous section. First, the issue faced by single biometric systems is addressed using a hybrid fusion methodology, while serial fusion strategy is proposed to address the issue of large class imbalance and population size. Major contributions of this research include:

- A hybrid fusion methodology that combines three classifiers built on feature and score level of fusion using a decision level fusion rule is proposed. first, feature fusion is performed using five standard global and local feature extractors for face and iris modalities separately. Secondly, weighted score fusion is performed between a global feature extractor for face and local feature extractor for iris. These three classifiers are combined to determine the claimed identity using a majority voting rule.
- In order to solve the problem of large population and imbalance class a serial fusion with BPSO and iRVM as a classifier is proposed. First the use of serial fusion reduces the number of imposter scores by selecting top matchers from the face unimodal system while iRVM is used to train the biometric data as it arrives in batches. BPSO has been used to select optimal mix of local and global feature extractors for improving the rank k recognition rate of face biometric system.

1.8 Thesis outline

This thesis is made up six chapters whose layout is given below:

Chapter 2 : Information fusion in multi-biometric systems

This chapter provides a literature review of biometric systems. These are split based fusion techniques using parallel and serial architecture. First, research works based on fusion techniques in parallel mode are explored according to the following branches: (a) score level fusion (b) feature level fusion (c) rank level fusion (d) decision level and (e) fusion based on Nature-Inspired (NI) algorithms. Then finally review works on serial fusion to conclude the chapter.

Chapter 3 : Feature Extraction Algorithms

The five standard feature extraction algorithms are discussed in this chapter. The algorithms have been divided into local and global methods. These methods extract feature vectors describing the texture and shape information from the face and iris images.

Chapter 4 : Hybrid fusion at feature, score and decision level

This chapter proposes the hybrid fusion methodology that combines three levels of fusion to determine the claimed identity. Herein, the formation of the three classifiers is discussed with experimental results showing the performance of the proposed scheme against other fusion methods in literature.

Chapter 5 : Serial fusion using BPSO and *i*RVM

The proposed serial fusion scheme using BPSO and *i*RVM are the point of discussion in this chapter. First, the advantages of serial fusion over parallel fusion is highlighted and the architecture along with algorithms utilised are discussed. Finally, simulation results are presented and discussed.

Chapter 6 : Conclusion and Future Work

Chapter 6 is the conclusion. In this chapter I review the research outputs, that is the achievements and limitations of the study. Recommendations for future research are presented.

Chapter 2

Information Fusion in Multi-Biometric Systems

2.1 Introduction

The success of multi-biometric systems depends on its design approach and its design depends on the type of fusion scheme employed. The thrust in this section is to discuss multi-biometric systems in literature based on parallel and serial architecture. Under the parallel architecture, works implemented with the following fusion schemes (a) score level fusion (b) feature level fusion (c) sensor level fusion (d) rank level fusion (e) decision level fusion (f) hybrid fusion and (g) fusion based on NI algorithms are considered as shown in Figure 2.1. Mean while works under serial architecture are also reviewed. Finally results obtained from the research both under parallel and serial architecture are reported.

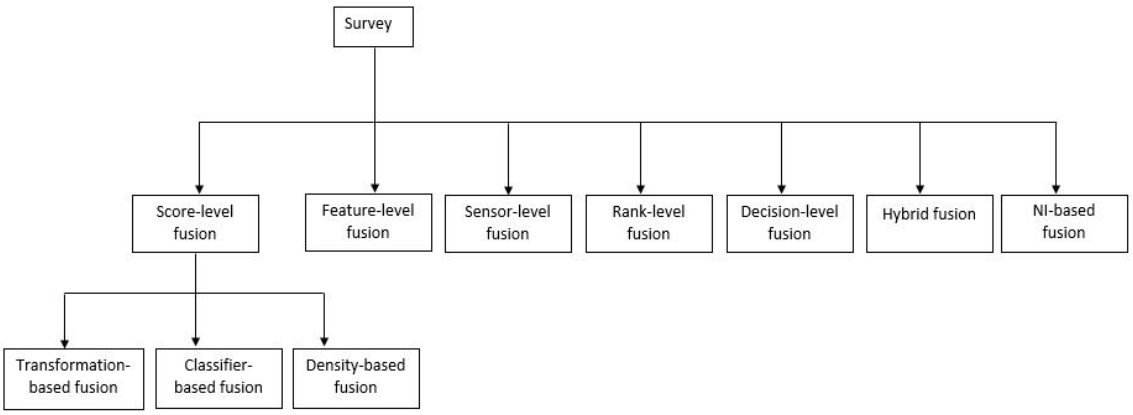


Figure 2.1: Categories of information fusion for multi-biometrics systems in parallel mode

2.2 Fusion in parallel mode

In this fusion mode, all modalities used for biometric fusion are processed simultaneously and combined at a later fusion stage. Figure 2.2 shows the general architecture for fusion in parallel mode. According Sanderson and Paliwal (2002) this fusion mode can be classified into two groups namely:

- *fusion before matching*
- *fusion after matching*

Such classification is necessary because the amount of information present in biometric templates reduces as it progresses to the matching module as shown in Figure 2.3. As the names imply, *fusion before matching* occurs when the templates are combined before getting to the matching module, while *fusion after matching* is quite the opposite. The first group includes fusion at *sensor level* and *feature level*, while the second group includes fusion at *score level*, *rank level* and *decision level*. Below, a description of studies conducted under these groups of fusion schemes are discussed.

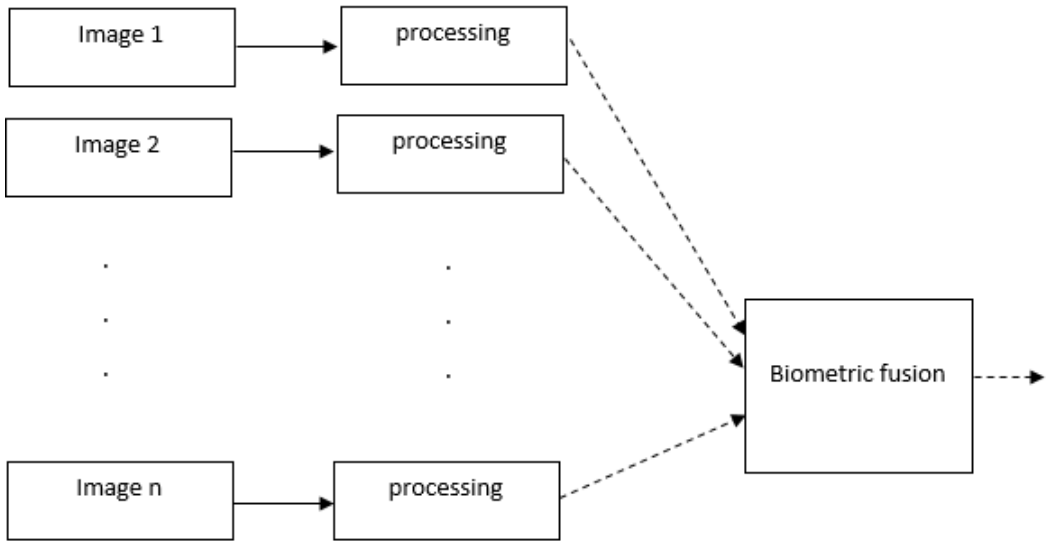


Figure 2.2: Parallel fusion architecture

2.2.1 Score-level fusion

In this section, research based on the different categories of score level of fusion are discussed. Work on transformation-based methods are reviewed first, ensued by the

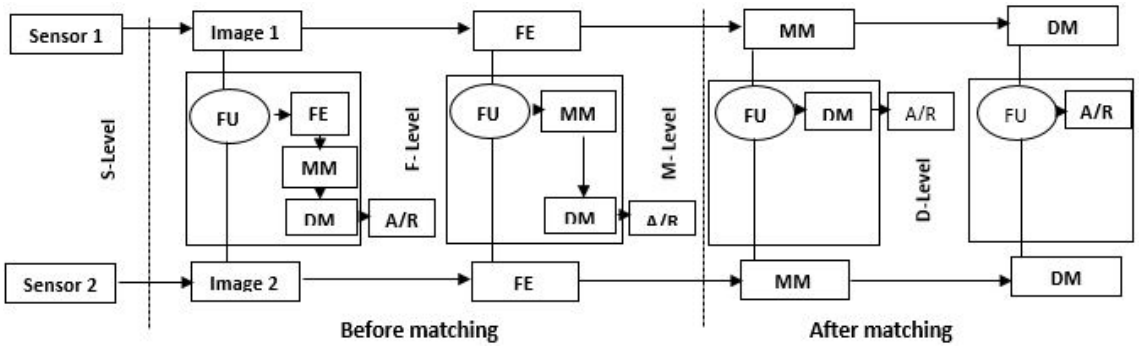


Figure 2.3: Types of fusion in Multi-biometric systems, Key: S-Level = Sensor- level fusion, F-Level = Feature level fusion, M-Level = Matching score fusion, D-Level = Decision level fusion; FU = Fusion, MM = Matching Module, DM = Decision Module, A/R = Accept/Reject

density-based methods and consequently classifier-based methods.

2.2.1.1 Transformation-based methods

In this method, matching scores evaluated from each modality are rescaled to the same domain usually in the interval $[0, 1]$ using normalization techniques and then combined with fusion rules. The commonly used fusion rules in literature include sum, product, minimum, maximum, mean and median rule which are discussed below (Fakhar et al., 2011; Eskandari et al., 2014; Connaughton et al., 2011).

Let $s = [s_1, s_2, \cdots, s_J]$ denote the score vectors for each uni-modal system J for the i^{th} sample. Where $i = [1, 2, \cdots, N]$ and N represents the number of probe samples, while $k = [1, 2, \cdots, J]$ uni-modal systems. Let m_i represent the combined match score belonging to the c_i membership class for the i^{th} sample. Therefore, fusion rules can be defined as follows:

(a) **Sum-rule**

Calculates the sum of the similarity scores from each modality. It performs better when the confidence levels obtained from all modalities are similar or with less variations (Duin, 2002; Eskandari et al., 2014).

$$m = \sum_{k=1}^{J} s_k \tag{2.2.1}$$

(b) **Product-rule**

Takes the product of the similarity scores obtained from each modality. This rule performs better if the modalities are un-correlated, as it assumes reliable confidence estimates from each modality. It however, fails when at least one of the match scores is equal to zero (Duin, 2002).

$$m = \prod_{k=1}^{J} s_k \tag{2.2.2}$$

(c) **Min-rule**

Picks the minimum matching score obtained from each modality as the best match for the probe image.

$$m = min_{k=1}^{J} s_k \tag{2.2.3}$$

(d) **Max-rule**

Picks the maximum matching score obtained from each modality as the best match for the probe image. However, this fails if any of the modalities are over trained than others skewing the final decision to the outcome of a particular modality. (Duin, 2002; Connaughton et al., 2011).

$$m = max_{k=1}^{J} s_k \tag{2.2.4}$$

(e) **Weighted sum-rule**

This combines matching scores from different modalities based on their noise levels. It does this by assigning weights to each scores obtained from each unimodal system to make its final decision (Eskandari and Toygar, 2014). Therefore a modality with less noise level will have higher weight as compared to those with high noise levels.

$$\begin{aligned} m &= \sum_{k=1}^{J} w_k s_k \\ where \sum_{k=1}^{J} w_k &= 1 \end{aligned} \tag{2.2.5}$$

(f) **Triangular-norms**

Triangular norms is a commutative, associative, monotonous operation $T : [0,1] \times [0,1] \rightarrow [0,1]$. It has been applied in fuzzy logic, fuzzy control systems

and artificial intelligence systems (Yang et al., 2012). There are three t-norms equations which can be divided into two groups and they are described below (Yang et al., 2012):

- non-parametric

$$\begin{aligned} Einsteins\ product : T(s_1, s_2) &= \frac{s_1 s_2}{(2 - (s_1 + s_2 - s_1 s_2)} \\ Hamacher : T(s_1, s_2) &= \frac{s_1 s_2}{(s_1 + s_2 - s_i s_2)} \end{aligned} \tag{2.2.6}$$

- parametric

$$Frank : T(s_i, s_{i+1}) = \begin{cases} T_M(s_1, s_2) = min(s_1, s_2)\ if\ p = 0 \\ T_P(s_1, s_2) = s_1, s_2\ if\ p = 1 \\ T_L(s_1, s_2) = max(s_1 + s_2 - 1, 0)\ if\ p = \inf \\ \log_p(1 + \frac{(P^{s_1} - 1)(P^{s_2} - 1)}{(p-1)}\ otherwise \end{cases} \tag{2.2.7}$$

Where p is a parameter in the range $[0, \infty]$ that determines the fusion operation to perform as seen above. With this is method any two scores s_i and s_{i+1} can be fused by performing $T(s_{i+1}, s_i)$ if more than two scores exist, a further fusion operation $T(s_{i+2}, T(s_i, s_{i+1}))$ performed with third score s_{i+2}. Therefore final fused score s_f is obtained from equation 2.2.8

$$s_f = T(s_{i+2}, T(s_i, s_{i+1})) \tag{2.2.8}$$

(g) **Score-normalization methods**

Prior to applying the above fusion rules, normalization techniques are applied to the matching scores obtained from each uni-modal system to rescale them to a domain $[0, 1]$ Jain et al. (2007). This is because matching scores from different modalities might not be similar in terms of their score distributions. For example a matcher might produce a similarity score, while another may produce a dissimilarity score. Furthermore, the match scores may not be in the same range e.g a face system producing match scores in range $[-1, 1]$ and finger-print gives match scores in range $[50, 100]$. Combining this systems using any of the fusion rules mentioned will not produce accurate results, hence the need for normalization. A couple of normalization techniques used in literature include:

- Min-Max method

$$s'_k = \frac{s_k - Min}{Max - Min}, \quad where\ k = 0, 1 \cdots k \tag{2.2.9}$$

Where Min and Max are minimum and maximum value of the matching score vector obtained from each modality. s'_k and s_k are the normalized score and corresponding matching score (Wang et al., 2011).

- Tanh method

$$s'_k = \left\{ \tanh\left(0.01\left(\frac{S_k - \mu_{gh}}{\sigma_{gh}}\right)\right) + 1 \right\} \tag{2.2.10}$$

where μ_{gh} and σ_{gh} represent the mean and standard deviation of the set of matching scores (Jain et al., 2005).

- Z-score

$$s'_k = \frac{s_k - \mu}{\sigma} \tag{2.2.11}$$

where μ and σ represent the mean and standard deviation of the set of matching scores (Aly et al., 2013).

- Decimal scaling

This can be applied to individual scores from different modalites in different logarithmic scale (Jain et al., 2005).

$$s'_k = \frac{s_k}{10^n} \tag{2.2.12}$$

where $n = \log_{10} max(s)$

- Median Absolute Deviation (MAD)

$$s'_k = \frac{s_k - median}{MAD} \tag{2.2.13}$$

where $MAD = median(|s_k - median|)$ (Ross and Govindarajan, 2005)

- Double sigmoid function

$$s'_k = \begin{cases} \frac{1}{1+\exp(-2((s_k - t)/r_1))} & if\ s_k < t \\ \frac{1}{1+\exp(-2((s_k - t)/r_2))} & otherwise \end{cases} \tag{2.2.14}$$

where t is the system threshold, r_1 and r_2 denote the extreme regions in which the double sigmoid function is linear (Jain et al., 2005).

These fusion rules together with score normalization techniques have been applied

for biometric fusion in literature, which will be explored during the course of this review. Ramli et al. (2011), studied the performance of sum-rule and weighted-sum rule on a multi-instance and multi-modal system. The multi-instance system was a combination of different verbal samples of the same individual for voice recognition, while multimodal was constructed by using both voice and face samples. SVM (Support Vector Machines) was used as the matching algorithm for both voice and face uni-modal systems, while matching scores obtained were combined using both sum and weighted sum rule. Results obtained from their experiments revealed that weighted sum rule outperformed sum rule for both multi-instance and multi-modal system. In another study, Connaughton et al. (2011) researched a multi-modal system of face and iris templates using a multi-sensor approach. The videos obtained from the three different IOM (Iris On the Move) sensors were stitched together to obtain a single unit. Both Viola Jones and modified Daugman's algorithm were used to extract features from the respective biometric templates (Viola and Jones, 2004; Daugman, 2004). Match scores from each modality were normalized and combined using weighted sum rule.

The first attempt to merge face and iris traits using an efficient feature extraction algorithm based on steerable pyramids (S-P) was investigated by (Fakhar et al., 2011). The S-P bands captured intrinsic geometric frameworks of face and Iris to compute the feature parameters (mean variance, energy or entropy) from each modality. The city block distance was used to compute matching scores of face and iris images obtained from FERET (The Face Recognition Technology) and CASIA (Chinese Academy of Sciences' Institute of Automation) biometric database after which both biometric templates were combined using the sum rule score fusion technique pre-processed by Z-score and Min-Max normalization (Phillips et al., 1998; Tieniu and Zhenan, 2010). Hanmandlu et al. (2011) proposed the use of triangular-norms for biometric score fusion for palm-print, hand-vein and hand geometry traits. Gabor wavelets were used to extract features for the first two traits while an independent component analysis was used for the last trait. Two sets of databases were used to validate this scheme, a self-acquired and chimeric dataset constructed from PolyU (Palmprint, hand vein and hand-geometry) database. The authors were able to show that the method performed better that min, mean and sum rule fusion scheme.

Yang et al. (2012) presented a multi-instance finger-vein recognition system using LBPH (Local Binary Patten Histogram) to extract features from two fingers of the same individual. Both sum rule and triangular-norms were used to combine match scores of each finger-vein instance. Results obtained by the authors showed that

triangular-norms rule performed better than sum-rule, as it considers the uncertainty in the relationship of different modalities. Similarly, Wang et al. (2013a) proposed a novel fusion scheme by merging dual iris, visible and thermal face images. The 1D Log-Gabor filter was used to extract features from the dual iris images while CGJD (Complex Gabor Jet Descriptor) was used to improve the feature representation of visible and thermal face images. The match scores were fused using a variant of triangular-norms to obtain the final decision. The results were validated using CASIA Iris-thousand and NVIE (Natural Visible and Infrared facial Expression) visible and thermal face database (Wang et al., 2010). The EER's obtained from the proposed scheme outperformed other fusion schemes in literature.

Sim et al. (2014) employed weighted score fusion to fuse scores from face and iris traits based on their weights availability. While, Principal component analysis and neural networks were used for feature extraction for face and iris template. Their results were validated using a self-acquired dataset UTMIFM (Universiti Teknologi Malaysia Iris and Face Multimodal), UBIRISv2 (University of Beira IRIS version 2) and ORL biometric databases (Proenca et al., 2010; Ahonen et al., 2004; Tan et al., 2010). The author's findings showed that the recognition rate obtained from the self-acquired dataset performed better than the chimeric dataset constructed using UBIRISv2 and ORL. A new approach for a multi-modal system of finger vein and geometry was explored by (Asaari et al., 2014). The authors proposed a new feature representation and matching algorithm for finger geometry and vein respectively. First, a BLPOC (Band Limited Phase Only Correlation) algorithm invariant to noise and occlusion was used to evaluate the match scores of the finger-vein images. While new geometric features for finger geometry images were generated by combining finger width with the CCD (Centeroid Contour Distance). Weighted sum rule was applied as the fusion scheme for combining match scores. Results of the author's experiment were validated using a self-acquired database of 123 users, which showed that the proposed scheme was able to attain an improved performance rate and processing time as compared to other methods in the literature.

Eskandari and Toygar (2014) proposed a face-iris multi-modal biometrics system by integrating features from local and global feature extraction algorithms for each face and iris sub-system. Matching scores were rescaled using Tanh method and fused using weighted sum rule. The proposed scheme was validated using ORL, FERET, BANCA and CASIA datasets and results achieved showed improved recognition accuracies as compared to its unimodal and other multi-modal systems in literature (Bailly-Bailliere et al., 2003). Recently, Yong et al. (2015) proposed a novel method

by merging both the left and right palm-print for human recognition. Three matching scores were generated as the first two consisted of scores obtained from both the left and right palm-print subsystem while the third score was generated by computing the similarity between left and right palm-print of the same individual using a specialized algorithm developed by the authors. Finally, the three matching scores were combined using weighted sum rule to calculate final fused score.

Farmanbar and Toygar (2015a) presented a multi-modal system that employed face and palm-print biometric. Local Binary Pattern was used to obtain features from the face and palm-print biometric while Backtracking Search Algorithm (BSA) was used to evaluate the optimal values for each feature that produced the best recognition rate. Match scores obtained from the selected feature vectors were combined at score level to determine the claimed identity. Experimental results demonstrated significant improvement for the proposed scheme when compared to the uni-modal systems.

2.2.1.2 Classifier based methods

In this approach, match scores retrieved from each modality are concatenated into a $d-$dimensional vector, the classifier takes the score vector as an input to determine if the user belongs to the genuine or impostor class. A description of some of these classifiers that have been employed in literature are given below:

(a) **Support Vector Machines**

SVM is a popular machine learning algorithm that creates a hyperplane in a multi-dimensional space (Vatsa et al., 2008). The success of SVM lies in its ability to perform both linear or non-linear classification. In linear form, the goal of SVM is to create a straight-line boundary that accurately separates the two classes. Linear form SVM can also be extended to its non-linear form through the application of the appropriate kernel function for dimension reduction (Woo and Kim, 2006). Consider a set of scores x with class membership $y_i \in \{-1, 1\}$, SVM represents a hyperplane that separates the two classes as:

- linear form

$$\vec{w}.\vec{x} + b = 0 \tag{2.2.15}$$

- non-linear form

$$\vec{w}.\varphi(\vec{x}) + b = 0 \tag{2.2.16}$$

where $\vec{w}$ represents the vector of weights and $\varphi(\vec{x})$ represents a kernel function.

(b) **Relevance Vector Machines**

RVM is a model with an identical form to SVM that employs Bayesian inference (Tipping, 2001; Tran and Le, 2015). One of the most preferred feature of RVM is that, it is capable of making probabilistic predictions from the learned model using relatively fewer kernel functions as compared to SVM. Generally the RVM model is described as follows:

$$y(x, w) = w^T \phi(x) \tag{2.2.17}$$

where x is the input vector, w is the vector of weights, $\phi(x)$ is a set of basis functions and y is the output.

(c) **Artificial Neutral Networks**

ANN are artificial Intelligence systems that mimic the human nervous system. It is presented as a system of interconnected neutrons which exchange information with one another. it is made up of an input layer which takes the training data as input, hidden layer in which the learned model is developed from the training data and an output layer which provides prediction value in the case of regression or prediction class in the case of classification (Cristianini and Shawe, 2000).

Popular machine learning classifiers have been adapted in the field of biometrics with improved performances recorded, a description of some works in literature that explored these methods follows.

Wang and Han (2009) presented a score fusion methodology using support vector machines. They achieved this by concatenating scores retrieved from both face and iris modalities and passed them as features to SVM to determine if the user was genuine or an impostor. Results obtained by the authors showed that equal error rate for the proposed scheme performed better than other fusion techniques like sum, product and fisher rule. Similarly, the study of four algorithms for biometric score fusion was conducted on XM2VTS face-voice database (Messer et al., 1999; Damousis and Argyropoulos, 2012). The algorithms studied by the authors include GMM (Gaussian Mixture Models), SVM, ANN and FES (Fuzzy Expert Systems) with SVM (lowest half total error rate) performing better than other three algorithms and other fusion methods in the literature.

Mehrotra et al. (2012) presented a biometric matching score fusion using relevance vector machines (RVM) for a multi-unit iris recognition system. They used RVM to estimate posterior probabilities for both left and right iris scores. The final fusion was done by performing weighted score fusion to determine the probability of the prediction. Results obtained showed that RVM had better generalisation properties than SVM during classification. Eskandari et al. (2013) presented a new score fusion methodology using face and Iris biometrics. They used five local and global feature extractors to obtain scores from face and iris individually and passed the concatenated scores as feature vectors into an SVM classifier. ORL, BANCA, CASIA and UBRISv2 database were used to validate their scheme, which was better than the accuracy obtained for feature level fusion (Bailly-Bailliere et al., 2003).

Mehrotra (2014) also proposed iGRVM for multi-unit iris recognition. The authors introduced the concept of incremental learning and granular computing, in order to handle class imbalance in biometrics score set. The fusion methodology used was the same as that in (Mehrotra et al., 2012). The results obtained showed that the classifier was capable of training with a large population and high-class imbalance between the genuine and imposter scores.

2.2.1.3 Density based methods

The core of this method is based on performing statistical test with prior estimation of the probability density functions of the match scores. Let S represent a random variable that a score is obtained from a matcher. Then the probability distribution functions for the genuine and impostor scores $F_{gen}(s)$ (with $f_{gen}(s)$ density function) and $F_{imp}(s)$ (with $f_{imp}(s)$ density function) are defined as:

$$\begin{aligned} F_{gen}(s) &= P(S \leq s \mid S \; is \; genuine) \\ F_{imp}(s) &= P(S > s \mid S \; is \; an \; impostor) \end{aligned} \tag{2.2.18}$$

Two statistical hypothesis are defined namely H_0: meaning that score S represents an impostor and H_1 represents a genuine user. The probability of not accepting H_0 when H_1 is holds is known as the FAR, while the probability of not accepting H_1 when H_0 is holds is known as FRR (Nandakumar, 2008). Using Nerman-person

thoerem, a likelihood ratio test is performed as follows:

$$P(\Psi(S) = 1 \mid H_0) = \alpha$$
$$\Psi(S) = \begin{cases} 1, & when \ \frac{f_{gen}(s)}{f_{imp}(s)} > \eta \\ 0, & when \ \frac{f_{gen}(s)}{f_{imp}(s)} < \eta \end{cases} \tag{2.2.19}$$

where Ψ is the test, η is the system threshold and α is the level of test Ψ.

