

Graduate Theses, Dissertations, and Problem Reports

2005

Generic multimodal biometric fusion

Yash Shah West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation

Shah, Yash, "Generic multimodal biometric fusion" (2005). *Graduate Theses, Dissertations, and Problem Reports.* 4191.

https://researchrepository.wvu.edu/etd/4191

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Generic Multimodal Biometric Fusion

by

Yash Shah

Thesis Submitted to the College of Engineering and Mineral Resources at West Virginia University in partial fulfillment of the requirements for the degree of

> Master of Science in Electrical Engineering

Natalia A. Schmid, D.SC., Chair Arun A. Ross Ph.D. Bojan Cukic, Ph.D. Lawrence Hornak, Ph.D.

Lane Department of Computer Science and Electrical Engineering

Morgantown, WV 2005

Keywords: Biometrics, Multimodal systems, Iris recognition, Feature Selection, Boosting, Support Vector Machines.

Copyright 2005 Yash Shah

ABSTRACT

Generic Multimodal Biometric Fusion

Yash Shah

Biometric systems utilize physiological or behavioral traits to automatically identify individuals. A unimodal biometric system utilizes only one source of biometric information and suffers from a variety of problems such as noisy data, intra-class variations, restricted degrees of freedom, non-universality, spoof attacks and unacceptable error rates. Multimodal biometrics refers to a system which utilizes multiple biometric information sources and can overcome some of the limitation of unimodal system. Biometric information can be combined at 4 different levels: (i) Raw data level; (ii) Feature level; (iii) Match-score level; and (iv) Decision level. Match score fusion and decision fusion have received significant attention due to convenient information representation and raw data fusion is extremely challenging due to large diversity of representation. Feature level fusion provides a good trade-off between fusion complexity and loss of information due to subsequent processing. This work presents generic feature information fusion techniques for fusion of most of the commonly used feature representation schemes. A novel concept of Local Distance Kernels is introduced to transform the available information into an arbitrary common distance space where they can be easily fused together. Also, a new dynamic learnable noise removal scheme based on thresholding is used to remove shot noise in the distance vectors. Finally we propose the use of AdaBoost and Support Vector Machines for learning the fusion rules to obtain highly reliable final matching scores from the transformed local distance vectors. The integration of the proposed methods leads to large performance improvement over matchscore or decision level fusion.

Table Of Contents

Chapter 1: Introduction	1
1.1 Biometric System	1
1.2 Unimodal Biometric Systems and Their Limitation	4
1.3 Multimodal Biometric System	7
1.3.1 Sources of information	8
1.3.2 Information Architecture	
1.3.1. Levels of fusion:	11
1.4 Thesis Contribution	
Chapter 2. Concris Information Fusion at Festure I avai	16
2.1 Mativation and Challenges	10
2.1 Mouvation and Chanenges	10
2.2 Generation of Multimodal Database	10
2.5 Fusion Framework	
2.4 Data Preprocessing	
2.4.1 Feature Transformation	
2.4.2 Feature Normalization	
2.4.3 Special Transformation	
2.5 Feature Selection	
2.6 Training and Testing Procedure	
2.7 Recognition Performance	
2.7.1 Implementation Methodology	
2.8 Summary	
Chapter 3: Boosting Biometric Features	55
3.1 Motivation and Challenges	55
3.2 Learning Boosted Function for Match-Score Generation	55
3.3 Training and Testing Architecture	
3.3.1 Intramodal Fusion	
3.3.2 Intermodal Fusion Error! Bookmark no	t defined.
Charten A. Sama art Vester Meshings for Fasion	74
Chapter 4: Support vector Machines for Fusion	
4.1 Motivation and Challenges	
4.2 Learning with Support Vector Machines	
4.3 Architecture of SVM Based Information Fusion	
4.3 Performance of SVM Fusion Scheme	
4.3.1 Intramodal Fusion	
4.3.2 Intermodal Fusion	
4.4 Summary	
Chapter 5: Summary and Future Work	