This approach is often preferred over fusion schemes under transformation and classifier based categories in that no score normalization technique is required. However to obtain improved performance with this approach the probability density functions of the genuine and impostor scores have to be estimated accurately (Mehrotra, 2014). Below a description of recent works based on these methods and their contribution to literature.

Nandakumar (2008) proposed a fusion technology based on Neyman-Pearson theorem for combining multiple biometric matchers. The likelihood ratio was used to maximize the GAR at any desired false acceptance rate. GMM estimated the probability density function of the genuine and impostor scores. The method was able to consistently achieve high recognition rate for the different biometric database without any parameter tuning. In another study, Vatsa et al. (2008) proposed a hybrid framework of likelihood ratio test and SVM for score level fusion of face biometrics. The probability density functions of the match scores were estimated by assuming they are Gaussian distributed, then the likelihood ratio was computed and used as an input for SVM. The results obtained by the authors showed that the proposed scheme achieved improved accuracy as compared to sum rule and SVM fusion method.

Qian and Veldhuis (2013) estimated the likelihood ratios of biometric scores of selected points in the individual ROC curves to construct the naive Bayes classifier. The use of the selected points in ROC diminished the overhead cost of the algorithm as compared to computing the density distribution function of the genuine and imposter scores.This fusion scheme performed better than the one based on SVM and GMM.

Table 2.1 shows the summary of results of literature reviewed above on score level fusion.

Table 2.1: Score level fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
Transformation-based methods					
2011	Ramli et al. (2011)	sum, weighted sum rule	face and voice	self acquired dataset	sum: EER:0.2778%, 2.0261% EER:0.0563%, 1.9904%
2011	Connaughton et al. (2011)	weighted sum rule	face and iris	self acquired dataset	Acc: 93.2%
2011	Fakhar et al. (2011)	sum rule	face and iris	FERET and CASIA	Acc: 99.17%
2011	Hanmandlu et al. (2011)	t-norms	palm-print, hand vein and hand-geometry	PolyU and IITD	GAR: 99.7%
2012	Yang et al. (2012)	sum-rule and t-norms	finger-veins	SDUMLA-HMT	EER: 1.42%, EER: 1.26%
2013	Wang et al. (2013a)	t-norms	dual iris, visible and thermal face	CASIA and NVIE	EER: 0.0289%
2014	Sim et al. (2014)	weighted sum rule	face and iris	self acquired, UBIRISv2 and ORL	GAR: 97%, GAR: 96%
2014	Asaari et al. (2014)	weighted sum rule	finger geometry and vein	Self acquired	EER: 1.78%
2014	Eskandari and Toygar (2014)	weighted sum rule	face and iris	ORL,FERET CASIA and UBIRISv2	EER: 0.5%
2015	Yong et al. (2015)	weighted sum rule	palm-print	PolyU and IITD	EER:0.53%
2015	Farmanbar and Toygar (2015a)	BSA and sum rule	face and palm-print	FERET and PolyU	Acc: 99.17%
Classifier-based methods					
2009	Wang and Han (2009)	SVM	face and iris	ORL and UBIRISv2	EER: 0.35%
2012	Damousis and Argyropoulos (2012)	SVM,GMM FES and ANN	face and voice	XM2VTS	HTER: 0.25%, HTER: 0.51% HTER: 0.45% HTER: 0.5%
2012	Mehrotra et al. (2012)	RVM	dual iris	CASIA	Acc: 98.81%
2013	Eskandari et al. (2013)	SVM	face and iris	ORL,FERET CASIA and UBIRISv2	Acc: 98.25%
2014	Mehrotra (2014)	*i*GRVM	dual iris	BATH and CASIA	GAR: 47.83%, GAR: 17.28%
Density-based methods					
2008	Nandakumar (2008)	likelihood ratio with GMM	Multimodal	NIST match score and XM2VTS	GAR: 99.1%, GAR: 98.7%
2008	Vatsa et al. (2008)	likelihood ratio with SVM	Face	Mixture	GAR: 94.98%
2013	Qian and Veldhuis (2013)	Naive Likelihood Ratio via ROC	Face	FRGC	EER: 1.75%

2.2.2 Feature-level fusion

In this section, works in literature that have employed fusion at feature level for consolidating evidence obtained from different modalities are discussed. Different strategies for combining features have been explored and they are discussed below:

Consider a multi-biometric system consisting of J modalities such that $k = 1, 2, \cdots, J$ and $i = 1, 2, \cdots, N$ where N is the number of probe samples for each k modality. Let X_{ik} is the feature vector of the i^{th} sample for k^{th} modality and Z_i be the resulting feature vector after fusion for i^{th} sample . Therefore Z_i can be obtained based following methods:

(a) **Feature concatenation**

The fused feature vector is obtained by simply combining features from each modality (Ross and Govindarajan, 2005; Kumari and Suma, 2014). Therefore Z_i is computed as:

$$Z_i = [X_{i1}, X_{i2}, \cdots, X_{ik}] \tag{2.2.20}$$

(b) **Feature update**

This method is used when features are of the same modality and obtained from the same feature extraction algorithm. The J number of features for each N samples are averaged to obtain the final fused vector Z_i (Jain et al., 2007)

$$Z_i = \frac{1}{J} \sum_{k=1}^{J} X_{ik} \tag{2.2.21}$$

(c) **Sum/Weighted sum rule**

In these methods, the final fused vector Z_i is obtained by performing an equal or weighted sum of feature vectors from each modality (Wang et al., 2011).

$$\begin{aligned} Z_i = \sum_{k=1}^{J} X_{ik} \; or \sum_{k=1}^{J} w_k X_{ik} \\ where \sum_{k=1}^{J} w_k = 1 \end{aligned} \tag{2.2.22}$$

(d) **Complex feature fusion**

This is often used when two modalities are to be combined. Here, the feature sets obtained from each modality is represented in the form of a complex

number. If the dimension of one feature vector is larger than other, the smaller feature vector is padded with zeros until its dimension equals its counterpart (Wang et al., 2013b). below describes the format for complex feature vector.

$$Z_i = wX_{i1} + (1-w)X_{i2}j \quad for\ k = 1, 2 \tag{2.2.23}$$

It should be noted that prior to performing feature level fusion, the process of normalisation discussed in section 2.2.1.1 will have to be carried out as individual feature values may exhibit significant variation in range and distribution (Ross and Govindarajan, 2005). Prior research where these feature level fusion techniques for biometric recognition have been employed are explored below.

(Ross and Govindarajan, 2005) presented a feature level scheme based on hand and face biometrics. Feature extraction was done by extracting unique geometric measurements from the hand biometric while LDA (Linear Discriminant Analysis) was used for face biometric. The two feature sets from each biometric were concatenated, followed by feature selection to optimise the feature set that will provide the best performance. Results showed that scheme proposed by authors performed better than fusion at match score using MSU (Michigan State University) multi-modal database (Ross and Govindarajan, 2005). Wang et al. (2011) proposed a face-iris multi-modal system using feature level fusion. Feature sets from both the face and iris biometrics were obtained using PCA (Principal Component Analysis) and Gabor filter respectively, followed by Z-score normalization to rescale feature set into the same domain. Three types of feature level fusion schemes were tested in the course of their experiment and these include concatenation, sum and weighted sum rule, with the feature concatenation method having the best performance. Ramachandra et al. (2012) presented a dual bimodal system based on face and fingerprint traits. Fingerprint images were pre-processed to find the ROI (Region Of Interest), followed by the DTCWT (Dual Tree Complex Wavelet Transform) algorithm to obtain features as the high and low-frequency components of the fingerprint ROI. Similarly, ROI for the face images were obtained and features were extracted as coefficients of frequency bands using Haar wavelets. Both features were merged to form a fused feature vector while the euclidean distance was used to evaluate the match score between the fused feature vector and the stored templates in the database. Experimental results obtained by authors showed the EER was lower for their proposed scheme compared to its unimodal systems.

A complex feature fusion strategy was proposed by Wang et al (Wang et al., 2013b) based on visible and thermal face biometric. 2DPCA and 2DLDA were applied for

extracting feature sets from both modalities and then combined using a weighted complex fusion method described above. Experimental results revealed that the proposed feature level scheme performed better than some fusion schemes in literature. A novel algorithm for fusing features using an SVM classifier was presented in (Gawande et al., 2013). Haar wavelet transform was used to obtain feature sets from both the iris and fingerprint modality, with Mahalanobis distance used to find top-match feature sets as compared to the probe sample. The values of the individual feature sets were normalised with the Tanh method to ensure equal contribution to fusion and combined using feature update for both iris and fingerprint biometrics. The fused feature vector was passed as an input to be trained by an SVM classifier. From the authors' simulation results, they were able to show that their proposed fusion scheme was able to obtain higher recognition accuracy and lower false rejection rate as compared to other existing approaches.

Chin et al. (2014) developed a biometric template protection technique for fingerprint and palm-print based on feature level fusion. The authors used Gabor filter to extract a feature set from both modalities and then merged to form a fused feature vector. The fused vector was then transformed into a binary template using random tilling and 2^N discretisation scheme. Results obtained by the authors showed that the proposed scheme had improved performance as compared to its uni-modal counterparts, while still providing template security. Muhammad et al. (2015) proposed a non-stationary feature fusion scheme based on face and palmprint images. Features from the individual modality were extracted using DCT algorithm to obtain local descriptors for both face and palmprint images. The authors combined features sets obtained from DCT (Discrete Cosine Transform) to form fused feature vector, which was trained using a GMM to obtain its probability density function. Experimental results obtained by the authors showed that the proposed technique outperformed existing feature, matching and decision level fusion schemes.

Table 2.2 shows the summary of results of literature reviewed above on feature level fusion.

2.2.3 Sensor-level fusion

This fusion scheme employs multiple sensors to capture single or multiple traits of the same individual which are consolidated to form a single image. At this level of fusion, the performance of the system is expected to be better than other levels fusion because of the availability of pixel information of the biometric image (Kisku

Table 2.2: Feature level fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
2005	Ross and Govindarajan (2005)	Concatenation	hand and face	MSU	ROC analysis
2011	Wang et al. (2011)	Concatenation, weighted sum and sum rule	face and iris	ORL, FERET and CASIA	EER: 1.94% EER: 2.5% EER: 4.22%
2012	Ramachandra et al. (2012)	Concatenation	face and fingerprint	PIDB and PODB	EER: 0.13%
2013	Wang et al. (2013b)	Complex fusion	face	NVIE	Acc: 97.38%
2013	Gawande et al. (2013)	Feature update	iris, and finger-print	CASIA and Self acquired	GAR: 94%
2014	Chin et al. (2014)	Concatenation	finger-print and palm-print	FVC, and PolyU	EER: 1.64%
2015	Muhammad et al. (2015)	Concatenation	face, and palm-print	ORL,FERET and palm-print	Acc: 99.7%

et al., 2010). Below is a multisensor biometric fusion technique used in literature:

(a) **Discrete Wavelet Transform (DWT)**

DWT decomposes an image repeatedly into different frequency levels, which contains transform values known as *wavelet coefficients*. At each level, it breaks down the image into frequency bands categorised as low-low, low-high, high-low, high-high bands (Kisku et al., 2010). The low-low bands represent the raw information about the image while the other three band show sharp changes in gradients such the edges, lines and boundaries. At the n^{th} level the image is decomposed as :

$$I_{n-1} = I_{LL_n} + I_{LH_n} + I_{HH_n} + I_{HL_n} \tag{2.2.24}$$

Most studies that have employed sensor-level fusion have its foundation based on wavelet transform, however, several variants of this method which include DWT with PSO (Particle Swarm Optimization) and DWT with Monotonic-Decreasing Graph (MDG) have been developed to improve its performance (Raghavendra et al., 2009; Kisku et al., 2009). The next set of paragraphs provides brief descriptions of some research work done on biometrics.

Raghavendra et al. (2009) proposed a novel sensor level fusion scheme based on face and palmprint images. Wavelet transform was used to decompose the images of the two modalities while PSO was used to select the optimal mix of wavelet coefficients

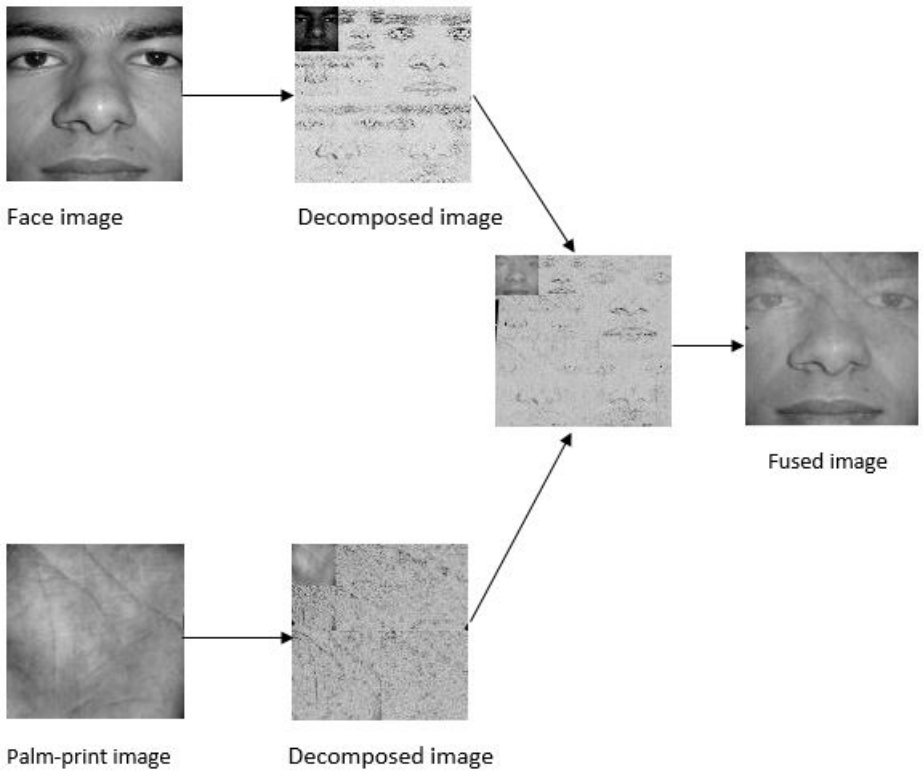


Figure 2.4: An example of DWT multi-sensor fusion for a face and palm-print image (Kisku et al., 2010)

from both face and palmprint images to produce the final fused image.The authors used Kernel Discriminant Analysis to extract features from the fused image and NNC (Nearest Neighbor Classifier) to obtain match scores. The authors demonstrated the efficacy of the proposed scheme through simulation, which revealed improved performance over match-score fusion. Another variant of wavelet transform was presented by Kisku et al. (2009) using the face and palmprint biometric. The authors approach was to combine multi-spectral images at different resolutions which was fused into a single image revealing a richer complementary image. After sensor fusion, features were extracted using SIFT (Scale Invariant Feature Transform) and recognition was done by matching the fused probe sample and fused template through a recursive decent tree transversal method. Experimental results showed that the proposed scheme outperformed its unimodal face and palmprint systems.

Raghavendra et al. (2011) presented a novel sensor fusion scheme based on PSO using visible and infra-red face images. PSO was used to pick the best mix of weights to perform a weighted linear combination of wavelet coefficients from both types of face images. The authors again used PSO to obtain an optimal collection of features from the visible and infra-red image. Results obtained from experiments by

Table 2.3: Sensor level fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
2009	Raghavendra et al. (2009)	DWT with PSO	face and palm-print	FRGC and PolyU	GAR: 94.26%
2009	Kisku et al. (2009)	DWT with MDG	face and palm-print	FRGC and PolyU	Acc : 98.19%
2011	Raghavendra et al. (2011)	DWT with PSO	face	FRGC and IRVI	Acc: 99.7%

the authors revealed that the proposed scheme was able to show stability to changes in environmental conditions and with better performance than that of score level technique.

Table 2.3 shows the summary of results of works reviewed above on sensor level fusion

2.2.4 Rank-level fusion

This level of fusion involves assigning ranks to every registered template (high rank indicates a good match) and final rank is obtained by consolidating ranks from different modalities. different techniques applied for fusion at rank level are described below:

(a) **Highest rank**

The fused rank of the user is computed as the lowest rank obtained from individual modalities (Monwar et al., 2013).

(b) **Borda count**

Is an election method used to rank options or candidates according to the best match criteria. In a multi-modal system, it is used to select the best match of scores from uni-modal systems by sorting all scores and giving the highest rank to the best match (Radha and Kavitha, 2012).

(c) **Weighted Borda count**

In this Borda count is extended by assigning weights to each matchers according to their performance. These weights are computed by using logistic regression (Kumar and Shekhar, 2011).

(d) **Bucklin majority voting**

This is a voting scheme in which a candidate with the highest median rating is chosen as the winner for the first choice votes. If otherwise the second choices are added to first choice votes and the candidate with the highest vote is taken as the winner and then again the procedure is repeated (Kumar and Shekhar, 2011) .

(e) **Nonlinear weighted rank**

This is a combination of ranks from different matcher that are non-linear weighted.

The following set of paragraphs describes works in literature that have explored the above methods. Kumar and Shekhar (2011) proposed a new non-linear weighted rank fusion scheme and presented a comparative study of this fusion scheme against other existing rank level fusion scheme like the Borda count, weighted Borda count, highest rank and Bucklin majority voting. Their multi-biometric system was based on multiple representation of palmprint matching scores and results obtained from their experiments revealed that the proposed non-linear weighted method outperformed other existing rank fusion schemes studied by the authors. Radha and Kavitha (2012) developed a multi-modal system which employed fingerprint and iris biometric. Features from each modality were extracted using FLD (Fisher Linear Discriminant) and combined using weighted Borda count rank fusion scheme for the determining the claimed identity. Simulation results obtained by the authors revealed that the performance of the multi-modal system outperformed uni-modal counter-parts.

A study of biometric fusion at the rank level, between facial thermograms and ear, was carried out by (Kumar et al., 2012). A self-acquired database was constructed by the authors by capturing images of the first modality with the aid of an infra-red camera and the second modality was captured using an ordinary digital camera. Features set obtained from each modality was extracted using Haar wavelets and SIFT respectively. Weighted Borda count rank fusion method was employed for consolidating ranks obtained both modalities and results indicated that the proposed system provided better performance than unimodal systems. Monwar et al. (2013) proposed the use of ocular biometrics (iris and retina scan) to tackle occlusion and illumination encountered during iris recognition. The authors studied the performance of rank level fusion on three ocular matching algorithms namely, PDM (Probability Distribution Model), modified SIFT (m-SIFT) and GOH (Gradient Oriented Gradient). The authors also demonstrated from the results they obtained,

that the application of existing rank fusion methods could lead to an improved performance rate.

The first attempt to apply fusion at rank level on cancellable biometrics was done by (Paul and Gavrilova, 2014). Multiple random projections were generated from the feature sets of each modality and were stored as templates in the database, then rank level fusion was applied to combine ranks obtained from each modality to determine the claimed identity during recognition. The authors presented a performance analysis of their proposed scheme with improved performance rate over unimodal systems. In another study, Talebi and Gavrilova (2015) proposed a novel reinforcement scheme for rank fusions based on frontal face, profiles face and ear images. The authors presented the rank-reinforcement approach using prior probability distribution of templates in the training data. Before performing rank fusion for recognition, the prior probability distribution is used to improve the rank list of each biometric matcher. The authors demonstrated that their proposed scheme had the capbility to enhance the performance rates of existing rank fusion schemes.

Table 2.4 shows the summary of results of works reviewed above on rank level fusion.

Table 2.4: Rank level fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
2011	Kumar and Shekhar (2011)	Highest rank, Borda count, Weighted Borda count, Bucklin majority voting, and nonlinear weighted rank	palm-print	NIST BSSR1	Acc: 100% Acc: 94.97% Acc: 96.32% Acc: 99.81% Acc: 100%
2012	Radha and Kavitha (2012)	Weighted Borda count	finger-print and iris	FVC2000 and CUHK	ROC analysis
2012	Kumar et al. (2012)	Weighted Borda count	face and ear	Self acquired	GAR: 98%
2013	Monwar et al. (2013)	Highest rank, Borda count and plural voting	Ocular	FOCS	ROC analysis
2014	Paul and Gavrilova (2014)	Border count	face and ear	FERET, VIDTIMIT, AT& T and USTB	Acc: 84%
2015	Talebi and Gavrilova (2015)	Rank-reinforcement	face and ear	FERET and USTB	ROC analysis

2.2.5 Decision-level fusion

In this level of fusion, each modality performs its own feature extraction and matching process, after which the outputs are combined to obtain the final decision. This fusion scheme together with score level fusion have generated most interest in the academia as less information is conveyed during the fusion process (only class labels are combined) (Jain et al., 2004). Different methods have been used to combine class labels for decision fusion, a brief description is given below:

(a) **Boolean operation**

Set of "AND" or "OR" based rules are used to make the final decision considering output from different classifiers (Tao and Veldhuis, 2009).

(b) **Majority voting**

With this technique, the correct class label with the highest median rating is selected as the claimed identity. In the case of a tie, the correct class label with the highest score is chosen as the claimed identity. (Islam, 2014).

(c) **Average voting**

Each classifier computes the confidence average for every class and the class with the highest value is selected as the claimed identity (Islam, 2014; Fridman et al., 2015). See equation 2.2.25

$$Q(x) = \underset{j=1}{\overset{N}{\operatorname{argmax}}} \left(\frac{1}{k} \sum_{i=1}^{k} y_{ij}(x) \right) \tag{2.2.25}$$

where N represents the number classes and y_{ij} represents the output of the i^{th} classifier for the j^{th} class.

(d) **Maximum voting**

In this technique the class with the highest overall score is selected as the claimed identity (Islam, 2014; Wanas, 2003). Where N represents the number classes and y_i represents the output of the i^{th} classifier and x is the input.

$$Q(x) = \underset{i}{\overset{k}{\operatorname{argmax}}}\, y_i(x) \tag{2.2.26}$$

(e) **Nash voting**

Each classifier assigns a number in the interval $[0, 1]$ to each candidate and the compares the product for all classifier values for each candidate. The highest value is the a winner (Wanas, 2003).Where N represents the number classes, y_{ij} represents the output of the i^{th} classifier for the j^{th} class and x is the input.

$$Q(x) = \underset{j=1}{\overset{N}{\operatorname{argmax}}} \prod_{i}^{k} y_{ij} \tag{2.2.27}$$

In the following paragraphs, a brief description of some studies that have employed the above voting schemes is given. Marcialis and Roli (2006) presented a multi-algorithm face recognition system using PCA and LDA. The predicted class labels obtained from both algorithms were combined by fusion at decision level to provide the final output. The average voting scheme was used as a fusion method, as simulation results revealed that the proposed recognition rate was comparable to best face matchers in literature. Veeramachaneni et al. (2008) designed a decision fusion scheme incorporating Likelihood Ratio Test (LRT) and Chair Varshney rule (CVR) for closely related classifiers. Both LRT and CVR was used to find the optimal threshold point and decision rule by minimizing the system error. Their scheme was validated on a biometrics score dataset revealing the importance of adding correlation structure in building classifiers for multi-biometric systems.

Tao and Veldhuis (2009) proposed a optimised "AND" and "OR" rule-based decision scheme. They showed that optimising the threshold values of the classifiers of individual modalities provided substantial improvements to the performance of the fusion system by balancing matching scores from individual classifiers. The benefit of such is that the matching score normalisation process performed in other fusion schemes will not be required, thus reducing the risk of dropping the performance of component classifiers considering significant differences in their individual performance. Experimental results showed improvements over original classifiers that were fused, with results comparable with other conventional fusion schemes.

Table 2.5 shows the summary of results of works reviewed above on decision level fusion.

2.2.6 Hybrid fusion

This fusion method combines two or more levels of fusion in order to determine the claimed identity. In general, hybrid fusion multi-biometric systems are built

Table 2.5: Decision level fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
2006	Marcialis and Roli (2006)	Average voting	face	Yale and ORL	Acc: 97.3%
2008	Veeramachaneni et al. (2008)	LRT and CVR	face and finger-print	BSSR1	EER : 0.0847%
2009	Tao and Veldhuis (2009)	"AND" and "OR" rule	face	FRGC	EER: 0.015% EER:0.003%

by combining two or more classifiers based on any of the fusion scheme discussed above. Architecture of hybrid fusion schemes that have been studied in literature are described below:

(a) **Feature-score hybrid** Here more than one multiple feature extraction algorithms are applied on each individual modality to extract features, which are combined to form a fused feature vector. Matching scores obtained are combined using any of the schemes discussed in section 2.2.1 (Farmanbar and Toygar, 2015b). Figure 2.5 shows a general architecture for this type of hybrid fusion.