Chapter 1: Introduction

1.1 Biometric System

Our society has out grown the geographical boundaries and computers have facilitated this. Human computer interactions (HCI) have become an integral part of modern living. Identity management refers to the challenge of providing authorized users with secure and desired access to information and services across a variety of networked systems. Examples of such applications include physical access control to a secure facility, e-commerce, access to computer networks and welfare distribution. The primary task in an identity management system is the determination of an individual's identity. Most conventional modes of authentication are based on 'what we know' (e.g., passwords) and/or 'what we have' (e.g., ID card). These methods have sufficed for a while in the past but today our lives are deeply tied with digital information and we cannot rely completely on these surrogate representations of the identity as they can easily be lost, shared or stolen.

The concept of biometrics has existed for over a century, fueled by the diversity and randomness in human physiology. Biometrics offers a natural and reliable solution to the problem of identity determination by recognizing individuals based on their physiological and/or behavioral characteristics that are inherent to the person. Some popular biometric traits are fingerprint, hand-geometry, iris, retina, face, palmprint, ear, DNA, voice, gait, signature and keystroke dynamics.

Biometric Characteristics: Any human physiological and/or behavioral trait can be used as a biometric characteristic as long as it satisfies the following requirements [2]:

Universality: Every person should possess the biometric characteristic.

- > Distinctiveness: No two persons should have the same characteristics.
- Permanence: The biometric characteristic should not change over time that is, it should be sufficiently invariant.
- > Collectability: The biometric characteristic should be easily collectable.
- Performance: The biometric characteristic should give reasonable performance in terms of recognition accuracy and speed.
- > Acceptability: The characteristic should be readily acceptable by the people.
- Circumvention: The characteristic should not be easily spoofed by using fraudulent methods.

Table 1.1 summarizes the different levels of desired individual characteristics of commonly used biometric traits [4]

Biometric identifier	Universality	Distinctiveness	Permanence	Collectability	Performance	Acceptability	Circumvention
DNA	Н	Н	H	L	Н	L	L
Ear	Μ	M	Н	Μ	М	Н	Μ
Face	Н	L	M	Н	L	Н	Н
Facial thermogram	Н	H	L	Н	Μ	Н	L
Fingerprint	M	Н	Н	М	Н	Μ	M
Gait	Μ	L	L	Н	L	Н	Μ
Hand geometry	Μ	М	M	Н	М	Μ	Μ
Hand vein	M	M	M	Μ	М	Μ	L
Iris	Н	Н	Н	Μ	Н	L	L
Keystroke	L	L	L	Μ	L	М	Μ
Odor	Н	Н	Н	L	L	Μ	L
Palmprint	M	Н	Н	Μ	Н	М	Μ
Retina	Н	Н	M	L	Н	L	L
Signature	L	L	L	Н	L	Н	Н
Voice	Μ	L	L	М	L	Н	Н

 Table 1.1: Different characteristics of the commonly used biometric traits. [4]

Biometric System: Figure 1.1 shows a generic biometric recognition system [5]. It can

be divided into five sub systems:



Figure 1.1: High level schematic of a generic biometric system. [5]

- Data collection: This subsystem acquires the image of the biometric trait for the user. It consists of a sensor that captures the image.
- Transmission: This subsystem receives data from the data collection module, compresses it and then transmits to the signal processing unit and data storage unit.
- > Data Storage: In this subsystem images and templates are stored.
- Signal Processing: This is the main block of the system. It performs image preprocessing feature extraction and pattern matching operations.
- Decision: This subsystem uses the match score to verify the identity claimed by the user or to identify the user

Biometric System functionality: Based on the information flow and setup shown in figure 1.1 different deployment scenarios are established. During enrollment, the true identity of the user is associated with his biometric template. Biometric systems can provide two main functionalities, namely, (i) verification and (ii) identification. Figure 1.2 shows a pictorial representation of each scenario.