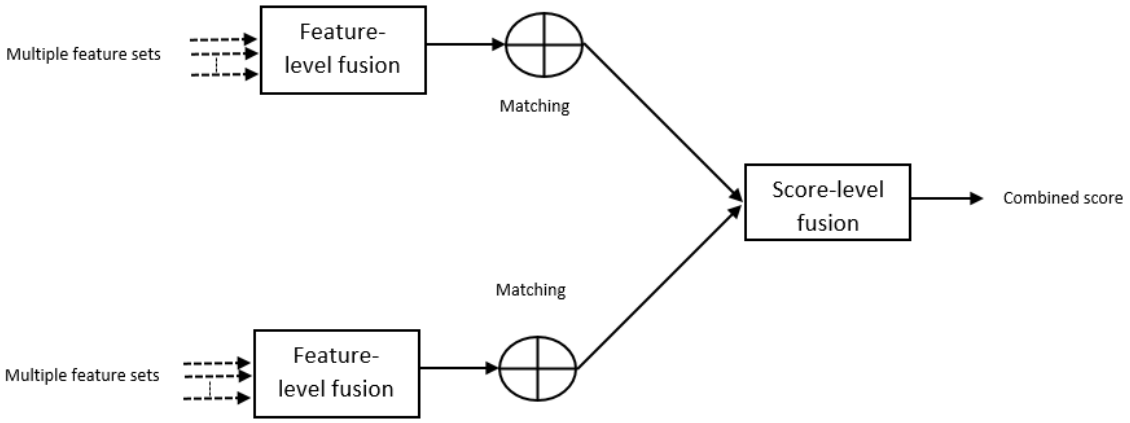


Figure 2.5: Feature-score fusion hybrid

(b) **Score-decision hybrid**

In this method, multiple score sets are generated for each modality either from multiple normalisation techniques or feature extraction algorithms. The match scores are fused and the claimed identity for each modality is determined. The final output is based on combining the class labels using any decision level fusion rule discussed in section 2.2.5. Figure 2.6 shows a general architecture for this type of hybrid fusion.

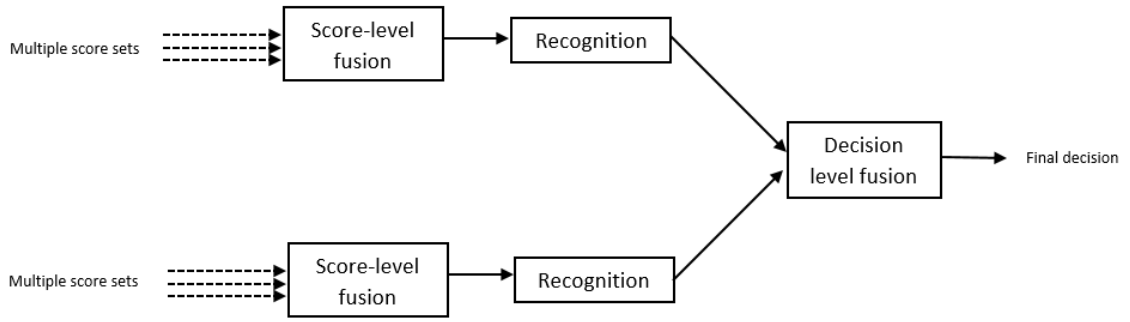


Figure 2.6: Score-decision fusion hybrid

(c) **Feature-decision hybrid**

This method is similar to the feature-score hybrid up until the matching stage, as the claimed identity is determined individually and the class labels are combined using a voting scheme or boolean operation to make the final decision as shown in 2.7.

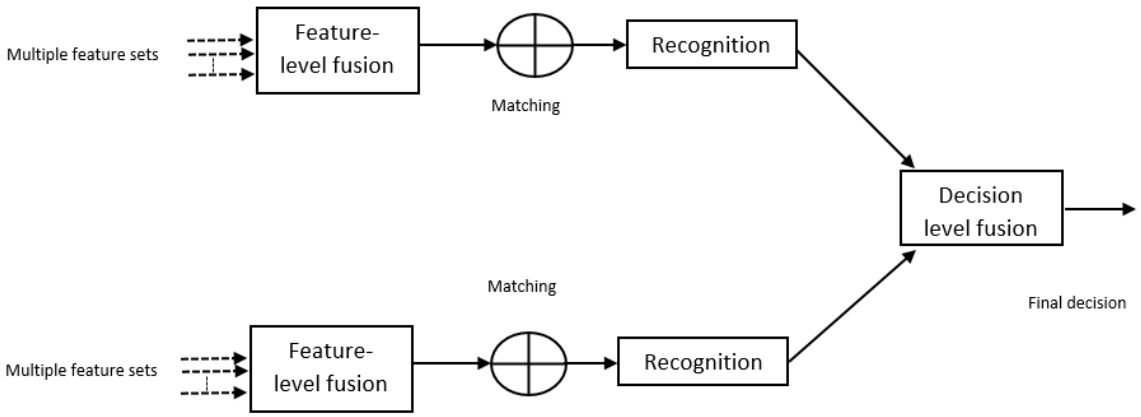


Figure 2.7: Feature-decision fusion hybrid

Below are recent works that have employed the above hybrid fusion techniques. Tao and Veldhuis (2008) proposed a score-decision hybrid fusion method. The authors enhanced the framework by providing an adaptability feature of switching between the two modes of fusion. The Receiver Operating Curves from components of matching scores were generated, with the optimal operating combined with "AND" or "OR" rule to provide the final output (Kekre et al., 2011). The authors presented a hybrid multi-modal system of face and iris traits using multiple feature extraction algorithms. Multiple feature extraction algorithms including 1D transform of row & column mean, Kekere wavelet and Kekere's Fast Codebook Generation (KFCG) were applied on the iris biometric, while Kekere's wavelet was used to extract features for the face biometric. The features obtained from the iris biometric were

combined together form the first classifier while features from the face biometric were combined with the fused iris feature vector to form the second classifier. The third classifier being the feature set from the face biometric. KNN algorithm was used to determine the class labels for finding the claimed identity for each classifier constructed. Finally the class labels were combined at decision level to obtain the final output.The proposed scheme outperformed its uni-modal counterparts.

The first attempt, of presenting a multi-biometric system that combined 3D templates of the ear and face at feature and score level was proposed by (Islam et al., 2013). Feature level fusion was performed by pairing face features with the most similar ear features while scores obtained after matching were fused using a closed iterative algorithm combined with weighted sum rule. Results obtained by authors showed that the proposed hybrid scheme was able to attain high-performance rate on large datasets. Fathima et al. (2014) studied a multi-sensor, multi-algorithm and multi-fusion based biometric system using face, finger-print and iris. Multiple sensors were used to capture the biometric images , with face images captured from visible and infra-red cameras were combined at sensor level to form fused face image. Multiple feature extraction algorithm such as Block-Independent Component Analysis (B-ICA), Kalman filter, DCT and FLD were applied on visible face images to obtain match scores, which was fused at score level to form the first classifier. While, B-ICA, Information orientation and Gabor filter were used as feature extraction algorithm on thermal face, finger-print and iris respectively to form the other three classifiers. Final scores from the four classifiers were combined at decision level using dynamic weighted function to determine the claimed identity. Simulation results improved performance of the proposed scheme over other fusion methods.

Islam (2014) proposed a new fusion approach for multi-unit iris recognition at feature and score level. He developed four Markov model classifiers based on feature and score level fusion of left-right iris templates along with individual matchers of left and right iris sub-system. After which the classifiers were combined with a voting rule to determine the claimed identity. Experimental results obtained by the authors show that the robustness of the proposed scheme were more efficient as compared to similar works in literature. Similarly, a hybrid fusion based on feature and score level using palmprint and face biometrics was proposed by (Farmanbar and Toygar, 2015b). The proposed method combined both local and global features of each modality using LBPH, Log-Gabor filter, PCA, and LDA, while, match scores obtained from feature concatenation of each modality were combined using sum rule. Simulations conducted by the authors showed significant improvement of

Table 2.6: Hybrid fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
2008	Tao and Veldhuis (2008)	Score-decision hybrid	face and finger-print	NIST BSSR1	ROC analysis
2011	Kekre et al. (2011)	Feature-decision hybrid	face and iris	CVRPFD and Phoneix	Acc: 99%
2013	Islam et al. (2013)	Feature-score hybrid	face and ear	UND-FRGC	Acc: 98.4%
2014	Fathima et al. (2014)	Score-decision hybrid	face, finger-print and iris	Self acquired, self acquired, CASIA	Acc:78.55%
2014	Islam (2014)	Feature-score hybrid	iris	CASIA	ROC analysis
2015	Farmanbar and Toygar (2015b)	Feature-score hybrid	face and palm-print	FERET and PolyU	Acc: 98.75% Acc: 99.06%

the proposed method over other multi-modal systems.

Table 2.6 shows the summary of results gathered in the studies reviewed above on decision level fusion.

2.2.7 Nature Inspired algorithm based fusion

NI algorithms are an evolutionary approach to learning, in which computing algorithms mimic social behaviours of natural beings for example ants and birds. In general NI algorithms have been applied to computing processes for optimisation purposes. In biometrics, NI algorithms are basically used for feature selection, selecting optimal fusion rules and minimising error rates to improve accuracy and speed of the recognition process. Therefore, popular NI algorithms used in biometrics and successes reported in the literature are explored. Examples include:

(a) **Genetic Algorithms (GA)**

Genetic algorithm belongs to a set of evolutionary algorithms (EA) which generate optimal solutions to search problems by mimicking the process of natural selection (Giot and Rosenberger, 2012). It initializes a population at the beginning of the process and evaluates each member of the population using a fitness function to determine the probability of selection in creating the next generation of off-springs. The process continues until a solution that satisfies minimum criteria is met.

(b) **Particle Swarm Optimization (PSO)**

It is a computational procedure that provides an optimal solution by improving the candidate solution to meet measured criteria. It was developed by Kennedy and Eberhart to imitate the social behavior of birds in the flock and means in which they exchange information to solve optimization problems (Kennedy and Eberhart, 1995; Adewumi and Arasomwan, 2014, 2015). More of PSO will be discussed in chapter four.

(c) **Ant Colony Optimization (ACO)**

It is one of the artificial intelligence techniques for solving complex optimization problems. It mimics the social behavior of ants in finding their way to-and-fro from a food source to their colony. The logic of the algorithm follows the manner ants walking to, and from, a food source, as they deposit chemical substances called *pheromones*. Other ants follow the path in which concentration of pheromones is strongest. This forms a *pheromone* trail that directs other ants to the food source (Saleh and Alzoubiady, 2014).

In the forthcoming section, a report on the performance of the algorithms explored above and how they have been applied in literature is discussed. Altun et al. (2008) proposed a multi-modal system based of fingerprint and iris biometric. They employed a feed-forward neural network for feature extraction of both modalities, due the size of the fused feature vector, GA was applied to pick the optimal set of features that effectively speed up and improve the recognition rate. The test results obtained from simulation by the authors showed the feature selection process improved both the accuracy and speed of the recognition process. The performance improvement of biometric systems by combining their error rates was proposed by (Giot et al., 2010). The authors focused on using GA to learn optimal parameters of score level fusion schemes for improving computation time and accuracy. They also proposed an algorithm for the EER of each uni-model system, which was used was as a fitness function for the GA to optimise the parameters of score fusion techniques like weighted sum and product rule. Simulation results reveal that the proposed scheme reduced the computation time and EER.

Genetic programming for multi-biometric systems was presented by (Giot and Rosenberger, 2012). They proposed the use of GA to select the optimal score fusion schemes (discussed in 2.2.1), in order to determine the claimed identity. The proposed scheme was compared with other fusion schemes like weighted-sum and SVM and it was found that its performance was similar or better than other fusion methods in literature when tested on different biometric datasets. Roy and Kamel (2012)

proposed a new adaptive multi-modal biometric fusion algorithm which is a combination of Bayesian decision fusion and PSO. Bayesian decision was used to fuse decision obtained from each individual modality while PSO was used to find the optimal system operating point (Threshold) that gives the desired system performance and achieve the required security level. Results obtained showed that the proposed system achieved the desired security level.

Khalifa et al. (2013) used Choquet integral and GA in the design of a multi-biometric system made up of face, fingerprint, and palm-print. They used the GA to obtain the optimal sets of fuzzy measurements from the Choquet integral for performing biometric score fusion of match scores. Results obtained by the authors revealed that the proposed scheme's EER was lower than other fusion schemes in literature as well its uni-modal counter-parts. An adaptive multi-modal system using PSO to automatically obtain the desired system performance based security level requirements was presented by (Aly et al., 2013). This was achieved by using PSO to dynamically select the optimal fusion rule which minimised the global cost of both rejecting a genuine user and accepting an impostor. The proposed scheme was validated using iris, fingerprint, and finger-knuckle, and results showed that multi-modal systems performed better than their uni-modal counter-part.

Eskandari et al. (2014) presented a face-iris multi-modal system using local and global feature extractors like LBPH, mPCA, spPCA, LDA, PCA respectively. PSO was used for feature selection to obtain the sets of local and global features that will give the best recognition rate for the face biometric. Match scores of the optimally fused face vector were combined with match scores of LBPH iris feature vector at score level using weighted sum rule. Results obtained by the authors showed performance improvement for the proposed system as compared to other similar works in literature. Gogoi and Bhattacharyya (2014) presented an efficient method for decision level fusion of fingerprint and iris using ACO. Match scores were obtained from both modalities and ACO was used to select the optimal parameters and fusion rule to derive the best system performance. Results obtained by the authors show that the proposed system performed satisfactorily after been tested on a public available dataset.

Canuto et al. (2015) investigated the importance of an optimised ensemble on cancellable multi-biometric systems using face and voice templates. They used biohashing, interpolation and BioConvolving algorithms to convert the features set obtained from face and voice biometric into encrypted templates. Afterwards GA was used for feature template selection to perform feature fusion and weight selec-

tion for biometric score fusion. Results showed that the inclusion of GA into the ensemble system improved its accuracy. In another study, (Amioy and Ajay, 2015) presented an adaptive multi-modal system using ACO for selecting key parameters like weights, decision threshold, and fusion rule, in order to obtain the desired system performances in accordance to changes in security levels. The proposed scheme employed a weighted product and sum rule for fusion and validated the scheme using multiple biometric sets including Palmprint-iris, face-voice, and fingerprint-face. Results presented by authors showed that the proposed adaptive scheme operated at a lower error rate comapared to other studies recoreded in literature.

Table 2.7 shows the summary of results of works reviewed above on NI algorithm based fusion techniques.

Table 2.7: NI algorithms based fusion methods in biometrics

Year	Literature	Approach	Modality	Database	Results
2008	Altun et al. (2008)	GA and ANN	finger-print and iris	self acquired dataset	Acc: 99:3%
2010	Giot et al. (2010)	GA and Weighted sum, product rule	face and keystroke	AR and GREYC	EER: 0.07%
2012	Giot and Rosenberger (2012)	GA, Weighted sum product rule	face and finger-print	BSSR1 and BANCA	HTER :0.4% HTER:0.075%
2012	Roy and Kamel (2012)	PSO and Baysian rule	face, iris and gait features	ORL,CASIA-iris and CASIA-gait	GAR: 96.40%
2013	Khalifa et al. (2013)	GA and Choqet integral	face, finger-print and palm-print	face94, HKPU, and HFR	EER: 0.46%
2013	Aly et al. (2013)	PSO	iris, finger-print and finger-knuckle	CASIA, PolyU and self-acquired	GAR: 97.12% GAR: 98.33% GAR: 98.58%
2014	Eskandari et al. (2014)	PSO and weighted sum rule	face and iris	FERET, BANCA UBIRISv2 and ORL	Acc: 97.5%
2014	Gogoi and Bhattacharyya (2014)	ACO, "AND" and "OR"	iris and finger-print	CASIA and FVC	EER:0.01%
2015	Canuto et al. (2015)	GA sum, K-NN and SVM	face and voice	AR and TIMIT	Acc: 96.86% Acc: 95.13% Acc: 94.67%
2015	Amioy and Ajay (2015)	ACO, weighted sum, and product	face, palm-print, iris, and finger-print	IITD, BSSR1 and XM2VTS	ROC analysis

2.3 Fusion in serial mode

This mode of fusion processes modalities in sequence rather than processing simultaneously as in the case parallel fusion. An extent of flexibility in terms of processing is introduced as the first biometric can be used to authenticate the user based on its confidence level without considering the second modality (Zhang et al., 2014). Furthermore, only images similar to the probe image are considered when processing the second modality, which improves the accuracy of the system (Ramadan et al., 2015). Figure 2.8 shows the general architecture for fusion in a serial mode using two modalities sequentially. Below are works under this fusion mode and they are discussed in the following paragraphs.

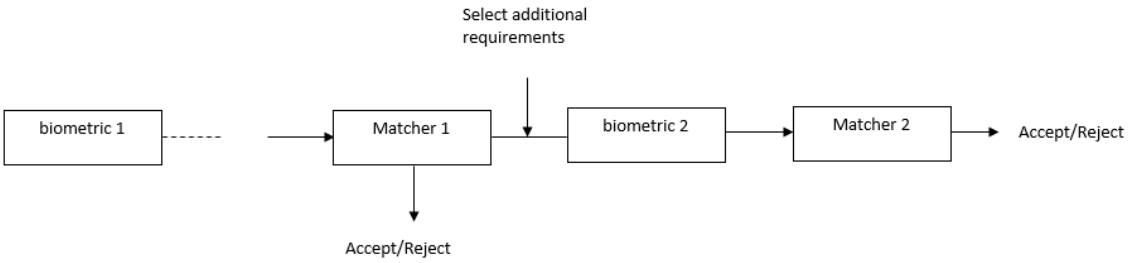


Figure 2.8: Architecture for fusion in serial mode

Marcialis and Roli (2007) proposed a serial fusion scheme consisting of the face and fingerprint matchers. The authors proposed a lower and upper threshold at zeroFRR to ensure genuine users are not rejected and zeroFAR to ensure impostors are not accepted. The two threshold values were used to determine if the second biometric will be required to find the claimed identity. This was achieved by evaluating whether the match score of the first biometric was greater than the upper threshold in which the user will be termed as genuine. In cases where the match score was less than the lower threshold, the user was termed as an impostor, otherwise, the second biometric is called up to make the final decision. Simulation results revealed that the sequential fusion of fingerprint and face performed better than that of face and fingerprint systems. A novel serial fusion scheme with a mathematical model for combining two sequential matchers was proposed by (Marcialis et al., 2009). The authors derived analytical relationships between the fused system and the individual matchers to predict the best sequence of unimodal systems that will provide the best trade-off between the performance of the system and the verification time. Simulation results validated with different multi-modal databases were compared with fusion schemes in parallel mode.

Table 2.8: Summary of serial fusion in biometrics

Year	Literature	Approach	Modality	Database	Results
2007	Marcialis and Roli (2007)	Threshold comparison	face and finger-print	AR and FVC2000-DB2	ROC analysis
2009	Marcialis et al. (2009)	Threshold analysis	face and finger-print	AR and FVC2000-DB2	ROC analysis
2010	Marcialis et al. (2010)	ROC and error estimation analysis	face and finger-print	NIST BSSR1	ROC analysis
2014	Zhang et al. (2014)	semi-supervised learning	face, finger-print and gait	SDUMLA	Acc:96.4% Acc: 98.6 Acc:99.0

Marcialis et al. (2010) presented a theoretical framework for a serial fusion of multiple unimodal systems for performance rating. The authors used the error rates (FRR and FAR) within the ROC curves of each matcher to analytically select the best sequence of unimodal systems that will provide the best accuracy. Also considering that the ROC curves were affected by error due to bad estimation of the error rates for each matcher, the authors proposed a mathematical model for the estimation of errors to accurately provide the level of acceptability of the system. A novel serial multi-modal system using semi-supervised learning was proposed by (Zhang et al., 2014). The basic concept behind the framework was to use more user acceptable traits at the beginning of the process chain and improve weaker traits using semi-supervised learning. This was achieved by coupling relationships between weaker and stronger traits. The proposed system was based on two prototype biometric traits; face-fingerprint and gait-fingerprint. Results obtained from simulation showed that the proposed system possessed the capability to boost user convenience and system performance.

Table 2.8 shows the summary of results from the works reviewed above for fusion in serial mode.

2.4 Summary

Different fusion schemes for combining two or more biometrics in order to improve accuracy have been proposed in several studies. Hybrid systems, are perceived to perform better than single fusion schemes because they combine two more levels of fusion to provide the final output while NI based algorithms increase computation speed by reducing the dimension of feature vectors. In the review, it emerged that

most hybrid systems are hinged on the combination of two levels of fusion to determine the claimed identity. In cognizance of that, this study proposes a hybrid fusion mechanism for face and iris biometrics that combines fusion at more than just two levels, including feature, score, and decision level. Furthermore, most studies reviewed only propose a parallel architecture in which both modalities are processed simultaneously; with less work on serial fusion. More so, serial fusion has not been used to tackle the problem of large class imbalance (to best of my knowledge) of biometric data in literature . Therefore, a serial fusion mechanism using BPSO and (iRVM which solves the problem of large class imbalance by reducing the number of impostor scores is proposed.

Chapter 3

Feature Extraction Algorithms

3.1 Introduction

Prior to the fusion of multiple pieces of evidence from different modalities, feature extraction will have to be performed to select the salient information in the biometric image. In most cases, a two-dimensional image is reduced to a one-dimensional feature vector or a set of keypoint descriptors that are stored as templates in the biometric database. This chapter therefore is an exploration of feature extraction algorithms used for obtaining respective features from face and iris modality. A face and iris modality has been chosen for this research work because face images are non-intrusive and iris traits are accurate and stable over time. Moreover, the iris is located on the face region, therefore, a single sensor can be used to capture both modalities reducing the time and cost of acquisition. Here, five standard feature extraction algorithms are considered and they are divided into two groups namely: local and global methods (Eskandari et al., 2014).

The local feature extraction methods used in this study extract features based on a sub-region defined within the image rather than considering the image as a whole. This operation is repeated across the whole image until all the sub-features are extracted. Global feature extraction methods, on the other hand, operate on the image as a whole to extract features for recognition. One of the main merits of using local feature extractors is that it is robust to partial illumination and occlusion that may occur at some regions in the image.

This chapter unfolds as follows; First a description of the global and local feature extractors is given, followed by, a description of techniques employed for face and iris recognition. The face and iris datasets used are also described, along with the data protocol used for performing experiments and finally, experimental results for the individual face and iris system are presented.

Figure 3.1: Mean image of the training set

3.2 Global methods

3.2.1 Principal Component Analysis (PCA)

In general, images are of high dimension given that they are represented in two dimension space of size $m \times n$. Therefore, for every image, there lies mn number of pixels in that space, in which only a few number of pixels carry useful information. Therefore, PCA is used to transforms a group of correlated variables to a smaller group of uncorrelated variables (Martinez and Kak, 2001). The logic behind this method is to maximise the total scatter of the centered images in the training set.

Given a set of N training samples of images $I_1, I_2, I_3, \cdots, I_N$. PCA algorithm is computed as follows:

1. Convert the set training samples to row vectors $x_1, x_2, \cdots, x_N$
2. Compute the mean of the column vectors. The mean image of the training set is given in Figure 3.1
$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3.2.1}$$
3. Calculate the covariance matrix S
$$S = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)(x_i - \mu)^T \tag{3.2.2}$$
4. Compute the eigenvectors and eigenvalues of the covariance matrix S
$$Sv_i = \lambda_i v_i \quad for\ i = 1, 2, \cdots, N \tag{3.2.3}$$

where v and λ are eigenvectors and eigenvalues respectively.

5. Select the eigenvectors or eigenfaces corresponding to the largest p eigenvalues. Figure 3.2 shows samples of eigenfaces computed.

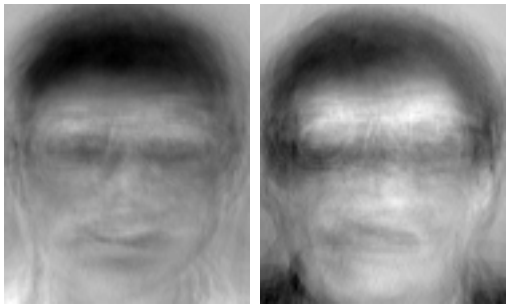


Figure 3.2: Sample of eigenfaces computed from training set

6. Compute the projections or principal components Y

$$Y_i = [v_1, v_2, \cdots, v_p]^T (x_i - \mu) \tag{3.2.4}$$

where Y_i represents the feature vector for image I_i

A feature vector of $1 \times p$ dimension is obtained, where $p \leq N$ and N is the number of training set.

3.2.2 Linear Discriminant Analysis (LDA)

LDA is similar to PCA in the sense that both try to project data into a given vector space. Unlike PCA, LDA attempts to model the class difference inherent in the training data. It acheives this by maximising variance between classes and minimising varinace within classes or in simple terms; it clusters similar classes together and seperates different classes far away from each other. (Martinez and Kak, 2001).

Given a set of N training samples of images $I_1, I_2, I_3, \cdots, I_N$ with image class $1, 2, \cdots, C$. LDA algorithm is computed as follows:

1. We convert the set training samples to row vectors $x_1, x_2, \cdots, x_N$ with image vectors from C classes defined as $X = \{X_1, X_2, \cdots, X_C\}$

2. Compute PCA on the dataset to project data to $(N - C)$ dimension space.

3. Compute the total mean of all images in dataset and mean of images for each class.

$$\begin{aligned} \mu &= \frac{1}{N}\sum_{i=1}^{N} x_i \\ for\ i &\in \{1, 2 \cdots, C\} \\ \mu_i &= \frac{1}{|X_i|}\sum_{x_j \in X_i} x_j \end{aligned} \tag{3.2.5}$$

4. Compute the within-class and between-class scatter matrix S_w and S_b

$$\begin{aligned} S_w &= \sum_{i=1}^{C}\sum_{x_j \in X_i}(x_j - \mu_i)(x_j - \mu_i)^T \\ S_b &= \sum_{i=1}^{C}(\mu_i - \mu)(\mu_i - \mu)^T \end{aligned} \tag{3.2.6}$$

5. Find the projection matrix W such that maximizes the class separability.

$$\underset{W}{\text{argmax}}\ \frac{|W^T S_b W|}{|W^T S_w W|} \tag{3.2.7}$$

   where W is computed as $W = W_{lda}^T W_{pca}^T$ (projects samples to $(C-1)$ dimension space. Where W_{lda} and W_{pca} are LDA and PCA eigenvectors.