A generic biometric verification system utilizes the presented biometric trait to verify the claimed identity by the user. The system user claims his/her identity using some form of hard or soft identifier e.g. ID card, smart card or PIN. The claimed identity is then verified by matching the presented biometric against the stored biometric under that identity. Thus, the matching is 1:1 in a verification system. From a pattern recognition perspective such biometric systems can be considered as a binary classification problem, where an input candidate has to be classified as either 'genuine user' or 'imposter user' using a scalar distance/similarity score. In case of an identification system, the end user does not claim any particular identity but instead the system determines his identity among the enrolled set of users. The user input is compared against all available database entries and the user is assigned the identity of the most similar stored template. This type of setup is analogous to a multi-class classification problem being treated with 1-Nearest Neighbor classifier.

1.2 Unimodal Biometric Systems and Their Limitation

Biometric systems relying on a single source of information for the purpose of identification or authentication are termed as unimodal biometric systems. Unimodal systems have to contend with many limitations while operating in real world conditions

and are not as secure and infallible as they are desired to be. The following drawbacks, as



Figure 1.2: Diagrammatic representation of Identification and Verification system setup

described by Jain et al. [3], of a unimodal system lay the motivation behind the use of multimodal biometrics.

Noise in input data: Most systems suffer from unavoidable imperfections in the sensed data, a fingerprint with a scar and a voice altered by a cold are examples of noisy inputs. Noisy data can be a result of anomalies at the user's end, improper environmental conditions or poorly maintained sensing device. The inherent noise in the sensed sample trickles its way further into the system and may cause inaccurate matching with templates in the database causing errors in the final output decision.

Intra-class variations: Due to the inherent physiology of different traits one cannot expect to acquire exactly similar biometric data at different instances of time. The biometric data acquired for authentication may be different from the data obtained during enrollment causing errors in the matching process. These variations are primarily caused by inconsistent interaction of users with sensing device or due to change in environment or the sensor setting itself.

Inter-class variations: Biometric traits are represented by quantitative feature sets. While the biometric trait itself is observed to vary significantly across large populations, the feature-sets used to represent them have large similarities. Also, since the feature-set is a finite representation, theoretically an upper bound exists on their distinctiveness.

Non-universality: Every biometric system works under an assumption that the user will be able to successfully present their biometric to the sensor. But in real life situations there is a percentage of the population for whom is it not possible to acquire reliable biometric data. A fingerprint biometric system, for example, may be unable to extract features from the fingerprints of certain individuals, due to the poor quality of the ridges. **Spoofing**: It can be defined as intentionally cheating the system in order to make it accept an artificially prepared biometric as a true biometric. For example, video recordings of a user's face can be played on a laptop and presented before a facial recognition camera; iris images of very high quality (2400 x 1200) may be used to fool an iris camera; artificial fingerprints can be made using gelatin, playdoh, etc.

Due to these practical problems, the error rates associated with unimodal biometric systems are quite high which makes them unacceptable for deployment in security critical applications. The state-of-the-art error rates associated with fingerprint, face and voice biometric systems are shown in the Table 1.2 [3, 64]. The overall identification performance can be improved by using multimodal systems.

1.3 Multimodal Biometric System

Identification systems that consolidate information from multiple biometric sources are called multimodal biometric systems. The concept of multimodal systems has existed for over a decade now, but has attracted a lot of attention in the recent past as these systems meet the performance requirements of commercial deployment. The term multimodal biometrics refers to very broad class of systems and hence is specified by the type of multiple cues used for authentication and the methods used for combining these cues. Integrating independent information from multiple sources surpasses the upper bound on performance imposed on any given unimodal system. The probability of simultaneous successful spoof attacks on all the sources is much lower compared to a unimodal system. A number of multimodal systems have been presented in literature that differ from one another in terms of there architecture, the number and choice of biometric modalities, the

level at which the evidence is accumulated, and the methods used for the integration or fusion of information.