A feature vectors of size $1 \times (C-1)$ is obtained, where C represents the number of classes inherent in the data. Figure 3.3 shows an example of fisher faces computed for the face biometric.

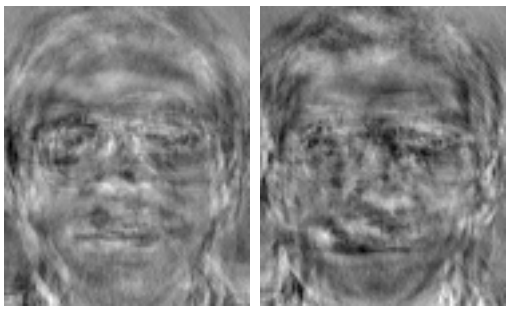


Figure 3.3: Sample of fisher faces computed from training set

3.3 Local methods

3.3.1 Local Binary Pattern Histogram

The basic concept of LBPH is an 8-bit or 16-bit operator used on an area of the image (3×3 *or* 9×9) such that the neighboring pixels greater in value than the center pixel are assigned a value of 1, otherwise 0 is assigned (Ahonen et al., 2004). Then the histogram is constructed for labels to be used as texture descriptors. A basic LBP operator is shown in Figure 3.4

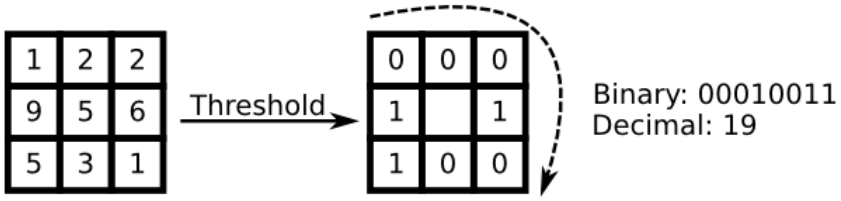


Figure 3.4: Basic LBPH operator

After this text descriptor was developed, it was noticed that fixed neighborhood could not efficiently describe details of different scales. Therefore, the operator was extended to use the circular neighborhood as shown in Figure 3.5 (Ojala et al., 2002). A circle with specified radius is defined within every block of the image in which the sampling points (neighbors) are located at the edge of the circle. For every

Figure 3.5: Extended LBP (8,2) oeprator

neighborhood defined within the image it is generally described as (P, R), where P is the number sampling points on the circle and R is the radius of the circle. For given point (x_c, y_c) the new values for the neighboring pixels (x_p, y_p) can be calculated as follows:

$$\begin{aligned} x_p &= x_c + Rcos(\frac{2\pi p}{p}) \\ y_p &= x_c + Rsin(\frac{2\pi p}{p}) \end{aligned} \tag{3.3.1}$$

Finally, spatial histograms are generated as feature vectors of size $(1 \times N \times K)$ are

obtained, where N represents the number of the training set and K represents the number partitions (Azom et al., 2015).

3.3.2 sub-pattern PCA (spPCA)

This is a variant of PCA in the sense that images are first turned into row vectors and divided into k sub-patterns (Liu and Lu, 2012). PCA is then performed on each sub-pattern group of the training set. Below describes the step by step process for

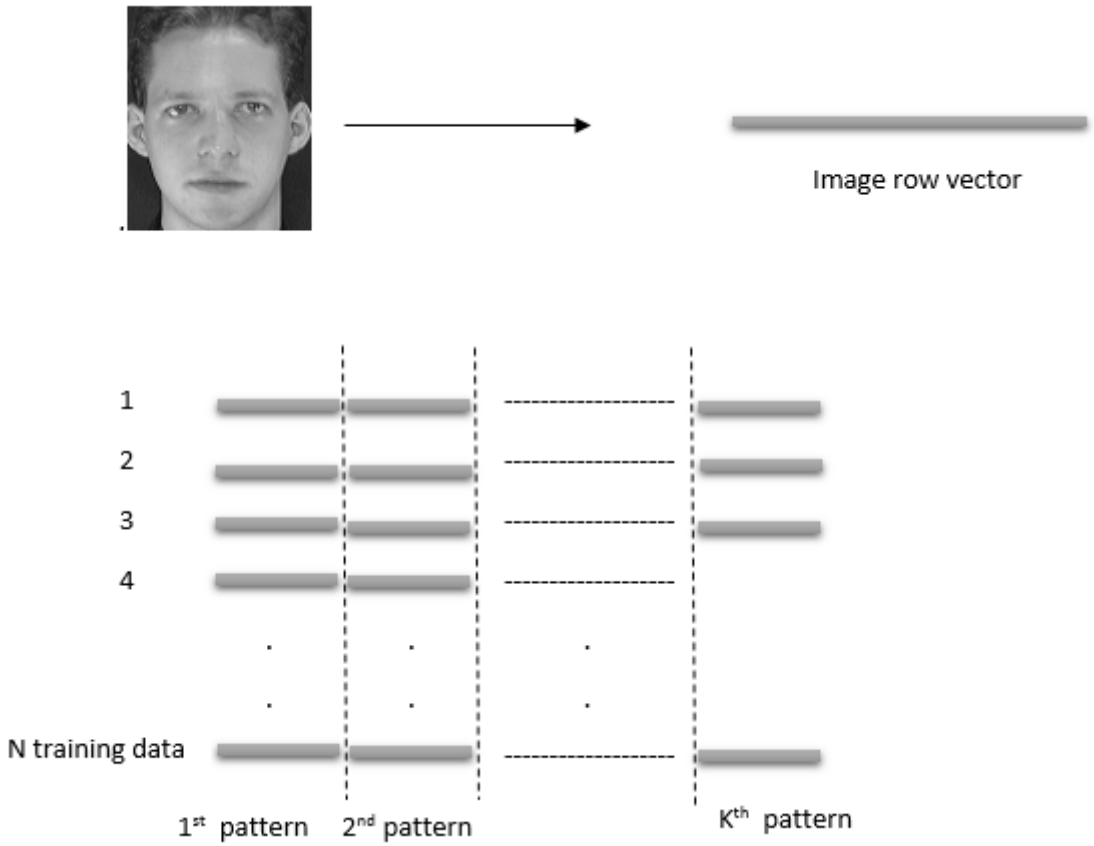


Figure 3.6: Images divided in sub-patterns of spPCA

implementing spPCA:

Given an N training set of images $I_1, I_2, I_3, \cdots, I_N$ of size $m \times n$

1. convert every image in the training set in row vectors $x_1, x_2, \cdots, x_N$
2. divde the images in the training set into K sub-patterns as shown in Figure 3.6

3. Compute the j^{th} sub-pattern mean as :

$$\mu_j = \frac{1}{N}\sum_{i=1}^{N} x_{ij} \quad for\ j = 1, 2, \cdots, K \tag{3.3.2}$$

4. Compute the j^{th} sub-pattern covariance matrix as :

$$S_j = \frac{1}{N}\sum_{i=1}^{N}(x_{ij} - \mu_j)(x_{ij} - \mu_j)^T \tag{3.3.3}$$

5. Then select the j^{th} sub-pattern eigenvectors $v_1^{(j)}, v_2^{(j)}, \cdots, v_p^{(j)}$ with the largest eigenvalues.

6. Since image I_i is partitioned into sub-patterns $\{x_{i1}, x_{i2}, \cdots, x_{iK}\}$. also the feature vector Y_i corresponding to image I_i will be partitioned into sub-patterns $\{Y_{i1}, Y_{i2}, \cdots, Y_{iK}\}$. Therefore:

$$Y_{ij} = [v_1^{(j)}, v_2^{(j)}, \cdots, v_p^{(j)}]^T(x_{ij} - \mu_j) \tag{3.3.4}$$

A feature vector of $1 \times p$ dimension is obtained for every sub-pattern group. where $p \leq N$ and N is the number of the training set.

3.3.3 Modular PCA (mPCA)

This is also a variant of PCA that divides the image into blocks and then performs PCA on the set of divided images as shown in Figure 3.7 (Liu and Lu, 2012). mPCA is implemented as follows:

Figure 3.7: Dividing an image into 3×3 block for mPCA

Given an N training set of images $I_1, I_2, I_3, \cdots, I_N$ of size $m \times n$

1. Divide every image I_i of size $m \times n$ into blocks of size $p \times q$

2. Reshape each block into a $\frac{mn}{pq}$ row vector x_{ij}

3. Compute the mean of all blocks for all images in the training set

$$\mu = \frac{1}{N \times p \times q} \sum_{i=1}^{N} \sum_{j=1}^{p \times q} x_{ij} \tag{3.3.5}$$

4. Compute the covariance matrix S as :

$$S = \frac{1}{N \times p \times q} \sum_{i=1}^{N} \sum_{j=1}^{p \times q} (x_{ij} - \mu)(x_{ij} - \mu)^T \tag{3.3.6}$$

5. Then select the eigenvectors $v_1, v_2, \cdots, v_p$ with the largest eigenvalues.

6. The we compute the feature vectors Y_{ij} by projecting sample x_{ij} to the eigenspace.

$$Y_{ij} = [v_1, v_2, \cdots, v_p]^T (x_{ij} - \mu) \tag{3.3.7}$$

A feature vector of $1 \times p$ dimension is obtained for every sub-pattern group. where $p \leq N$ and N represents the number of the training set.

3.4 Face recognition

Both local and global feature extraction algorithms have been applied to the face to obtain shape and texture based features. For the global methods, eigenvectors belonging to non-zero eigenvalues are extracted for computing feature vectors of the global methods. With regards to local methods, N = 9 partitions have been used for computation purposes. Meaning that each image in the training set was divided into 3×3 block. For the LBPH method, we have employed the (P, R) circular neighbourhood operation. In this study, $P = 8$ and $R = 1$. Finally, Euclidean distance was employed in calculating match scores between the test and training images using equation 3.4.1

$$\sum_{j=1}^{N} \parallel X - Y_j \parallel_2 \tag{3.4.1}$$

where X represents the projection of the test sample Y_j is the projection of the j^{th} image in the training set for $j = 1, 2, \cdots, N$ and N represents the number of images in the training data.

3.5 Iris recognition

Iris images are retrieved under different conditions for example under visible and infra lights, which determine the level of noise in these images. However, in ideal situations, the iris and pupil boundaries form a uniform concentric circles, in which using an edge map together with Hough transform works perfectly (Sim et al., 2014). However, this is not always the case as most iris images are captured with noise due to specular light and occlusion by the eyelids. In order to provide a solution to this problem, first, noise due to specular light and occlusion will have to be removed and secondly, the iris and pupil boundaries will then be extracted as contours. In achieving this, two pre-processing techniques will have to be undertaken, namely:

3.5.1 Segmentation

In this study, the iris and pupil boundary are extracted the Viterbi algorithm (Sutra et al., 2012). Owing to the existence of noise due specular light in the iris image, the white holes present in the pupil (as shown in Figure 3.8) were filled using morphological opening. After which the pupil became the darkest region in the eye image, therefore, pupil center (x_p, y_p) was roughly estimated. This guided the search for Viterbi algorithm to estimate the sequence of radii $\{R_1, R_2, \cdots, R_n\}$ and $\{\theta_1, \theta_2, \cdots, \theta_n\}$ that accurately defined the pupil contour. Once the pupil contour was defined least square fitting was applied to define the pupil circle (the accurate pupil center was determined). Furthermore, the accurate pupil center was used as a rough estimate for the Viterbi algorithm to find the accurate iris contour and its was found using the same method employed for finding the pupil center. Once the iris center was determined, an iris mask was built to extract the iris concentric circles from the eye image (as shown Figure 3.9).

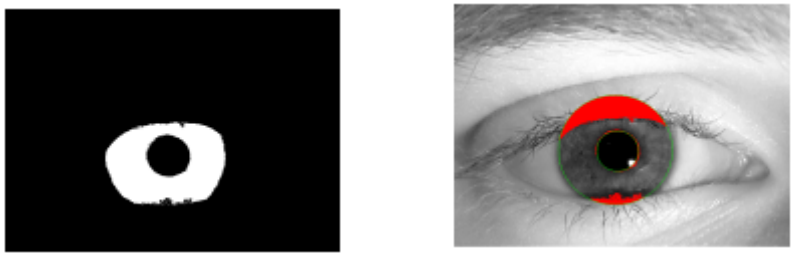
Figure 3.8: Iris image captured with noise due specular light

Prior to applying Viterbi algorithm, anisotropic smoothing was performed to preserve important edges followed by a Sobel filter to create an edge map. After the

pupil and iris contour estimation, the noise due to the occlusion of the eyelashes was removed by using Full Width at Half Maximum (FWHM) based on Gaussian density function (Sutra et al., 2012). Pixel $I_{i,j}$ is an occlusion if:

$$| I_{i,j} - \mu | = 2.35\sigma \tag{3.5.1}$$

Where μ and σ refers to the mean and variance of the pixel intensities of the



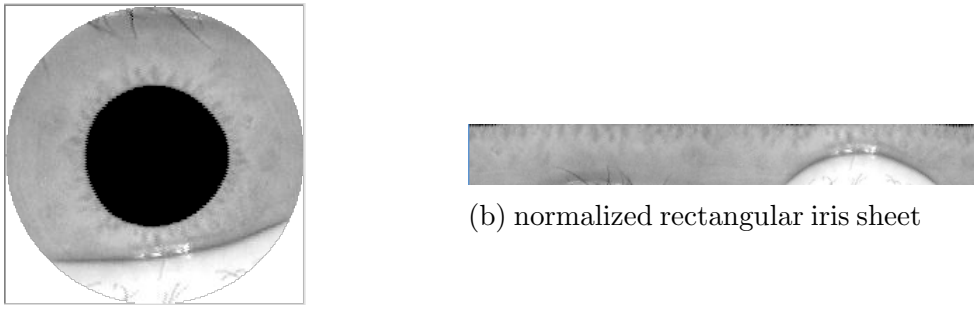
(a) (b)

Figure 3.9: Binary iris mask (a) and Segmented iris image (b)

segmented iris image.

3.5.2 Normalisation

In this process, Daugman's rubber sheet model was adopted, as it remaps and unwrap the iris template from cartesian coordinates to a non-concentric polar coordinates (Daugman, 2004). Coarse pupil and iris contours are detected using Viterbi algorithm by selecting the minimum number of noisy points and angles that define a closed contour.



(a) extracted iris image

(b) normalized rectangular iris sheet

Figure 3.10: Iris normalization

Let L and B represent the length and breadth of the rectangular sheet. compute

$\theta_i \in [0, 2\pi]$ such that:

$$\theta_i = 2\pi \frac{i}{B}, i = 0, 1 \cdots B \tag{3.5.2}$$

Let (x_p, y_p, ϕ_p) and (x_r, y_r, ϕ_r) represent coordinate points of the coarse pupil and iris contours respectively where x and y are coordinate point of the radius relative to the center and ϕ is the angle of non-regular sampling. An estimate (X_i^r, Y_i^r) corresponding to θ_i for two nearest points s, $s+1$ as shown in Figure 3.11. The (X_i^r, Y_i^r) are computed as follows:

$$X_i^r = (1-\alpha)x_r^s + \alpha x_r^{s+1} \tag{3.5.3}$$

$$Y_i^r = (1-\alpha)y_r^s + \alpha y_r^{s+1} \tag{3.5.4}$$

$$\alpha = \frac{\theta_i - \phi_r^s}{\phi_r^{s+1} - \phi_r^s} \tag{3.5.5}$$

The same process is carried out to estimate the pupil points (X_i^p, Y_i^p) corresponding to θ_i. In Figure 3.12 a displays the mapping of the extracted iris in cartesian coordinate to polar coordinates. The normalised iris image can be passed to any of the feature extraction algorithms described above to obtain its feature vectors. Moreover, euclidean distance in equation 3.4.1 is employed to obtain match scores for the iris modality.

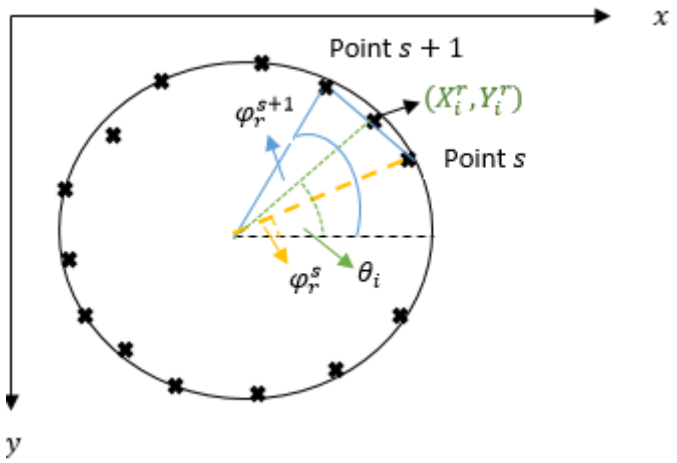


Figure 3.11: Computing coordinate points (X_i^r, Y_i^r)

3.6 Datasets

In order to validate the above methods, two face datasets and an iris dataset were employed . Also a chimeric dataset was constructed from these three datasets to

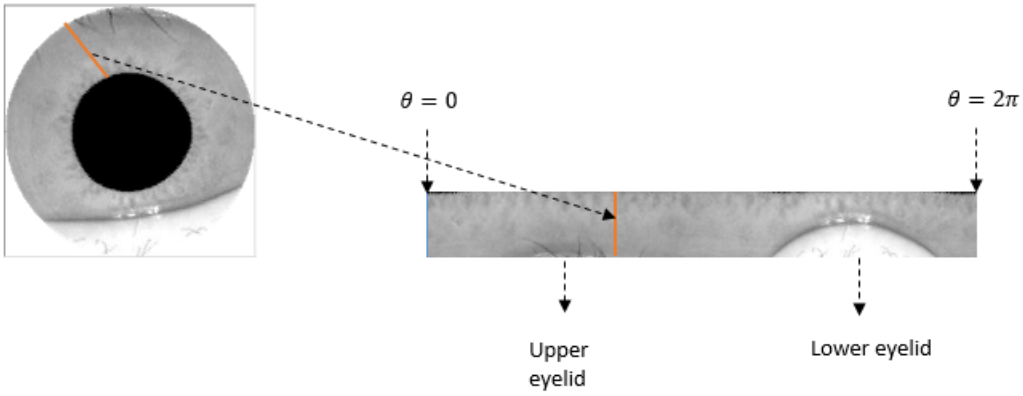


Figure 3.12: Mapping of iris image to rectangular rubber sheet

test the proposed fusion schemes which will be discussed in chapter four and five. brief description of datasets used in this study is provided underneath.

3.6.1 Face datasets

ORL and Feret database has been used to construct both training and test face data (Ahonen et al., 2004; Phillips et al., 1998). ORL contains 10 images each for 40 subjects. Each subject has images of size 112×92 with a different pose, facial expression, and illumination conditions. The Feret database contains thousands of images of size 512×768, also with varying illuminations and poses angle. The images selected from Feret dataset were resized to that of its ORL counterpart using a bi-cubic interpolation. The face database was constructed by combining ORL and Feret dataset in the ratio 1:5 that is the ORL dataset makes up one fifth of the face database. A mixed face dataset made of four images each for 200 users (total of 800 images) have been constructed. Here we have used three images for training and the rest for testing.

3.6.2 Iris datasets

The CASIA iris-interval database contains images of the left and right iris of 249 subjects captured under the infra-red light (Tieniu and Zhenan, 2010). Here, 200 users have been selected from the database and constructed the training and test set same as the face database.

3.7 Experimental results

As mentioned earlier the two groups of feature extraction methods have been applied on both face and iris modality and the performance of each unimodal system has been expressed in terms of recognition accuracy and equal error rate.

Table 3.1: Recognition rates for uni-modal systems.

RR for face and iris (%)		
Method	ORL/FERET	CASIA
PCA	70.38	72.36
LDA	69.96	**84.56**
LBP	**76.13**	83.87
spPCA	71.88	75.92
mPCA	**76.15**	77.13

Table 3.1 displays the recognition rate for both modalities. It is evident that the local feature extraction methods like LBP and mPCA (bold numbers) performed better than other methods for face recognition, while LDA, a global method produced the best performance for iris recognition. This is because local methods operate on sub-regions in the images rather the whole image as with global methods, making it robust to varying illuminations and partial occlusion present in an image, which is prominent with face images. It is seen that for both modalities, the variants of PCA (spPCA and mPCA) perform better than the parent method PCA, because both spPCA and mPCA both operate on sub-regions of an image rather than the whole image in case of PCA.

Table 3.2: Error rates for uni-modal systems.

Error rates for face and iris (%)				
Method	ORL/FERET		CASIA	
	EER	min. TER	EER	min. TER
PCA	0.148	0.296	0.136	0.272
LDA	0.150	0.300	**0.077**	**0.154**
LBP	**0.119**	**0.238**	0.081	0.162
spPCA	0.141	0.282	0.120	0.240
mPCA	**0.119**	**0.238**	0.119	0.238

Table 3.2 shows the EER and minimum TER (twice the EER) for every feature extraction algorithm, once again LBP and mPCA had the lowest EER for face unimodal system while LDA obtained the lowest EER for the iris unimodal system. The ROC curve showing the performance of both modalities at different system operating points as shown in Figure 3.13 and 3.14 are provided. For the face unimodal system, the area under the curve for both LBP and mPCA is lower than all other methods confirming that both methods give the best performance for face recognition. Likewise, for the iris uni-modal system, the area under the curve for LDA is lower than any other method used.

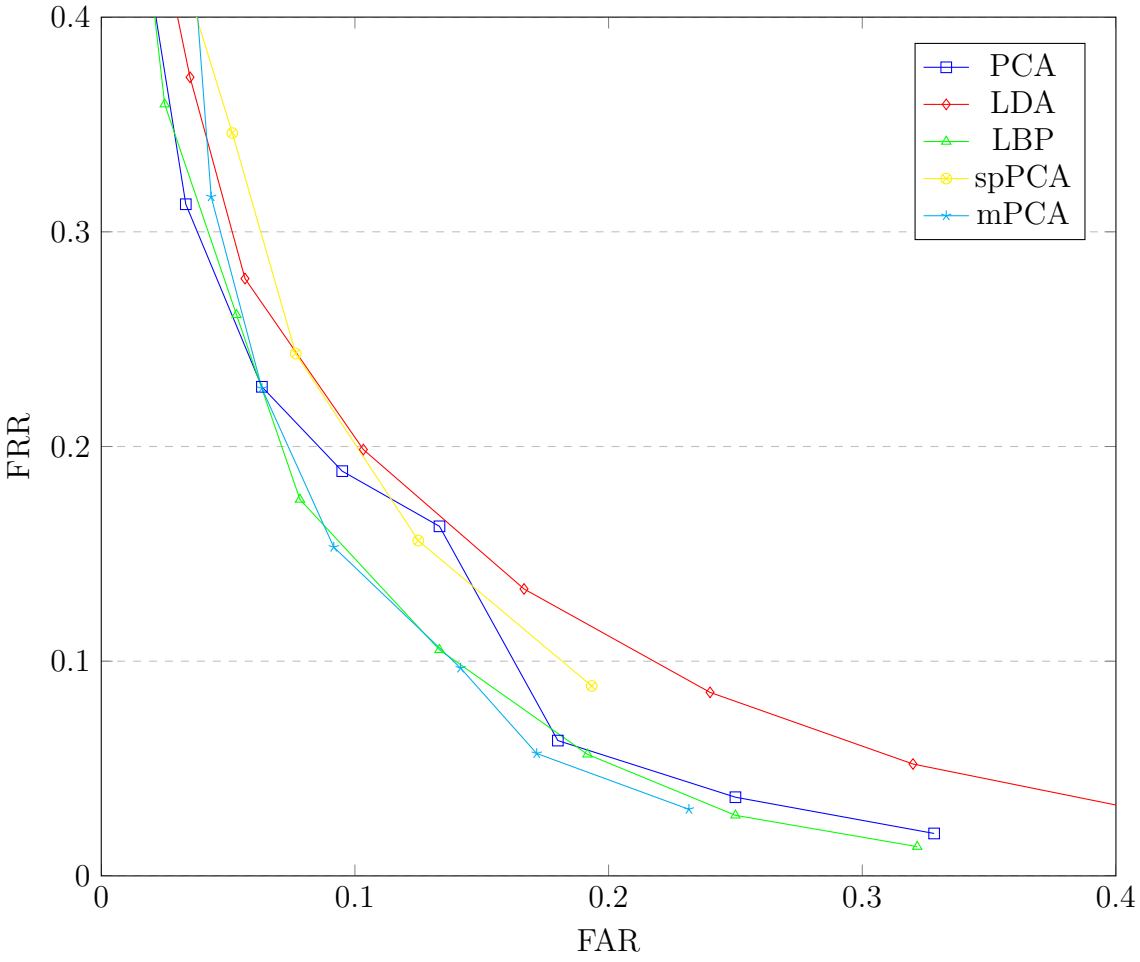


Figure 3.13: ROC curves for face uni-modal system

Figure 3.15 and 3.16 shows plots of FRR and FAR against the system threshold operating points for the best matchers LBPH for face and LDA for Iris. In the two plots it is evident that the graphs for FRR and FAR intersect and this point of intersection defines the EER of the system. As seen from Figure 3.15 the point of

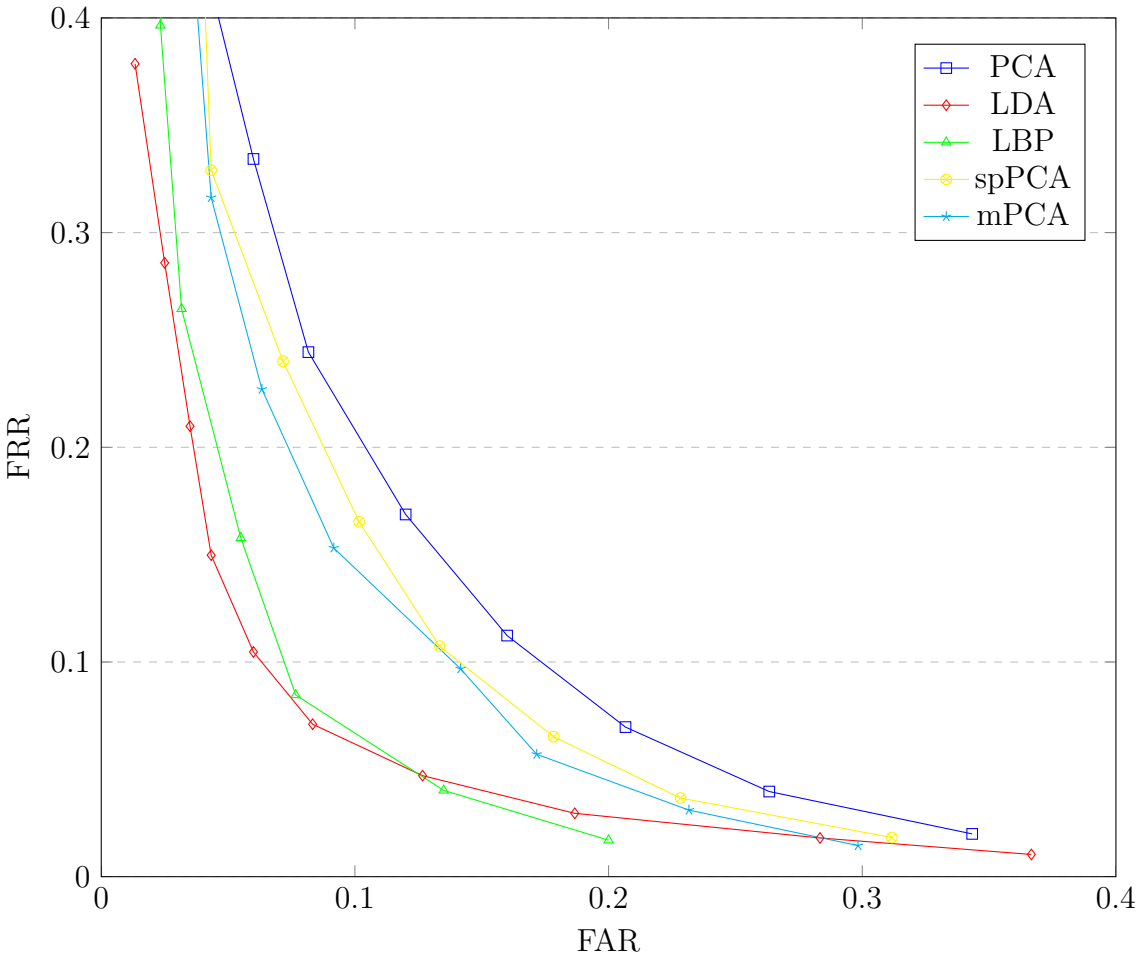


Figure 3.14: ROC curves for iris uni-modal system

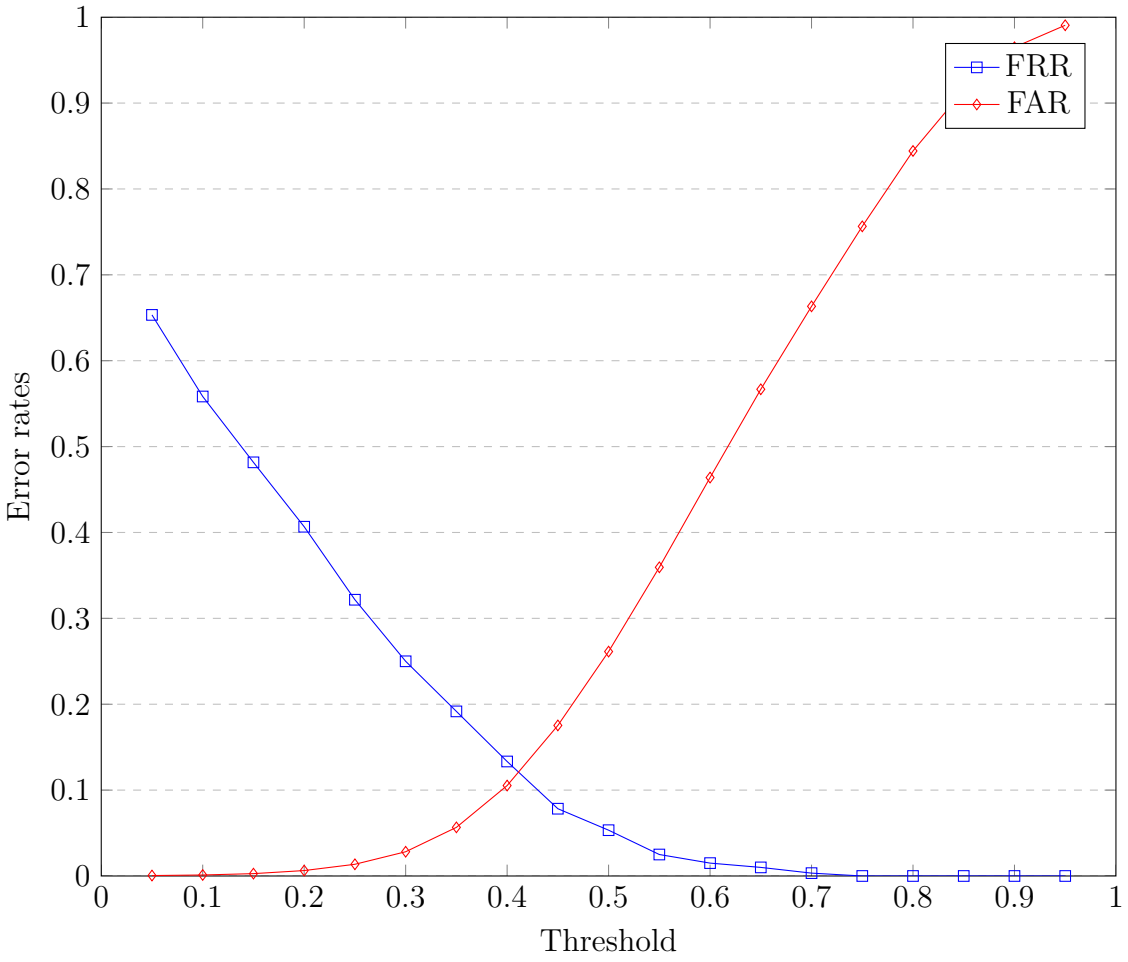


Figure 3.15: Plot of FRR and FAR against different system operating points for LBPH face

intersection falls between 0.1 and 0.2 as the EER of LBP face is given as 0.119 as shown in Table 3.2, while from Figure 3.16 the point of intersection falls between 0 and 0.1 as the EER for LDA iris is given as 0.077 as shown in Table 3.2.

3.8 Summary

This chapter has given a description of feature extraction algorithms used to extract features from the face and iris. Also described the pre-processing methods in segmenting the iris patterns from an eye image. Finally, experimental results of the local and global feature extraction algorithms on the individual modalities are presented and it is shown that local methods perform best for face modality as they are more robust to varying illuminations and partial occlusions that may occur in

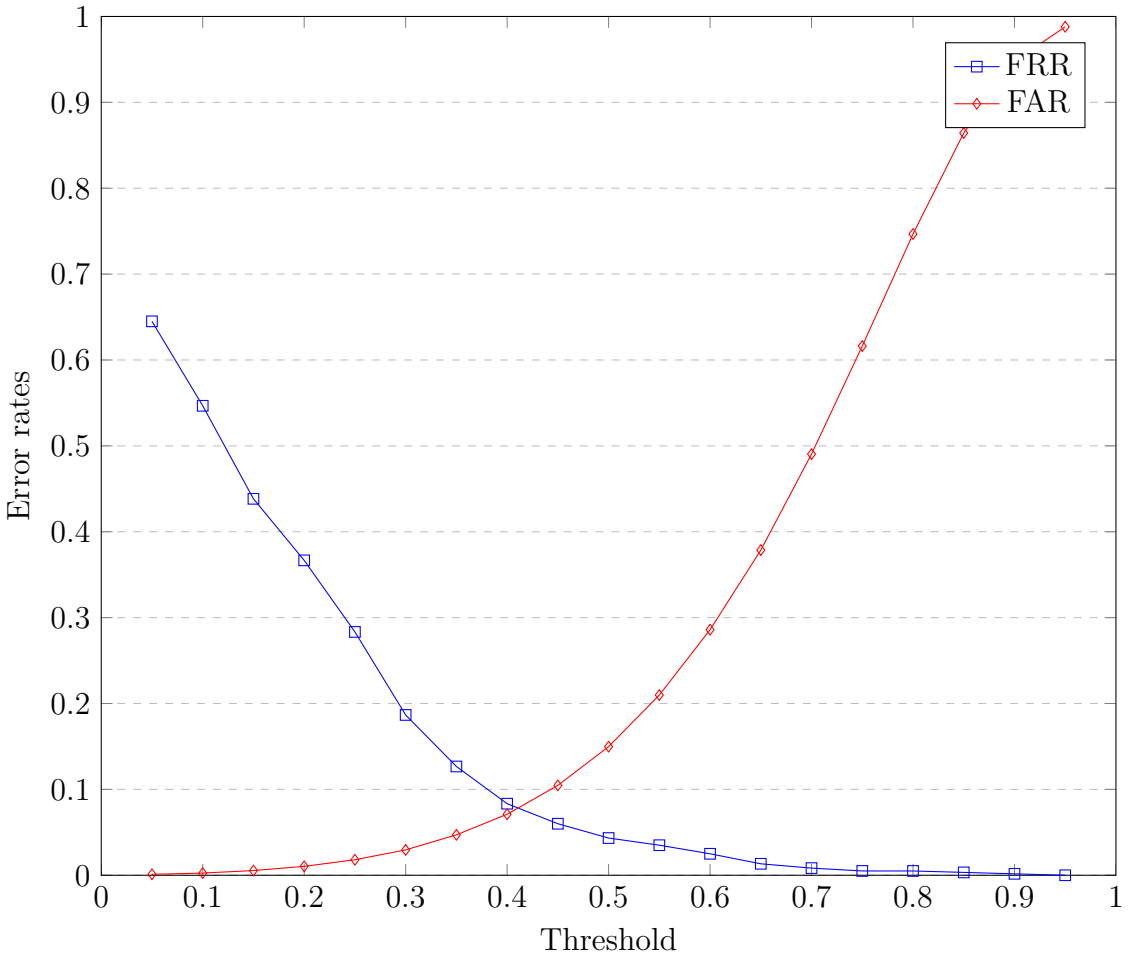


Figure 3.16: Plot of FRR and FAR against different system operating points for LDA Iris

some regions in the image, while global methods perform best for the iris modality.

Chapter 4

Hybrid Fusion at Feature, Score and Decision level

4.1 Introduction

The core of this chapter, is to present and discuss the proposed hybrid fusion which combines three level of fusion namely; the feature, score, and decision level fusions. As images pass through different modules in a biometric system, the level of information extracted from the image decreases, therefore, fusion at feature level is expected to give that best accuracy, as it contains more information about each modality. However, this is always not the case as features may contain redundant values that reduce the recognition accuracy (Ross and Govindarajan, 2005). Fusion at score level contains less information about modalities as only scores (scalar value) are combined, however, provides the capability to combine modalities at different noise levels present in each modality (Eskandari and Toygar, 2014). Decision level fusion contains the least information about modalities because it only captures the class labels, but it is often preferred because it is easy to implement (Tao and Veldhuis, 2009). These three fusion level can be combined to form a hybrid fusion scheme which compensates for the shortcomings of individual fusion methods to enhance the performance of multi-biometric systems.

Hybrid fusion mechanisms have been studied in literature (mostly based on two levels) and produced promising results as they harness the capabilities of different levels of fusion to produce the final result (Fathima et al., 2014; Islam, 2014; Farmanbar and Toygar, 2015b). This chapter commences with a description of the proposed hybrid fusion and this is ensued by a presentation of the results obtained from the experiment. Results obtained from comparing the proposed scheme with other fusion mechanisms in literature will make the conclusion.

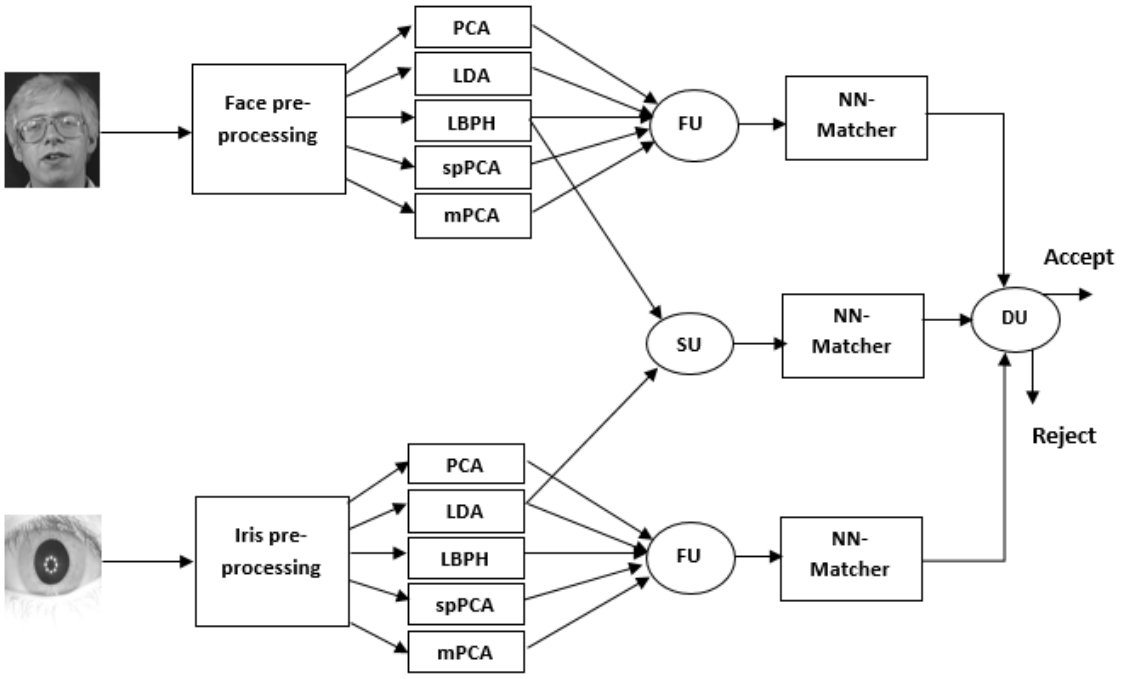


Figure 4.1: Proposed hybrid fusion scheme, Key: FU = Feature level fusion, SU = Score level fusion, DU = Decision level fusion

4.2 Proposed hybrid fusion scheme

Figure 4.1 shows the architecture for the proposed fusion scheme.For the individual face and iris modality, the five feature extraction algorithms discussed in chapter three are applied. Feature level fusion is then performed for each modality on the feature vectors obtained from the five methods to form the first two classifiers. Next a weighted score level fusion using matching scores retrieved from LBPH for face and LDA for Iris is performed to form the third classifier. Outputs obtained from each of the classifiers is then fused at decision level using the majority voting rule to determine the claimed identity (genuine or imposter user). In the following sections, the stages involved in building the proposed scheme are outlined.

4.2.1 Feature level fusion of face and iris vectors

After performing the pre-processing and feature extraction stages for both face and Iris biometric templates, feature level fusion is performed for each of the face and Iris templates separately. Considering that the magnitude and dimensions of features obtained from each feature extraction methods might not be same. Z-score normalisation technique discussed in chapter two (equation 2.2.14) has been employed to

bring all feature values into the same range. For spPCA and mPCA methods each image in the training data is divided into nine sub-patterns and a 3×3 block respectively. Therefore, nine feature vectors per image are obtained from this methods, together with single feature vector each from PCA, LDA, and LBPH to bring the total number for feature vectors after fusion to twenty-one as shown in Figure 4.2.

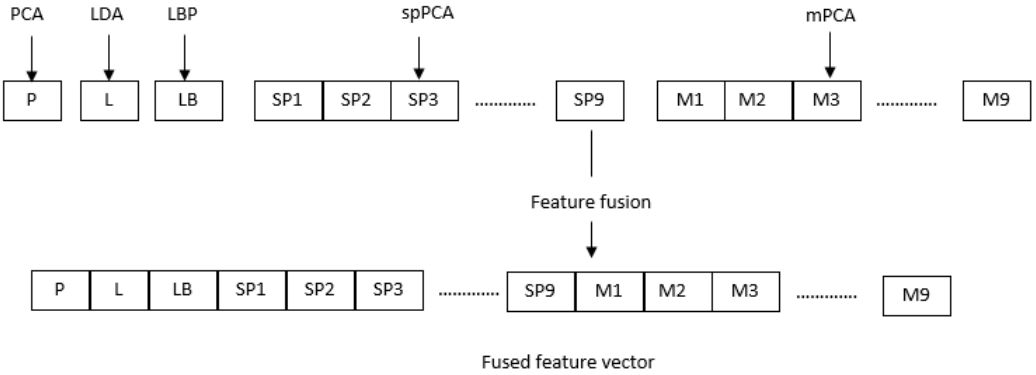


Figure 4.2: Feature fusion of face and iris vectors, Key: P,L,LB,SP,M represent features from PCA, LDA, LBP, spPCA and mPCA respectively

Let X_i and Y_i represent features extracted from face and Iris respectively; where $i = 1, 2 \cdots N$ where N refers to the number of features extracted from each methods. In this study $N = 21$, as single feature vectors are obtained per image for PCA, LDA and LBPH methods, while nine feature vectors per image are each extracted for spPCA and mPCA methods. Then the fusion of feature vectors Z_f and Z_I for face and Iris respectively are expressed as:

$Z_f = \{X_1, X_2 \cdots X_N\}$ and $Z_I = \{Y_1, Y_2 \cdots Y_N\}$ In order to perform recognition on this method we consider two instances of the fused face vectors $Z_{1f} = \{X_{11}, X_{12}, \cdots X_{1N}\}$ and $Z_{2f} = \{X_{21}, X_{22}, \cdots X_{2N}\}$. An average Euclidean distance d is obtained as:

$$d = \frac{S_1 + S_2 + \cdots + S_N}{N} \tag{4.2.1}$$

Where $S_1 = \parallel X_{21} - X_{11} \parallel_2$ and $S_N = \parallel X_{2N} - X_{1N} \parallel_2$ The process is also performed for the iris template, then NNC is now used for classification.

4.2.2 Score level fusion of local and global methods

Score level of fusion is implemented based on LBPH for face and subspace LDA for Iris because of their performance on the individual unimodal systems. Weighted score fusion has been employed at this stage and it is computed as follows:(Eskandari

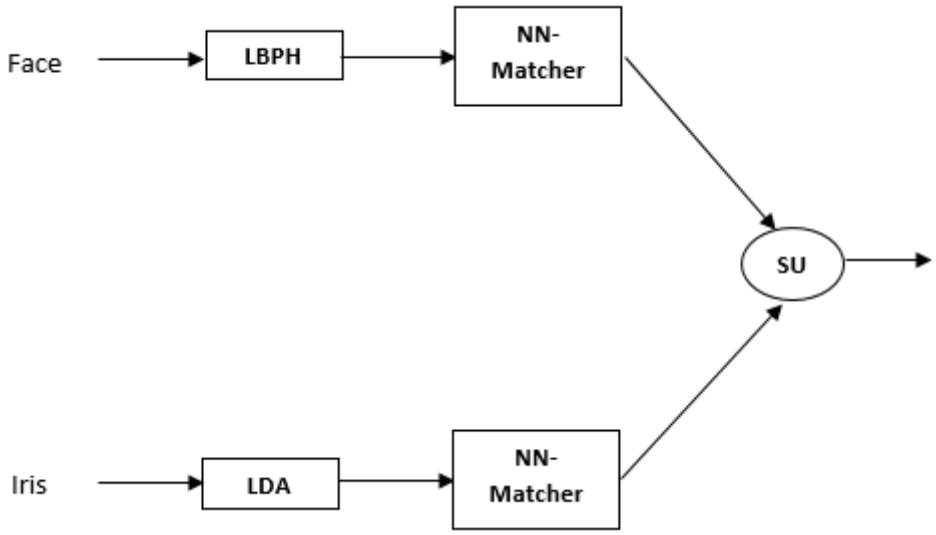


Figure 4.3: Score level fusion LBPH face and LDA iris

and Toygar, 2014)

$$WS = W_F \times S_F + W_I \times S_I, \ where \ W_F + W_I = 1 \tag{4.2.2}$$

Where W_F and W_I represents the reliability of each template towards the fusion process. The weights W is evaluated as follows:

$$W = \frac{\frac{1}{EER_i}}{\sum_{i=1}^{N} \frac{1}{EER_i}} \ \text{for } i = 1, 2, \cdots, N \tag{4.2.3}$$

Where N represents the number of matchers. In this case $N = 2$, the first weight W_F for face is computed with EER obtained for the LBPH matcher (see Table 3.2) and that of iris W_I is computed using EER obtained form LDA matcher. Prior to combining the scores, each score set from both LBPH and LDA matchers are rescaled using $Tanh$ normalisation discussed in section 2.2.1.1

4.2.3 Decision level fusion

In order to merge the three classifiers generated by feature and score level fusion, a majority voting decision rule is used to combine the class labels obtained from the classifiers. It picks the class label with highest median rating as the claimed identity.

4.3 Experimental results

In this section, results based on the proposed hybrid fusion scheme discussed above are presented. The face and iris datasets used in section 3.6 are used to construct a chimeric multi-modal dataset of 200 users to validate the hybrid fusion scheme. The metrics for the training and test data is similar to that of unimodal systems discussed in section 3.6. The performance of the system has been expressed in terms of Recognition Rate (RR %)and Equal Error Rate (EER) as shown Table 4.1.

Table 4.1: Recognition rates for multi-modal systems.

Method	RR (%)
Face-fused vector (Face-FV)	79.14
Iris-fused vector (Iris-FV)	82.56
Weighted-Score fusion (W-score)	94
Proposed hybrid scheme	**96.3**

The proposed method outperforms any other feature or score level fusion techniques because it combines information from three different classifiers using the voting method to make the final decision. This is in constrast to other methods which make a decision based on either the feature level fusion or score level fusion. The proposed system takes advantage of the rich information inherent in feature fusion and ability to combine modalities based on noise level through weighted score fusion to enhance the overall performance of the system. More so, the experiment above indicates that fusing feature vectors with face and iris does not necessarily improve performance of unimodal systems. This is evidenced by comparing the recognition rates obtained for face and iris-fused vector in Table 4.1 and with top performers like LBPH and LDA in Table 3.1

Table 4.2: Recognition rates for multi-modal systems.