	Test	Test Parameter	False Reject Rate	False Accept Rate
Fingerprint	FVC 2004 [11]	Exaggerated skin distortion, rotation	2%	2%
	FpVTE 2003 [12]	U.S. government operational data	0.1%	1%
Face	FRVT 2004 [13]	Varied lighting, outdoor/indoor	10%	1%
Voice	NIST 2004 [14]	Test independent, multilingual	5-10%	2-5%

Table 1.2 Error rates associated with different unimodal biometric systems (Fingerprints, Face and Voice). Each system is evaluated under different test conditions [3, 64].

1.3.1 Sources of information

Based on the type of information being combined multimodal systems can be classified as the follows [3]

- Multiple sensors for the same trait: In this case information is fused using multiple digital samples belonging to the same trait but from multiple type of sensor technology. This setup can address the problem of noisy sensor data. (For example, fingerprint data from optical and solid-state fingerprint sensors.)
- Multiple instances of the same biometric: In this case information is combined from biometric samples acquired from the same trait but multiple setting and/or at different time instances. (For example, multiple face images of a person obtained

under different pose/lighting conditions).

- Multiple algorithms for the same biometric: Different recognition algorithms exploit different discriminatory properties in a given biometric sample. The combination of information from different algorithms can aid improving performance of an existing database or system. (For example, multiple iris matchers like PCA, ICA and Gabor phase information).
- Multiple units of the same biometric: Multiple physiological units of a single individual have been known to be distinctive. Also a recognition system that works on multiple units of the same biometric can ensure the presence of a live user by asking the user to provide a random subset of biometric measurements. (For example, left index finger followed by the right middle finger).
- Multiple biometric traits: In this case information is derived from different biometric traits. This type of fusion has received a lot of attention because a system based on multiple traits is more robust to noise, addresses the problem of non-universality, improves the matching accuracy and is more difficult to spoof. (For example, face, fingerprint and iris based system).

In the serial architecture, each information source is activated sequentially and the result at any stage affects the evoking of the next stage. An example of such a system was proposed by Hong and Jain [17], for fusion of face and fingerprint systems. In the parallel architecture, all information sources contribute independently and a fusion scheme is devised to combine the input information. This architecture achieves lower error rates and hence has been widely used, details about this can be found in [16] and [18]. The cascaded system can be more efficient and convenient to the user

but the design such a system is more complex and needs more attention from the research community.

1.3.2 Information Architecture

Biometric data goes through several transformations from acquisition till matching. In case of multiple information sources, information fusion and processing can either have a cascade or parallel architecture (see figure 1.4).



Figure: Architecture of multimodal biometric systems; (up) serial and (down) Parallel

The cascade architecture sequentially utilizes additional biometric sources if a reliable decision cannot be made with the available information. There is still much to be done in structural design of such an arrangement. The parallel architecture although more

exhaustive, provides more robust performance.

1.3.1. Levels of fusion:

Fusion in multimodal biometric systems can take place at four major levels [3] (see figure 1.5) raw data level, feature level, score level and decision level. As seen in figure 1.5, the information can be combined at different stages; information versatility reduces after every stage of processing.



Figure 1.5: Different levels of information fusion in multimodal systems

Decision Level Fusion: Individual biometric systems, as final output, give out a binary

decision or a rank based decision. Decision level fusion involves integrating the results emerging from individual decision modules of each constituent source. Several methods exist in the literature to achieve a final decision from multiple outcomes such as, majority voting [28], behavior knowledge space [29], and weighted voting based on Dempster-Shafer theory of evidence [30]. Decision level is the poorest level in terms of available information. Thus performance improvement due to fusion that can be achieved, under optimal processing, is less than that can be achieved at preceding levels of fusion.

Match-Score Level Fusion: When the biometric matchers output a set of possible matches along with the quality of each match (matching score), fusion can be achieved at the match-score or confidence level. This process involves integration of different match score, generated from each component modality, into a consolidated match score. Next to the feature vectors, the matching scores generated by the matching module contain the richest information about the input pattern. Match-score fusion is fairly simple because most available systems generate a scalar similarity or dissimilarity score. In the past, match-score fusion has received a lot of attention due to ease of fusion and availability of matching scores from various commercial biometric products. The fused match-score is either generated through (i) rule based approach (ii) or classification based approach.