	Error rates for multi-biomtric systems	
Method	EER	min TER
Face-fused vector (Face-FV)	0.104	0.208
Iris-fused vector (Iris-FV)	0.088	0.176
Weighted-Score fusion (W-score)	0.022	0.044
Proposed hybrid scheme	**0.017**	**0.034**

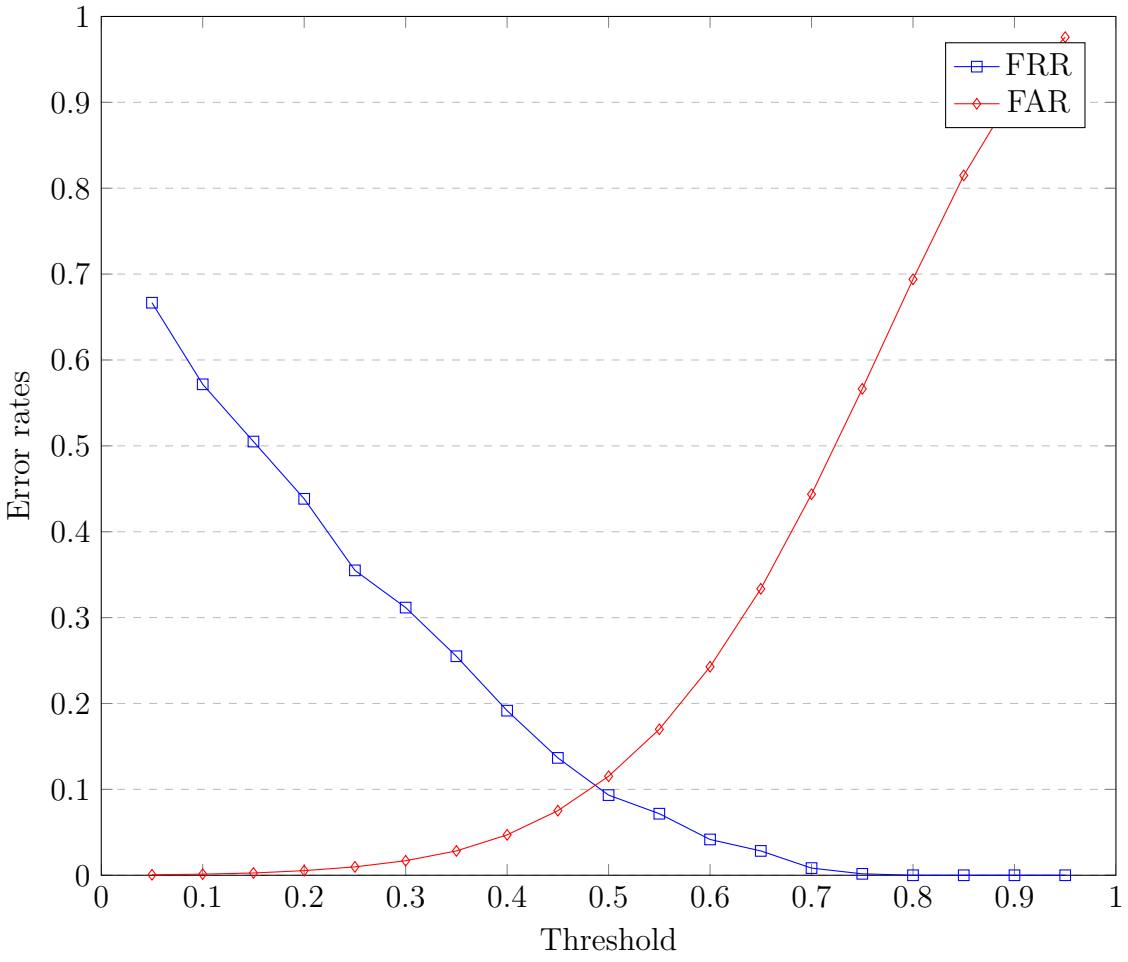


Figure 4.4: Plot of FRR and FAR against different system operating points Face-FV vector

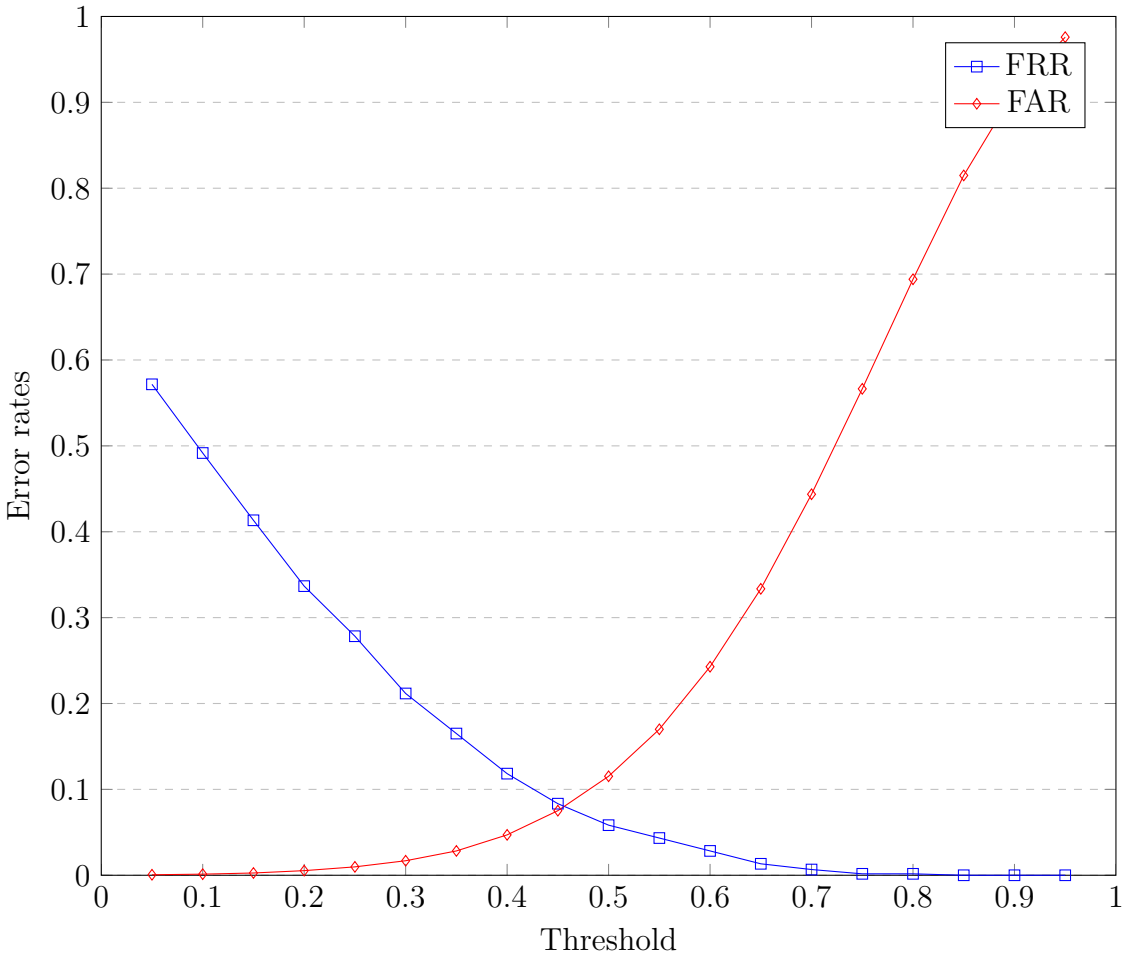


Figure 4.5: Plot of FRR and FAR against different system operating points Iris-FV vector

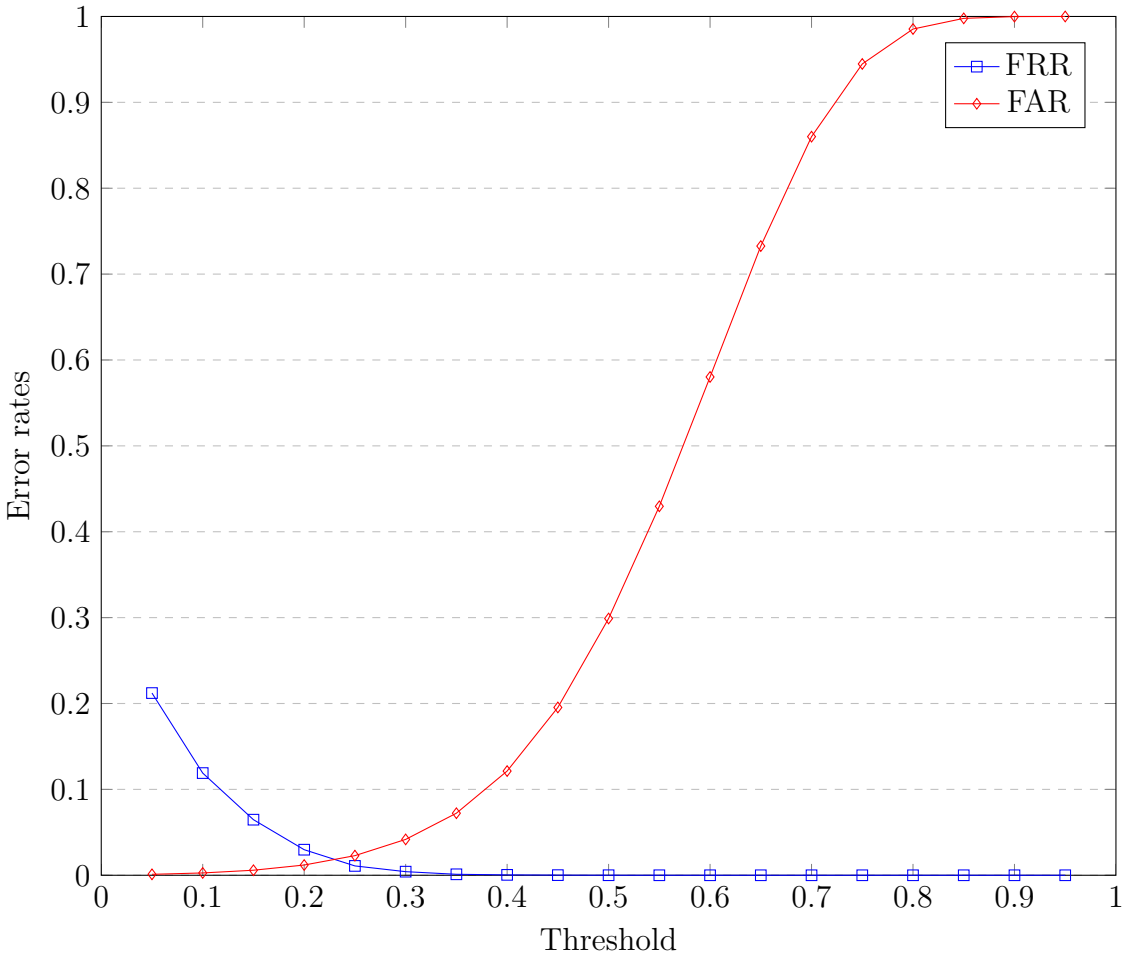


Figure 4.6: Plot of FRR and FAR against different system operating points proposed hybrid fusion

Table 4.2 shows the EER's and min TER's for the multi-biometric systems, as indicated the proposed hybrid scheme obtained the lowest EER as compared other multi-biometric systems, exhibiting its superiority. Furthermore, Figure 4.4, 4.5 and 4.6 shows the plot of FRR and FAR against the system operating threshold, the point were the two graphs intersect defines the EER of the Face-FV, Iris-FV and proposed hybrid fusion scheme. However, the point intersection for the proposed hybrid fusion scheme (Figure 4.6) is lower than that of Face-FV (Figure 4.4) and Iris-FV (Figure 4.5), showing that both FRR and FAR were reduced as compared to Face-FV and Iris-FV. The ROC curves for the multi-biometric systems presented above are also given in Figure 4.7. As seen the area under the curve of the proposed system is smaller than the curve of any other method, showing that proposed scheme gives the best performance at different operating points of the system. It should be noted that the recognition of the proposed algorithm is performed offline in the sense that output (including feature template, scores and class labels) of every algorithm combined have been stored in the system and only retrieved by the proposed algorithm upon request. This takes care of any time concerns on the proposed algorithm.

Finally, the performance of the proposed hybrid fusion scheme is compared with the other fusion schemes as seen in Figure 4.8 with the proposed giving the best performance in term recognition accuracy and error rate over these fusion schemes because it combines three levels of fusion to determine the claimed identity, while other methods use a single or at most double fusion rule. It should be noted that fusion compared here has been implemented to compare with the proposed scheme due the variation in database size used in this research compared to the original research papers.

Table 4.3: Comparison of recognition rates for state-of-the-art fusion schemes with proposed scheme.

Method	RR (%)
Sum-rule (Fakhar et al., 2011)	94.95
Product-rule Eskandari et al. (2014)	94.86
Weighted-sum rule Sim et al. (2014)	93.93
triangular-norm rule Hanmandlu et al. (2011)	94.75
Feature-score hybrid Farmanbar and Toygar (2015b)	91.58
Proposed hybrid fusion scheme	**96.3**

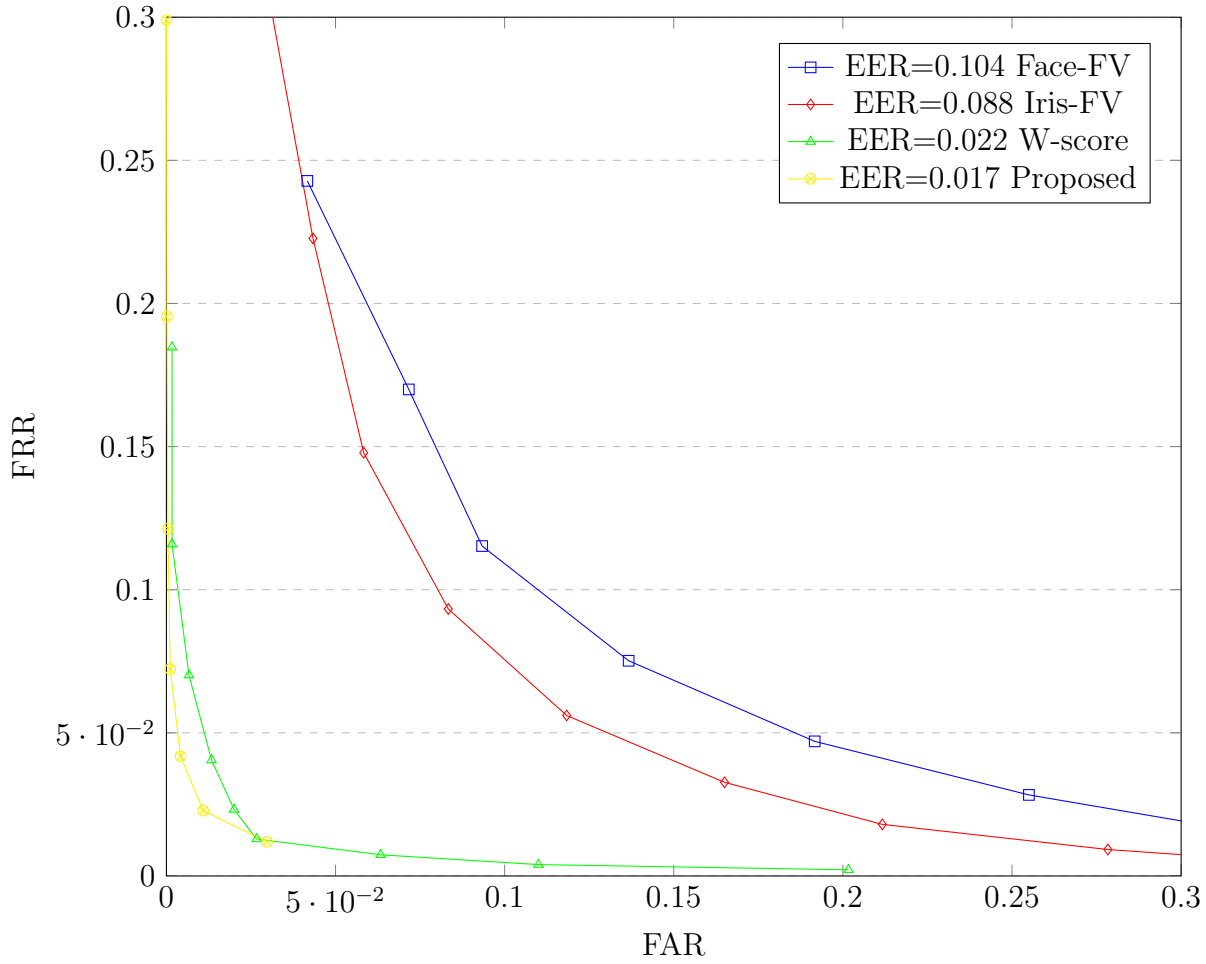


Figure 4.7: ROC curves for Multi-biometric systems

4.4 Statistical Analysis

The aim of the statistical analysis is formally conclude at 95% confidence level ($\alpha = 0.05$) that the proposed hybrid fusion scheme performs better than some fusion schemes in literature as shown in Table 4.3. Here, a paired t-test is used as the number of samples is less than 30. Furthermore, the selection of training and test set have been randomized for each simulation of the fusion schemes and results shown in table 4.4.

(a) **Parameters**

proposed hybrid fusion: $\overline{X_1} = 96.3\%$, $n_1 = 10$, $S_1 = 0.106$

sum rule, product rule, Weighted sum rule, T-norms rule, Feature-score hybrid : $\overline{X_2} = 94.95, 94.86, 93.93, 94.75, 91.58$, $n_2 = 10$,

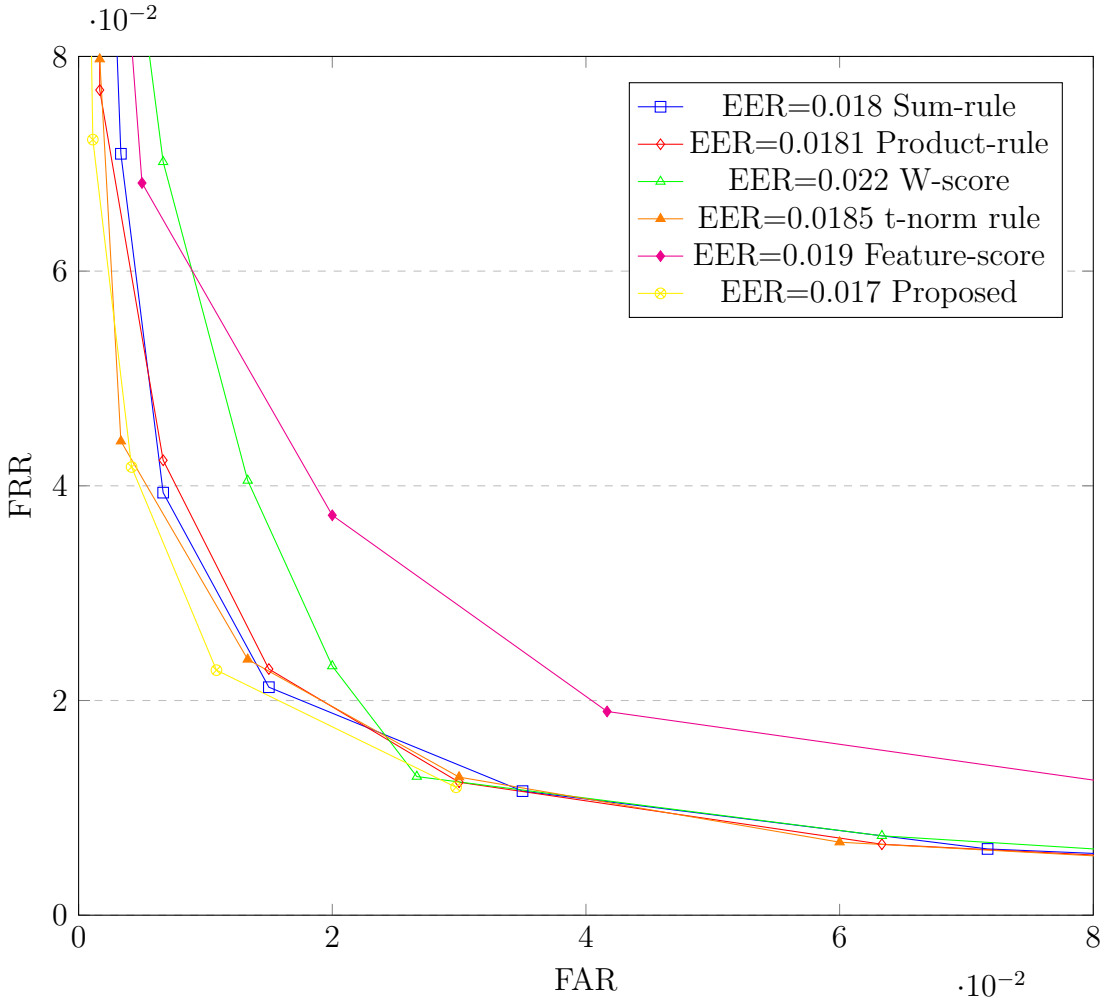


Figure 4.8: ROC curves comparing the proposed scheme with other fusion schemes

$S_2 = 0.563, 0.625, 0.991, 0.715, 1.827$

Where $\overline{X_1}$ and S_1 represents the average and variance of the recognition rate for the proposed hybrid fusion scheme.

$\overline{X_2}$ and S_2 represents the average and variance of the recognition rate for any sum rule, product rule, Weighted sum rule, T-norms rule, Feature-score hybrid

(b) **Hypothesis**

Let H_0 define the null hypothesis and H_1 define the alternate hypothesis, where H_0 as the measure of changing from any of the fusion rule in T 4.3 to proposed hybrid fusion scheme, and H_1 is what is expected to be true, if the null hypothesis does not hold.

$H_0 : \overline{X_1} = \overline{X_2}$

$H_1 : \overline{X_1} > \overline{X_2}$

Table 4.4: Recognition Rate for proposed hybrid fusion and other methods over 10 runs

trials	Proposed(%)	Sum rule(%)	product rule(%)	Weighted-sum(%)	T-norms(%)	Feature-score(%)
1	96.63	96.37	96.2	96.04	96.28	94.43
2	96.17	94.10	94.09	93.69	94.15	90.11
3	96.87	95.69	95.69	94.75	95.6	93.33
4	96.65	95.19	95.36	93.91	95.08	91.17
5	96.07	94.51	94.33	93.51	94.41	90.16
6	95.88	94.16	94.08	93.07	94.11	91.04
7	95.96	94.38	94.19	93.53	94.09	91.51
8	96.14	94.48	94.14	92.44	93.52	91.23
9	96.25	95.52	95.51	94.58	95.39	90.95
10	96.38	95.15	94.98	93.80	94.90	91.89
Mean						
1	96.3	94.95	94.86	93.93	94.75	91.58
Variance						
1	0.106	0.563	0.625	0.991	0.715	1.827

Table 4.5: Result of statistical analysis

values	Proposed vs Sum rule	Proposed vs product rule	Proposed vs Weighted-sum	Proposed vs T-norms	Proposed vs Feature-score
p-value	9.93×10^{-5}	7.46×10^{-6}	3.45×10^{-6}	1.33×10^{-5}	2.079×10^{-5}
t-stat	8.109	8.404	9.235	7.813	12.896

H_1 is true iff, $\overline{X_1} - \overline{X_2} > 0$

(c) **Calculation of t-statistic**

The t-statistic is calculated to test the above hypothesis. The formula is given by:

$$T = \frac{\overline{d}}{\frac{S_D}{\sqrt{n}}} \tag{4.4.1}$$

where $\overline{d}$ represents mean of the difference $X_1 - X_2$ and S_D is the standard deviation of the difference $X_1 - X_2$

From Table 4.5 it is seen the following *p-values* calculated by comparing the proposed hybrid fusion scheme with sum, product, weighted-sum, t-norm and feature-score hybrid fusion rule is less than $\alpha = 0.05$ (confidence level)in each case. Based on this criteria, the null hypothesis H_0 is rejected, therefore it can be concluded that at 95% confidence level that results produced by the proposed hybrid fusion scheme is statistically significant.

4.5 Summary

This chapter was aimed at discussing the hybrid fusion methodology at feature, score and decision level. The hybrid fusion involved a combination of classifiers based on feature and score level fusion with a decision level fusion scheme. The performance of the system was evaluated and it was revealed that proposed system obtained improved performance as compared to its unimodal systems and other fusion methods studied in literature.

Chapter 5

Serial fusion using BPSO and iRVM

5.1 Introduction

In literature, a fusion of multi-biometric systems is often implemented in parallel mode that is two or more modalities are processed simultaneously and then combined either at feature, score, decision or rank level of fusion. The acquisition time for multi-modal systems during identification is sometimes not favorable in terms user convenience as each modality will have to be captured before recognition whereas on the other hand, serial fusion is a cascaded approach that processes modalities in sequence. The advantage of this fusion approach over the parallel mode is that the total number of modalities used before the next modality is reduced which in turn provides a balance between verification time, performance and acceptability (Ramadan et al., 2015). Although, this fusion mode has been poorly investigated literature, some work in this area show promising results for further research (Marcialis et al., 2009; Zhang et al., 2014).

It is generally assumed in biometrics that the training data is available all at once and might not change with time, leading to high recognition rates. However, according to Mehrotra (2014) these assumptions are faulted due the following reasons:

- The training data may increase with time as users are enrolled into the system in batches.
- Data classification in multi-biometrics involves making a binary decision between genuine and impostor score sets. However, biometric training data have a high rate of class imbalance between the genuine and the impostor scores. In biometrics, the number of genuine scores is $O(n)$ (n is the number of users) is always under-represented when compared to the number of impostor scores that is $O(n^2)$ as seen in Figure 1.5. Large class imbalance is said to occur in training data when the imbalance ratio IR is greater than 100 (Garcia and

Herrera, 2009) :

$$IR = \frac{N^-}{N^+} \tag{5.1.1}$$

Where N^- represents the number of negative samples in this study impostor scores, while N^+ represents number of positive samples in this study genuine scores.

A serial fusion methodology using BPSO and Incremental Relevance Vector Machines (iRVM) is proposed in this chapter (Kennedy and Eberhart, 1997; Mehrotra, 2014; Arasomwan and Adewumi, 2014). Here, the serial fusion mechanism reduces the number impostors scores prior to the selection of the second modality, thereby reducing the effect of large class imbalance, while iRVM provides the capability to train biometric data in batches. Finally, BPSO is used for feature selection to find the best possible mix of local and global features that produce the best accuracy.

This layout of this chapter is presented as follows— Section 5.2 the architecture of proposed serial fusion is presented, along with feature selection techniques and iRVM formulation. Finally in Section 5.3, results from a simulation of the proposed scheme are presented in detail.

5.2 Proposed serial fusion scheme

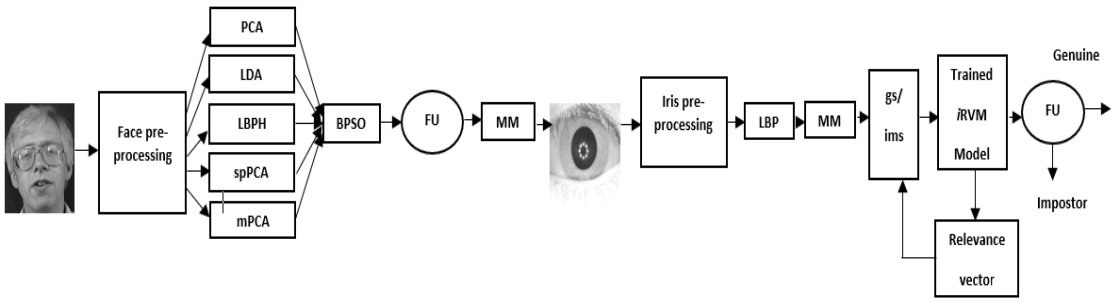


Figure 5.1: A serial fusion architecture for face and iris biometrics

Figure 5.1 shows the system structure for the proposed serial fusion scheme. First the face images are pre-processed using histogram equalisation to improve the contrast of the face images. Afterwards, features are extracted using global and local methods alluded to in chapter three. Performing feature fusion with feature vectors obtained each algorithm to form a fused feature vector of large dimension. Considering that some features might be redundant and not contributing to the recognition accuracy, BPSO is employed to select the optimal mix of local and global features that will

provide the best accuracy. The selected features are combined at feature level and a euclidean distance is used to obtain match scores. The iris images corresponding to the top-k matchers are selected and pre-processed as in section 3.5 and LBP feature extractor is used to obtain the feature vectors for the iris images. The genuine and impostor scores are computed and they are passed as features to trained by the iRVM classifier. The iRVM classifier makes a binary decision either to accept user as genuine or reject the user an impostor. It also provides the Relevance Vectors(RVs) which concatenated with data of the next training batch. Below is a description of different modules that make up the proposed system architecture.

5.2.1 Feature selection with BPSO

PSO was first presented by Kennedy and Eberhart in 1995 (Kennedy and Eberhart, 1995). It is a nature inspired algorithm that mimics the social behaviour of birds in a flock. Each particle (representing a bird in flock) is considered as point in n-dimensional feature space. The i^{th} particle is represented as $X_i = \{x_{i1}, x_{i1}, \cdots, x_{id}\}$ where d represents the number of feature extraction algorithms used. Each particle has the memory to hold its best position visited during the search process and its denoted as p_{best} (the position with best fitness). The particle's position amongst all the population that has the best fitness is denoted as global best g_{best}. The velocity of the i^{th} particle is denoted as $v_i = \{v_{i1}, v_{i1}, \cdots, v_{id}\}$ and it is updated at each iteration using equation 5.2.1

$$v_{id} = wv_{id} + c1 \times rand()(p_{id} - x_{id}) + c2 \times rand()(p_{gd} - x_{id}) \tag{5.2.1}$$

$$x_{id} = \begin{cases} 1, & \text{if } \frac{1}{1+\exp^{-v_{id}}} > c1() \\ 0, & \text{otherwise.} \end{cases} \tag{5.2.2}$$

Where w represents the inertia weight and set 1 to balance global and local exploration. Accelerations $c1$ and $c2$ represents stochastically the pull of each particle toward the global and local best positions, these values has been set as 2. rand() is a random value uniformly distributed in the interval $[0, 1]$. At the end of iteration, if $x_{id} = 1$ then all the features for the d_{th} feature extraction algorithm are chosen, otherwise it not chosen as shown in Figure 5.2.

10 particles have chosen with 20 iterations performed. Here, the fitness function is

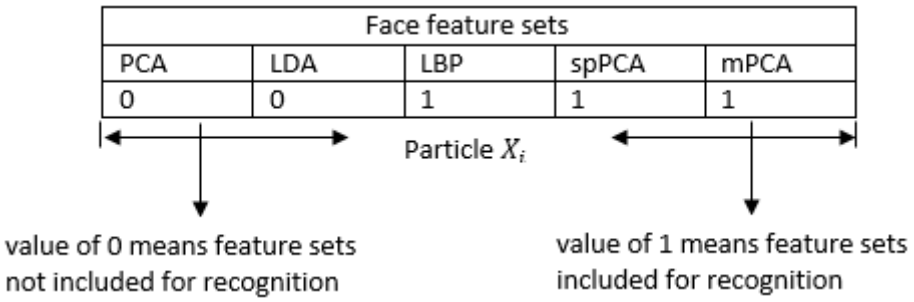


Figure 5.2: An example PSO bit particle

to maximize the rank-k identification rate as in equation 5.2.3

$$E = \frac{R_k}{N} \tag{5.2.3}$$

where E is the rank-k recognition rate, R_k is the number of correctly matched labels in the selected k-top matchers and N refers to the number of images in the database. Once the optimal features have been determined, feature level fusion is performed using technique outlined in section 4.2.1 to obtain a fused feature vector. Below in algorithm 1 shows pseudo-code for how BPSO is applied for feature selection.

Algorithm 1 BPSO algorithm

for each i : P intialize population of swarm along with their velocity and position **do**
    Fitness = Rank-k recognition rate for X_i from eq 5.2.3
    if Fitness > BestFitness of the i^{th} particle
    Update the BestFitness of the i^{th} particle
    $P_i = X_i$
    if Fitness > GlobalBestFitness
    Update the GlobalBestFitness
    $P_g = X_i$
    for $j = i : d$ **do**
        Update v_{id} using eq 5.2.1
        Update x_{id} using eq 5.2.2
    end for
end for

5.2.2 Relevance Vector Machines (RVM)

In this section, the RVM algorithms are presented for a classification problem, which forms building blocks for iRVM that will be discussed in the next section. RVM model which provides sparse solutions comparable to SVM using fewer basis func-

tions was first proposed by (Tipping, 2001). The model exploits Bayesian learning framework to provide accurate probabilistic predictions on classification problems. Although SVM is considered as an advanced technique for regression and classification, it still suffers from the following demerits (Tipping, 2001):

- Predictions made by SVM are not probabilistic, because it outputs a binary value for classification. Considering that conditional probability provides information on the uncertainty of predictions made by the classifier
- The number of support vectors is quite large, which increase as the size of dataset increases.
- SVM requires tunning of the regularization parameter "C" during the training phase.
- The kernel function must pass the test of Mercer's condition.

On the other hand, RVM is a fully probabilistic classifier which introduces prior probabilities over each weight governed by a set of hyperparameters. Due to its sparse nature, the number of relevance vectors required to perform classification is relatively small, hence reducing the testing time. RVM also requires fewer parameters to be optimized during the training phase.

According to Tipping (2001) RVM is of the form:

$$y(x, w) = w^T \phi(x) \tag{5.2.4}$$

where x is the input vector, w is the vector of weights,$\phi(x)$ is a set of basis functions and y is the output. RVM is a probabilistic classifier that does not suffer from the limitation of SVM listed above according to (Tipping, 2001). It considers weights with non-zero values and these weights are called the relevance vectors. RVM is perceived to perform better than SVM in terms of classification as it uses fewer kernel functions.

Given an input-target pair $\{x_n, t_n\}_n^N$, a two class Bernoulli distribution is required to predict the posterior probability of class labels given the input x. The generalization of the linear model is performed by using a sigmoid function $\sigma(y) = 1/(1 + e^{-y})$ to $y(x)$,the likelihood ratio can be given as:

$$P(t|w) = \prod_{n=1}^{N} \sigma\{y(x_n, w)\}^{t_n} [1 - \sigma\{y(x_n, w)\}]^{1-t_n} \tag{5.2.5}$$

The aim of RVM is to estimate the most probable weight w_{MP} that maximises the likelihood ratio as defined in equation 5.2.5. This is achieved by using the laplacian approximation method. Then w_{MP} can be defined as (Tipping, 2001):

$$\Sigma = (\phi^T B\phi + A)^{-1} \tag{5.2.6}$$

$$w_{MP} = \Sigma\phi^T Bt \tag{5.2.7}$$

where $A = diag(\alpha_1, \alpha_2, \cdots, \alpha_n)$, and $B = diag(\beta_1, \beta_2, \cdots, \beta_n)$ are diagonal of matrices with α a vector of hyper-parameters and $\beta_n = \sigma\{y(x_n)\}[1 - \sigma\{y(x_n)\}]$. Σ is the posterior convariance matrix over weights centered at w_{MP}. The hyper-parameters α_i and β are updated as :

$$\begin{aligned} \alpha_i &= \frac{\gamma_i}{w^2_{MP(i)}} \\ \text{where } \gamma_i &\equiv 1 - \alpha_i\Sigma_{ii} \end{aligned} \tag{5.2.8}$$

where $w_{MP(i)}$ is the i^{th} posterior mean weight as computed from 5.2.7 and Σ_{ii} is the i^{th} diagonal of the posterior convariance matrix computed in 5.2.6, also β_i is updated as :

$$\beta = \frac{N - \sum_i \gamma_i}{\| t - \phi w \|^2} \tag{5.2.9}$$

The convergence criteria to stop the iteration is given below:

$$\delta = \sum_i \alpha_i^{n+1} - \alpha_i \tag{5.2.10}$$

Termination of the algorithm stops when $\delta < \delta_\tau$ where δ_τ is the threshold value. Finally to predict a new value x' , the output y is computed as : 5.2.11

$$y = w^T_{MP}\phi(x') \tag{5.2.11}$$

Below the algorithm for RVM is presented in algorithm 2.

5.2.3 Incremental Relevance Vector Machines (iRVM)

In real life applications, training data arrives in batches i.e as new users are captured into the system, both the old and new data is retrained to provide a model for recognition. Resultantly therefore, the training time will continue to increase as the number of users increase. RVM alone will not be suitable as matrix inversion become a costly operation with a large dataset.

Algorithm 2 RVM algorithm

Input the training set of scores vector s with dimension d and labels t
Generate the basis function $\phi = [\phi(x_1), \phi(x_2), \cdots, \phi(x_n)]$
initialize δ_τ, α and β
for $\delta > \delta_\tau$ **do**
    $A = diag(\alpha)$ and $B = diag(\beta)$
    $\Sigma = (\phi^T B \phi + A)^{-1}$
    $w_{MP} = \Sigma \phi^T B t$
    $\gamma_i \equiv 1 - \alpha_i \Sigma_{ii}$
    $\alpha_i = \frac{\gamma_i}{w^2_{MP(i)}}$
    $\beta = \frac{N - \sum_i \gamma_i}{\|t - \phi w\|^2}$
    $\delta = \sum_i \alpha_i^{n+1} - \alpha_i$
end for
Output relevance vectors $R = x(w_{MP(index)})$

When a new batch arrives, *i*RVM uses only the relevance vectors obtained from the training phase of the old batch, to be trained along with the new batch (Mehrotra, 2014). This approach is built on the assumption that only relevance vectors required to make an accurate prediction and other training samples from the old batch can be safely removed (Tipping, 2001). The concatenation of relevance vectors with the new training batch helps to improve the decision boundary and avoids over-fitting. a step by step process is given below:

- *Step1 :* Use RVM to train the initial data T_1 and obtain the relevance vectors RV_1
- *Step2 :* When a new training batch T_2 arrives, the relevance vectors RV_1 obtained from the last training phase is concatenated with T_2 to form the new training data $(T_2 \oplus RV_1)$. Where $\oplus$ represents concatenation.
- *Step 3:* The operation RVM-Train$((T_2 \oplus RV_1))$ is carried out on the new dataset to obtain the new relevance vector RV_i with weights w_i.
- *Step 4:* The above steps are repeated for training batch T_i for $i = 1, 2, 3, 4, \cdots, m$ where m is the number of incremental training batches. See Figure 5.3

The *i*RVM model can be described as :

$$y = w^T_{MP} \phi(T_m \oplus RV_{m-1}) \tag{5.2.12}$$

The algorithm for *i*RVM is presented in algorithm 3.

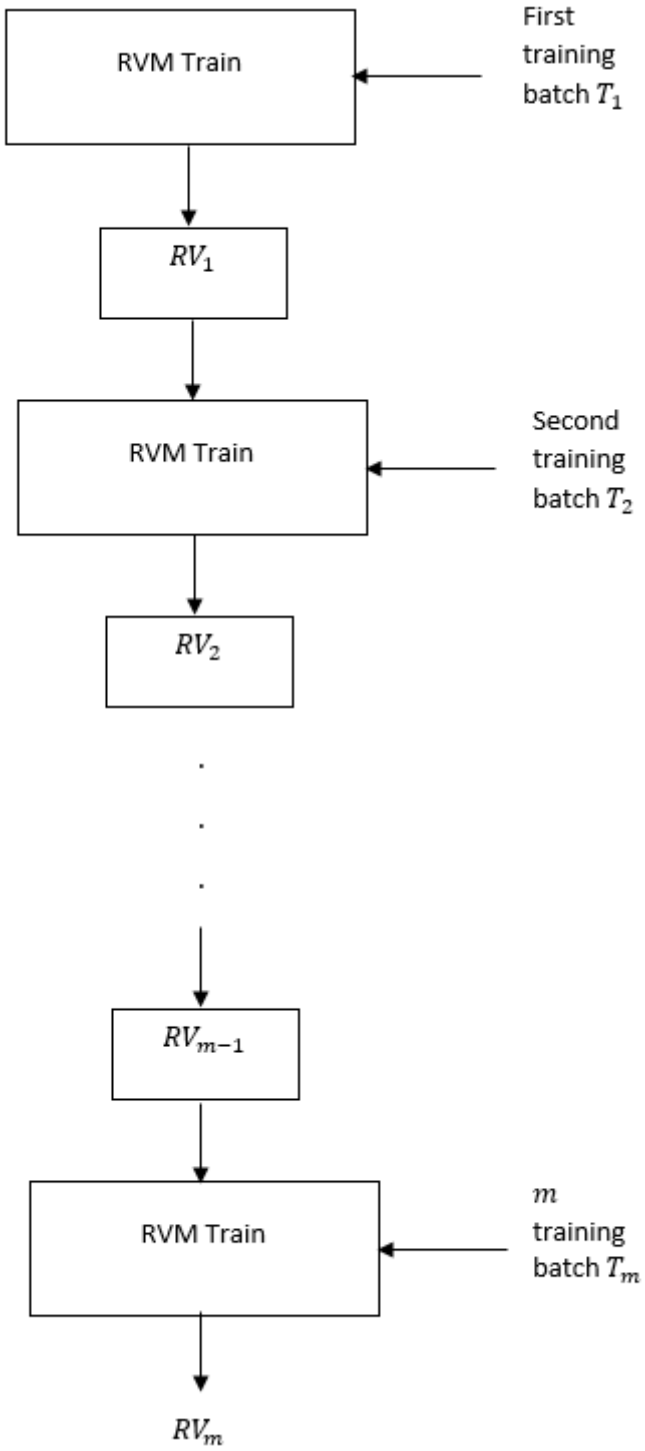


Figure 5.3: Block diagram showing training process for iRVM, RV is Relevance Vectors

Algorithm 3 *i*RVM algorithm

1: **Input** the training set batch T_m with data labels t_m
2: **for each** batch m of the training set **do**
3: $T_m = T_m \oplus R_{m-1}$ // *concatenate new training batch with RV's from the previous training batch*
4: $R_m = \text{RVM}(T_m, t_m)$
5: **end for**
6: **Output** relevance vectors R_m

5.3 Experimental results

This section was aimed at presenting the results obtained through simulation. The same face and iris dataset used in chapter three were used to validate the proposed scheme. The genuine and impostor scores that were generated and passed to the *i*RVM classifier, Table 5.1 shows the individual number of genuine, impostor scores and IR before serial fusion is applied for face and iris modality. It is revealed that there is a large class imbalance between the genuine and impostor scores, therefore, classification based on these scores is likely to favor the impostor class. The table shows the distribution of the genuine and impostor scores after the iris images corresponding to the top-k face matchers have been selected, these values given for different values of k. It also seen at $k = 100$ all the genuine scores are selected with reduced number of impostor scores with IR of 50 which less than 100 (required for large class imbalance) as in equation 5.1.1. The scores in Table 5.2 are divided into test and training data, here 67% of the scores have been used for training and rest for test as seen Table 5.3.

Table 5.1: Distribution genuine and imposter scores in parallel mode

Modality	Genuine scores	Imposter scores	IR
face	600	119400	199
iris	600	119400	199

Table 5.2: Distribution genuine and imposter scores in serial mode

Modality	k top-matchers	Genuine scores	Impostor scores	IR
iris	25	597	8892	15
iris	50	597	16947	28
iris	75	597	23943	40
iris	100	600	30297	50

Table 5.3: Training and test set for face and iris scores

		Training set		Testing set	
Method	k top-matchers	genuine	impostor	genuine	impostor
Iris	25	400	5958	197	2934
Iris	50	400	11354	197	5593
Iris	75	400	16041	197	7902
Iris	100	402	20299	198	9998

Table 5.4: Recognition rate for all k top-matchers

k top-matchers	RR (%)
25	96.78
50	98.82
75	99
100	**99.06**

Simulation using the above training and test set has been performed with three times random cross validation. Futhermore, the training set has been divided into ten batches to perform incremental learning with iRVM. The performance of the proposed serial fusion scheme based on the top-matchers selected are shown Table 5.4, obviously the best accuracy is obtained for $k = 100$ as contains all the genuine scores needed for classification. Figure 5.4 shows the ROC curves for different top-matchers and again the lowest EER is obtained when $k = 100$. The performance of the proposed scheme with an iRVM classifier is compared with using SVM and RVM classifiers (Vapnik, 1999; Tipping, 2001). As revealed in Table 5.5, the proposed scheme with iRVM performs better than RVM classifier, while its ROC curve intersects with that of the SVM classifier, meaning that at some operating points the propsed scheme with iRVM performs better, while at some other point SVM classifier performs better. However, the training time required to fit a model is longer than that of SVM (RVM the longest) because of the costly matrix computations associated with RVM training process. iRVM takes less training time as compared to RVM because the data is trained in batches rather than being considered as a whole. Contrarily, SVM takes more time for making prediction as more SV's are used while iRVM and RVM takes lesser time for testing because few RV's are used for prediction. Finally, Figure 5.5 shows the ROC curve to compare the peformance of the proposed serial fusion with iRVM and SVM classifier with the lowest EER obtained for the iRVM classifier.

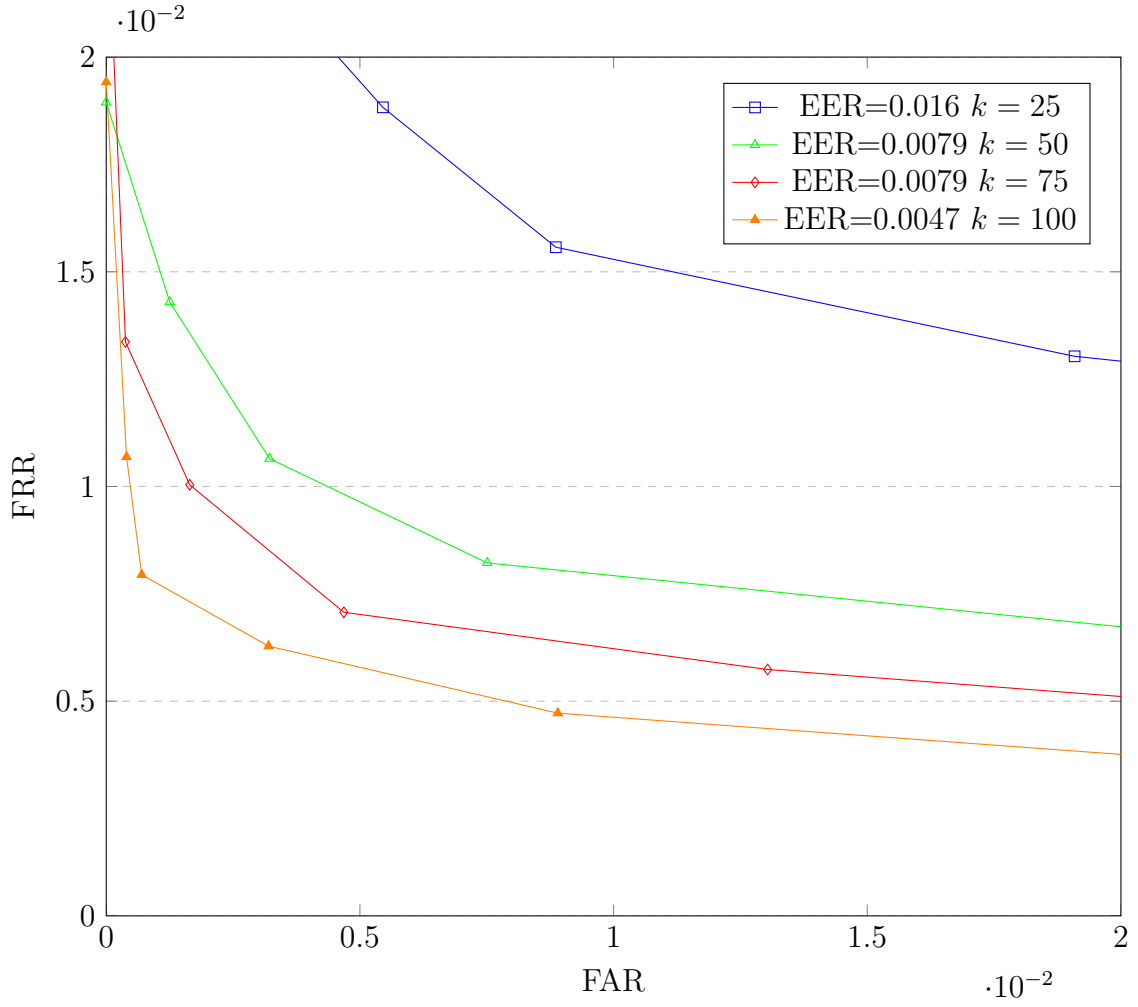


Figure 5.4: ROC for different k top-matchers of the proposed system

Table 5.5: Time taken for different learning classifiers

Approach	RR(%)	Training time (secs)	Testing time (secs)	SV's\RV's
SVM	99	272	5.37	471
RVM	98.25	1892	0.026	20
proposed	**99.06**	386	0.025	20

5.4 Statistical Analysis

The aim of the statistical analysis is formally conclude at 95% confidence level ($\alpha = 0.05$) that the proposed serial fusion scheme outperforms other classifiers in literature as shown in Table 5.5. Here, a paired t-test is used as the number of samples is less than 30. The EER has been used as the metric to perform this test.

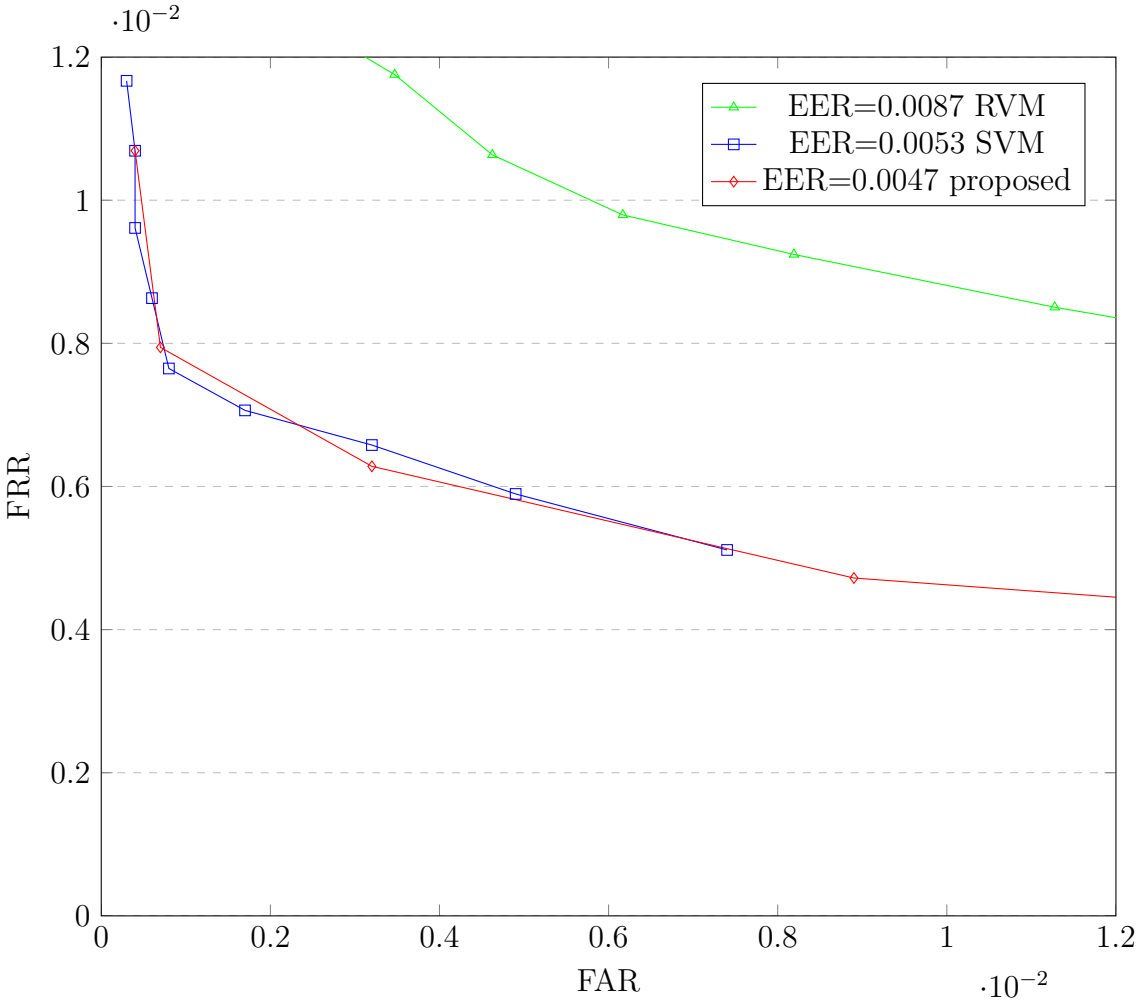


Figure 5.5: Comparsion of iRVM with SVM classifiers

(a) **Parameters**

SVM and RVM : $\overline{X_1} = 0.00545, 94.86$, $n_2 = 10$,

$S_2 = 2.05 \times 10^{-8}, 5.38 \times 10^{-8}$

proposed serial fusion scheme: $\overline{X_1} = 0.0048$, $n_1 = 10$, $S_1 = 6.88 \times 10^{-8}$.

$\overline{X_1}$ and S_1 represents the average and variance of the EER for SVM and RVM classifiers.

Where $\overline{X_2}$ and S_2 represents the average and variance of the EER for the proposed serial scheme.

(b) **Hypothesis**

Let H_0 define the null hypothesis and H_1 define the alternate hypothesis, where H_0 as the measure of changing from proposed serial fusion scheme to fusion

Table 5.6: Equal Error Rate for proposed hybrid fusion and other methods over 10 runs

trials	Proposed	SVM	RVM
1	0.0047	0.0053	0.0087
2	0.0047	0.0056	0.0087
3	0.0043	0.0053	0.0085
4	0.0046	0.0053	0.0088
5	0.0047	0.0055	0.0087
6	0.0048	0.0056	0.0087
7	0.0046	0.0053	0.0081
8	0.0047	0.0056	0.009
9	0.0051	0.0054	0.0086
10	0.0047	0.0056	0.0087
Mean			
1	0.00469	0.00545	0.00865
Variance			
1	6.88×10^{-8}	2.05×10^{-8}	5.38×10^{-8}

Table 5.7: Result of statistical analysis

values	SVM vs proposed	RVM vs proposed
p-value	3.99×10^{-7}	2.66×10^{-12}
t-stat	11.950	46.103

using SVM or RVM classifier , and H_1 is what is expected to be true, if the null hypothesis does not hold.

$H_0 : \overline{X_1} \leq \overline{X_2}$

$H_1 : \overline{X_1} > \overline{X_2}$

H_1 is true iff, $\overline{X_1} - \overline{X_2} > 0$

(c) **Calculation of t-statistic**

The t-statistic is calculated to test the above hypothesis using equation 4.4.1.

From Table 5.7 shows the results for the pair t-test for SVM and RVM classifiers against proposed serial fusion scheme. This comparison has been done in the reverse case as compared to section 4.4 because the EER is used as a metric for the test rather than the recognition rate. The *p-values* obtained shows that it is less than $\alpha = 0.05$ for both cases, meaning that the null hypothesis is not accepted. This shows that EER obtained in both cases is greater than the EER of the proposed serial fusion, Therefore the results obtained for the proposed is statistical significant.

5.5 Summary

This chapter proposed a serial fusion methodology using BPSO and *i*RVM. First the optimal mix of local and global feature was obtained using the BPSO algorithm and the face recognition was performed to obtain the top-k matchers. The iris images corresponding with the top-k matchers are selected and used to obtain the genuine and impostor scores for classification with *i*RVM classifier. The peformance of the proposed scheme was compared with other classifiers like SVM and RVM. The best performance was obtained with the proposed fusion scheme using *i*RVM, however the training time was higher than that of the SVM classifier. Findings indicate that proposed scheme is capable of reducing the effects of imbalance data and changes in training data.

Chapter 6

Summary, Conclusions and Future work

6.1 Summary

A lot of research dedicated towards improving the accuracy of the single biometric system has been conducted. To achieve this, concerted efforts have been made to minimise the drawbacks associated with the single biometric system through fusion of more than one modality. In this research, work based on serial and parallel architectures and have been reviewed, alongside with fusion schemes employed. However, fusion schemes in parallel mode, with improved performance over its unimodal systems were given more attention. it was gathered that, fusion schemes under this mode have some limitations such as inconveniences caused to the user when capturing multiple biometric templates, an increase in processing time and imbalanced dataset. In order to address these challenges associated with parallel fusion, a serial fusion has been proposed because it allows sequential fusion of two or more matchers which provides a balance between processing time and performance of the biometric system. Nonetheless, promising as it may, this fusion mode has not received adequate scholarly attention. Thus, two fusion schemes based on both the parallel and serial modes have been proposed in this study. First, a hybrid fusion scheme (parallel mode) that combines fusion at the feature, score and a decision level for face and iris modality. Results obtained from experiments performed on this scheme shows performance improvement over non-hybrid fusion schemes. Secondly, a serial fusion using BPSO and iRVM is presented, which reduces the problem of imbalance dataset during classification and allows for incremental learning on encountering new batch of data.

To measure the performance of the two techniques, three different datasets namely, ORL face, FERET face and CASIA iris have been employed. For the first set, local and global feature extraction methods for individual face and iris modality are used for experiment. Results revealed that local feature method provides the best accuracy of face recognition while the global feature methods give the best performance

for iris recognition. The next set of experiments were performed on the hybrid scheme with recognition accuracy of 96.34% and EER of 1.7% showing improved performance over non-hybrid methods on the same dataset. Lastly, simulation results on the proposed serial fusion obtained a recognition rate of 99.06% and EER of 0.47%, revealing improved performance over other fusion scheme alongside the proposed hybrid fusion scheme.

6.2 Conclusion

The urgent need for more research into serial fusion schemes cannot be over-emphasized as promising results have been obtained in previous works based on this fusion scheme. Furthermore, a hybrid method of the nature-inspired algorithm with machine learning techniques applied on this can improve performance. One of the two major issues affecting the performance of multi-biometric systems include the rapidly changing training data (leading to large training data) and class imbalance inherent in training data. Therefore, the need to develop adaptive fusion schemes that can handle changes in training data and reduce the effect of class imbalance inherent in the training data. In this research, two fusion schemes have been presented a hybrid fusion scheme in parallel mode and sequential fusion of face and iris modalities using BPSO and *i*RVM. The outcome from the hybrid fusion showed performance improvement over single fusion schemes. The serial fusion using a hybrid of nature-inspired algorithm and machine learning technique revealed that it was capable of handling changes in training data and reduce the effect of class imbalance. The results obtained from simulation showed improved performance over other fusion schemes.

6.3 Future work

There are different areas that show room for enhancing the accuracy of multi-biometric systems and during the course of this research, these areas have been identified and categorized into three groups, which are discussed below:

6.3.1 Machine learning approach

Machine learning methods have proved to be more effective when classifying a user as genuine or an impostor for biometric systems. Academic work reviewed, shows that trained classifiers provide better performance than that transformation-based methods. However, this approach is mostly affected by class imbalance problem and change in training data size. In this research, serial fusion together with iRVM were used to tackle this problem. It was, however noticed that the training time for iRVM was slow and more so sets of experiments had to be conducted to select the best number of top-matchers for the iris template. This is not so favourable, as biometric systems are real-time applications. The imperative therefore will be investigating how the training time of iRVM can be reduced and propose techniques that automatically select the top-matchers for the iris image. Furthermore, it would be important to investigate the performance of multi-biometric systems using semi-supervised learning methods in the advent of unlabeled training data given that , supervised learning methods have mostly been used for classifying genuine users and impostors.

6.3.2 Hybrid fusion approach

In this study, a hybrid fusion at the feature, score, and decision level was investigated. However, there are little or no works reported on hybrid systems based on sensor-feature, sensor-score, sensor-feature-score and feature-score-rank level of fusion schemes. Research into this method will evaluate their performance against existing hybrid systems.

6.3.3 Fusion mode approach

Although serial fusion schemes have been poorly investigated in literature, it will be interesting to study their performance in both serial and parallel mode. Given that only a few studies based on this dual mode have been conducted, systems can combine the flexibility of serial based methods and rich information content of parallel based systems to improve performance rates.

References

Adewumi, A.O. and Arasomwan, A.M. Improved particle swarm optimization based on greedy and adaptive features. In *IEEE Symposium on Swarm Intelligence (SIS), 2014*, pages 1–6, Orlando, Florida, 2014.

Adewumi, A.O. and Arasomwan, A.M. Improved particle swarm optimizer with dynamically adjusted search space and velocity limits for global optimization. *International Journal on Artificial Intelligence Tools*, 24(5):1550017, 2015.

Ahonen, T., Hadid, A., and Pietikainen, M. Face recognition with local binary patterns. In *Computer Vision - ECCV 2004: 8th European Conference on Computer Vision*, pages 469–481. Prague, Czech Republic, 2004.

Altun, A.A., Kocer, H.E., and Allahverdi, N. Genetic algorithm based feature selection level fusion using fingerprint and iris biometrics. *International Journal of Pattern Recognition & Artificial Intelligence*, 22(3):585–600, 2008.

Aly, O.M., Salama, G.I., and Onsi, H.M. An adaptive multimodal biometrics system using pso. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 4(7):158–165, 2013.

Amioy, K. and Ajay, K. Adaptive management of multimodal biometrics fusion using ant colony optimization. *Information Fusion*, 10(16):1–15, 2015.

Arasomwan, A.M. and Adewumi, A.O. Improved particle swarm optimization with a collective local unimodal search for continuous optimization problems. *The Scientific World Journal*, 2014.

Asaari, M.S.M., Suandi, S.A., and Rosdi, B.A. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert Systems with Applications*, 41(7):3367–3382, 2014.

Azom, V., Adewumi, A.O., and Tapamo, J.R. Face and iris biometrics person identification using hybrid fusion at feature and score-level. In *Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*, pages 207–212, Port-Elizabeth, South-Africa, 2015.

Bailly-Bailliere, E., Bengio, S., Bimbot, F., Hamouz, M., Kittler, J., Mariethoz, J., Matas, J., Messer, K., Popovici, V., and Poree, F. The banca database and evaluation protocol. In *Audio and Video-Based Biometric Person Authentication: 4th International Conference*, pages 625–638, Berlin, Heidelberg, 2003.

Canuto, A.M., Pintro, F., and Fairhurst, M.C. Genetic algorithm and ensemble systems for multi-biometric cancellable recognition. *Journal of Biometrics and Its Applications*, 1(1):1–14, 2015.

Chin, Y.J., Ong, T.S., Teoh, A.B., and Goh, K. Integrated biometrics template protection technique based on fingerprint and palmprint feature-level fusion. *Information Fusion*, 18:161–174, 2014.

Connaughton, R., Bowyer, K.W., and Flynn, P.J. *Fusion of Face and Iris Biometrics from a Stand-Off Video Sensor*. PhD thesis, University of Notre Dame, 2011.

Cristianini, N. and Shawe, T.J. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.

Damousis, I.G. and Argyropoulos, S. Four machine learning algorithms for biometrics fusion: A comparative study. *Applied Computational Intelligence and Soft Computing*, 2012:8, 2012.

Daugman, J. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21–30, 2004.

Duin, R.P. The combining classifier: to train or not to train. In *16th International Conference on Pattern Recognition, 2002. Proceedings.*, volume 2, pages 765–770. IEEE, 2002.

Eskandari, M. and Toygar, O. Fusion of face and iris biometrics using local and global feature extraction methods. *Signal, Image and Video Processing*, 8(6): 995–1006, 2014.

Eskandari, M., Toygar, O., and Demirel, H. A new approach for face-iris multimodal biometric recognition using score fusion. *International Journal of Pattern Recognition and Artificial Intelligence*, 27(3):1356004, 2013.

Eskandari, M., Toygar, O., and Demirel, H. Feature extractor selection for face–iris multimodal recognition. *Signal, Image and Video Processing*, 8(6):1189–1198, 2014.

Fakhar, K., Aroussi, M.E., Saadane, R., Wahbi, M., and Aboutajdine, D. Fusion of face and iris features extraction based on steerable pyramid representation for multimodal biometrics. In *2011 International Conference on Multimedia Computing and Systems (ICMCS)*, pages 1–4, Ouarzazate, 2011.

Farmanbar, M. and Toygar, O. Feature selection for the fusion of face and palmprint biometrics. *Signal, Image and Video Processing*, pages 1–8, 2015a.

Farmanbar, M. and Toygar, O. A hybrid approach for person identification using palmprint and face biometrics. *International Journal of Pattern Recognition and Artificial Intelligence*, 29(6):1–15, 2015b.

Fathima, A.A., Vasuhi, S., Babu, N.N., Vaidehi, V., and Treesa, T.M. Fusion framework for multimodal biometric person authentication system. *IAENG International Journal of Computer Science*, 40(1):132–141, 2014.

Fridman, L., Stolerman, A., Acharya, S., Brennan, P., Juola, P., Greenstadt, R., and Kam, M. Multi-modal decision fusion for continuous authentication. *Computers & Electrical Engineering*, 41(2015):142 – 156, 2015.

Garcia, S. and Herrera, F. Evolutionary undersampling for classification with imbalanced datasets: Proposals and taxonomy. *Evolutionary Computation*, 17(3): 275–306, 2009.

Gawande, U., Zaveri, M., and Kapur, A. A novel algorithm for feature level fusion using svm classifier for multibiometrics-based person identification. *Applied Computational Intelligence and Soft Computing*, 2013:9, 2013.

Giot, R. and Rosenberger, C. Genetic programming for multibiometrics. *Expert Systems with Applications*, 39(2):1837–1847, 2012.

Giot, R., El-Abed, M., and Rosenberger, C. Fast learning for multibiometrics systems using genetic algorithms. In *High Performance Computing and Simulation (HPCS), 2010 International Conference on*, pages 266–273, Caen, France, 2010.

Gogoi, M. and Bhattacharyya, D.K. Fusion of fingerprint and iris biometrics using binary ant colony optimization. In *Proceedings of the Third International Conference on Soft Computing for Problem Solving*, pages 601–613, New Delhi, 2014.

Griaule, B. Biometric applications. available: http://www.griaulebiometrics.com/en-us/book/understanding-biometrics/introduction/applications-2/, 2014. Online; accessed 31 October 2015.

Hanmandlu, M., Grover, J., Gureja, A., and Gupta, H.M. Score level fusion of multimodal biometrics using triangular norms. *Pattern Recognition Letters*, 32 (14):1843–1850, 2011.

Islam, M. Feature and score fusion based multiple classifier selection for iris recognition. *Computational intelligence and neuroscience*, 2014:10, 2014.

Islam, M., Davies, R., Bennamoun, M., Owens, R.A., and Mian, A.S. Multibiometric human recognition using 3d ear and face features. *Pattern Recognition*, 46(3):613–627, 2013.

Jain, A. K., Flynn, P., and Ross, A.A. *Handbook of biometrics*. Springer Science & Business Media, 2007.

Jain, A.K., Ross, A., and Prabhakar, S. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

Jain, A.K., Nandakumar, K., and Ross, A. Score normalization in multimodal biometric systems. *Pattern recognition*, 38(12):2270–2285, 2005.

Kekre, H.B., Bharadi, V.A., Singh, V.I., Kaul, V., and Nemade, B. Hybrid multimodal biometric recognition using kekre's wavelets, 1d transforms & kekre's vector quantization algorithms based feature extraction of face & iris. In *2nd International Conference and workshop on Emerging Trends in Technology (ICWET)*, number 3, pages 29–34, 2011.

Kennedy, J. and Eberhart, R. Particle swarm optimization. In *IEEE International Conference on Neural Networks Proceedings.*, volume 4, pages 1942–1948, Perth, WA, 1995.

Kennedy, J. and Eberhart, R. A discrete binary version of the particle swarm algorithm. In *IEEE International Conference on Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation.*, volume 5, pages 4104–4108, Orlando, Florida, 1997.

Khalifa, A.B., Gazzah, S., and BenAmara, N.E. Multimodal biometric authentication using choquet integral and genetic algorithm. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 7(3), 2013.

Kisku, D.R., Singh, J.K., Tistarelli, M., and Gupta, P. Multisensor biometric evidence fusion for person authentication using wavelet decomposition and monotonic-decreasing graph. In *Seventh International Conference on Advances in Pattern Recognition*, pages 205–208, Kolkata, India, 2009.

Kisku, D.R., Rattani, A., Gupita, P., Singh, J.K., and Tisteralli, M. Biometrics sensor fusion. In Thomas, Ciza, editor, *Sensor Fusion and its Applications*, chapter 17, pages 395–401. 2010.

Kumar, A. Incorporating cohort information for reliable palmprint authentication. In *Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, pages 583–590, Bhubaneswar, India, 2008. IEEE.

Kumar, A. and Shekhar, S. Personal identification using multibiometrics rank-level fusion. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 41(5):743–752, 2011.

Kumar, A., Hanmandlu, M., and Vasikarla, S. Rank level integration of face based biometrics. In *Ninth International Conference on Information Technology: New Generations (ITNG)*, pages 36–41, Las Vegas, Nevada, 2012.

Kumari, P.A. and Suma, G.J. A novel multimodal biometric scheme for personal authentication. *International Journal of Research in Engineering & Technology*, 2(2), 2014.

Liu, W. and Lu, C. Face recognition based on rearranged modular 2dpca. In Huang, De-Shuang, Gan, Yong, Gupta, Phalguni, and Gromiha, M.Michael, editors, *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence*, volume 6839 of *Lecture Notes in Computer Science*, pages 395–403. Springer Berlin Heidelberg, 2012.

Marcialis, G.L. and Roli, F. Decision-level fusion of pca and lda-based face recognition algorithms. *International Journal of Image & Graphics*, 6(2):293 – 311, 2006.

Marcialis, G.L. and Roli, F. Serial fusion of fingerprint and face matchers. In Haindl, Michal, editor, *Multiple Classifier Systems: 7th International Workshop, MCS 2007*, chapter Serial Fusion of Fingerprint and Face Matchers, pages 151–160. Springer, Berlin, Heidelberg, 2007.

Marcialis, G.L., Roli, F., and Didaci, L. Personal identity verification by serial fusion of fingerprint and face matchers. *Pattern recognition*, 42(11):2807–2817, 2009.

Marcialis, G.L., Mastinu, P., and Roli, F. Serial fusion of multi-modal biometric systems. In *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS)*, pages 1–7, 2010.

Martinez, A.M. and Kak, A.C. Pca versus lda. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(2):228–233, 2001.

Mehrotra, H. *On the Performance Improvement of Iris Biometric System*. PhD thesis, National Institute of Technology Rourkela, 2014.

Mehrotra, H., Vatsa, M., Singh, R., and Majhi, B. Biometric match score fusion using rvm: A case study in multi-unit iris recognition. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 65–70, Providence, RI, 2012.

Messer, K., Matas, J., Kittler, J., Luettin, J., and Maitre, G. Xm2vtsdb: The extended m2vts database. In *Second international conference on audio and video-based biometric person authentication*, volume 964, pages 965–971, washington D.C, 1999.

Monwar, M.M., Vijayakumar, B.V., Boddeti, V.N., and Smereka, J.M. Rank information fusion for challenging ocular image recognition. In *12th IEEE International Conference on Cognitive Informatics Cognitive Computing (ICCI-CC)*, pages 175–181, New York, 2013.

Muhammad, I.A., Wai, L.W., and Satnam, D. Non-stationary feature fusion of face and palmprint multimodal biometrics. *Neurocomputing*, 177(2016):49–61, 2015.

Nandakumar, K. *Multibiometric systems: Fusion strategies and template security*. ProQuest, 2008.

Ojala, T., Pietikainen, M., and T.Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

Paul, P.P. and Gavrilova, M. Rank level fusion of multimodal cancelable biometrics. In *IEEE 13th International Conference on Cognitive Informatics Cognitive Computing (ICCI-CC)*, pages 80–87, London, 2014.

Phillips, P.J., Wechsler, H., Jeffery, H., and Rauss, P.J. The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(5):295–306, 1998.

Proenca, H., Filipe, S., Santos, R., Oliveira, J., and Alexandre, L.A. The UBIRIS.v2: A database of visible wavelength images captured on-the-move and at-a-distance. *IEEE Trans. PAMI*, 32(8):1529–1535, 2010.

Qian, T. and Veldhuis, R. Robust biometric score fusion by naive likelihood ratio via receiver operating characteristics. *IEEE Transactions on Information Forensics and Security*, 8(2):305–313, 2013.

Radha, N. and Kavitha, A. Rank level fusion using fingerprint and iris biometrics. *Indian Journal of Computer Science and Engineering*, 2(6):917–923, 2012.

Raghavendra, R., Rao, A., and K.G.Hemantha. Multisensor biometric evidence fusion of face and palmprint for person authentication using particle swarm optimisation (pso). *International Journal of Biometrics*, 2(1):19–33, 2009.

Raghavendra, R., Dorizzi, B., Rao, A., and Kumar, G.H. Particle swarm optimization based fusion of near infrared and visible images for improved face verification. *Pattern Recognition*, 44(2):401–411, 2011.

Ramachandra, A.C., Abhilash, S.K., Raja, K.B., Venugopal, K.R., and Patnaik, L.M. Dual transformation bimodal biometrics based on feature level fusion. In *Fourth International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom2012)*, pages 131–134, Bangalore, India, 2012.

Ramadan, G., Ayman, E., and E. Nawal, M. Zorkany. Multi-biometric systems: A state of the art survey and research directions. *International Journal of Advanced Computer Science and Applications*, (6):128–138, 2015.

Ramli, D.A., Rani, N.H., and Ishak, K.A. Performances of weighted sum-rule fusion scheme in multi-instance and multi-modal biometric system. *World Applied Sciences*, 12(11):2160–2167, 2011.

Rattani, A. and Tistarelli, M. Robust multi-modal and multi-unit feature level fusion of face and iris biometrics. In *Advances in Biometrics*, chapter Robust Multi-modal and Multi-unit Feature Level Fusion of Face and Iris Biometrics, pages 960–969. Springer, Berlin, Heidelberg, 2009.

Ross, A.A. and Govindarajan, R. Feature level fusion of hand and face biometrics. In *Proceedings of SPIE Conference on Biometric Technology for Human identification II*, pages 196–204, Orlando, USA, 2005.

Roy, K. and Kamel, M.S. Multibiometric system using level set method and particle swarm optimization. In *Image Analysis and Recognition*, chapter Multibiometric System Using Level Set Method and Particle Swarm Optimization, pages 20–29. Springer, Berlin, Heidelberg, 2012.

Saleh, I.A. and Alzoubiady, L.M. Decision level fusion of iris and signature biometrics for personal identification using ant colony optimization. *International Journal of Engineering and Innovative Technology (IJEIT)*, 3(11):35–42, 2014.

Sanderson, C. and Paliwal, K.K. Information fusion and person verification using speech and face information. *Research Paper IDIAP-RR, 02-33*, 2002.

Sim, H.M., Asmuni, H., Hassan, R., and Othman, R.M. Multimodal biometrics: Weighted score level fusion based on non-ideal iris and face images. *Expert Systems with Applications*, 41(11):5390–5404, 2014.

Sutra, G., Garcia, S.S., and Dorizzi, B. The viterbi algorithm at different resolutions for enhanced iris segmentation. In *5th IAPR International Conference on Biometrics (ICB)*, pages 310–316, New Delhi, 2012.

Talebi, H. and Gavrilova, M.L. Prior resemblance probability of users for multimodal biometrics rank fusion. In *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pages 1–7, Hong Kong, 2015.

Tan, T., He, Z., and Sun, Z. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *Image and vision computing*, 28(2):223–230, 2010.

Tao, Q. and Veldhuis, R. Hybrid fusion for biometrics: Combining score-level and decision-level fusion. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPRW'08.*, pages 1–6, Anchorage, AK, 2008.

Tao, Q. and Veldhuis, R. Threshold-optimized decision-level fusion and its application to biometrics. *Pattern Recognition*, 42(5):823–836, 2009.

Tieniu, T. and Zhenan, S. Casia database. avaliable http://biometrics.idealtest.org/dbDetailForUser.do?id=4, 2010. Online; accessed:31 October 2015.

Tipping, M.E. Sparse bayesian learning and the relevance vector machine. *The journal of machine learning research*, 1(2001):211–244, 2001.

Tran, L.B. and Le, T.H. Person authentication using relevance vector machine (rvm) for face and fingerprint. *International Journal of Modern Education and Computer Science*, 5:8–15, 2015.

Vapnik, V.N. An overview of statistical learning theory. *Neural Networks, IEEE Transactions on*, 10(5):988–999, 1999.

Vatsa, M., Singh, R., Ross, A., and Noore, A. Likelihood ratio in a svm framework: Fusing linear and non-linear face classifiers. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops CVPRW '08*, pages 1–6, Anchorage, AK, 2008.

Veeramachaneni, K., Osadciw, L., Ross, A., and Srinivas, N. Decision-level fusion strategies for correlated biometric classifiers. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops CVPRW'08.*, pages 1–6, Anchorage, AK, 2008.

Viola, P. and Jones, M.J. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.

Wanas, N. *Feature based architecture for decision fusion.* PhD thesis, University of Waterloo, 2003.

Wang, F. and Han, J. Multimodal biometric authentication based on score level fusion using support vector machine. *Opto-electronics review*, 17(1):59–64, 2009.

Wang, N., Li, Q., El-Latif, A., Xuehu, Y., and Niu, X. A novel hybrid multibiometrics based on the fusion of dual iris, visible and thermal face images. In *International Symposium on Biometrics and Security Technologies (ISBAST)*, pages 217–223, Chengdu, China, 2013a.

Wang, N., Li, Q., El-Latif, A.A., Peng, J., and Niu, X. Multibiometric complex fusion for visible and thermal face images. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 6(3):1–16, 2013b.

Wang, S., Liu, Z., Lv, S., Lv, Y., Wu, G., P.Peng, Chen, F., and X.Wang. A natural visible and infrared facial expression database for expression recognition and emotion inference. *Multimedia, IEEE Transactions on*, 12(7):682–691, 2010.

Wang, Z., Wang, E., Wang, S., and Ding, Q. Multimodal biometric system using face-iris fusion feature. *Journal of Computers*, 6(5):931–938, 2011.

Woo, N. and Kim, H. Multiple-biometric fusion methods using support vector machine and kernel fisher discriminant. In *Proceedings of the 6th International Conference on Recent Advances in Soft Computing (RASC 2006)*, pages 428–433, Canterbury, UK, 2006.

Yang, Y., Yang, G., and Wang, S. Finger vein recognition based on multi-instance. *International Journal of Digital Content Technology & its Applications*, 6(11): 86–94, 2012.

Yong, X., Lunke, F., and Zhang, D. Combining left and right palmprint images for more accurate personal identification. *IEEE Transactions on Image Processing*, 24(2):549–559, 2015.

Zhang, Q., Yin, Y., Zhan, D., and J.Peng. A novel serial multimodal biometrics framework based on semisupervised learning techniques. *Information Forensics and Security, IEEE Transactions on*, 9(10):1681–1694, 2014.