Feature Level Fusion: Feature level fusion refers to consolidating information from different feature-sets or templates obtained from participating modalities or information sources. Usually, when the feature vectors are in the same space, a single fused template is created from multiple input feature-sets. This template is then used as the input to the matcher to generate match scores, which is then fed to the decision making module. Since the features contain richer information about the input biometric data than the matching

score or the decision of a matcher, integration at the feature level should provide better recognition results than other levels of integration. The intuition towards better performance by feature level fusion faces several practical challenges. Also, most commercial biometric products do not make feature sets available to researchers due to the proprietary nature of there processing algorithm. Hence very few researchers have studied integration at the feature level.

Sensor Level Fusion: The digital data from the sensors are directly combined in *sensor level fusion* [20]. Although sensor level is the richest information source among all other levels, the representation of this information is extremely diverse. Sensor level fusion is, in most cases, possible when the same biometric trait is captured at multiple instances or setting using highly compatible or same sensors. For example, the face images obtained from several cameras can be combined to form a 3D model of the face. Another example of sensor level fusion is the mosaicking and super-resolution of multiple fingerprint impressions to form a more complete and detailed fingerprint image [21, 22]. Data registration and normalization is critical for effective sensor level fusion, unfortunately this is a very challenging task. Sensor level fusion may not be possible if the data instances are incompatible (e.g., it may not be possible to integrate face images obtained from cameras with different resolutions).

1.4 Thesis Contribution

A review of the proposed multimodal systems shows extensive research in decision level and match-score level fusion. Information fusion at both these levels has shown limited improvement in performance. Match-score represents the composite distance/similarity between two arbitrary templates. Based on this information a decision is then made. Due to processing there is a significant loss of information which originally existed in the templates. In this thesis, we turn towards the feature level fusion and propose the following methodology:

- We propose using a novel transformation on the original feature vectors to deal with the problem of incompatible templates. Since feature vectors obtained for different information sources/ modalities are heterogeneous, normalization is required to transform the feature vector into a common domain before combining them. We define *local distance kernels* for this purpose. These kernels can be adjusted to be biometric specific for example, processing fingerprint minutia data or Iris binary codes.
- Noise in input biometric sample leads to anomalies or noisy feature in the biometric templates. In this thesis, we propose two new feature selection (or noise removal) techniques so that only the most meaningful features contribute to the consolidated matching score.
- In this thesis we use *AdaBoost* for multimodal fusion. AdaBoost is a learning theory based method for feature selection and classifier design, which utilizes boosting and large margin theory in order to design a strong classifier by linearly combining the best *n* individual features. Each distance feature coupled with a scalar threshold forms a weak classifier. AdaBoost then selects the classifiers and there corresponding weights.
- In order to calculate the most discriminating matching score we propose a completely learnable match score calculation technique based on *Support*

Vector Machines (SVM's). SVM's rely on the theory of *'large margin analysis'* to obtain an optimal separating hyperplane which directly maximizes the minimum distance between the hyperplane and the distance vectors. The matching score generated with this method symbolizes the confidence generated by the trained classifier.

The recognition performance of all proposed techniques is empirically evaluated for verification mode and identification mode. Results are shown to be superior to the match-score level fusion.

This thesis is organized in the following manner. In Chapter 2, we describe the local distance transformation and the proposed feature selection scheme. Chapter 3 explains the theory behind AdaBoost and our proposed method for fusion. Chapter 4 focuses on the application of support vector machines for generating fused matching scores. Conclusions and future work are provided in Chapter 5.

Chapter 2: Generic Information Fusion at Feature Level

2.1 Motivation and Challenges

The feature-set which quantitatively represent a given biometric modality is the richest source of information. With a matching score obtained from fused templates, one may expect performance of a MMB system to be improved compared to the system that uses unimodal matching scores. While match-score and decision level fusion have been extensively studied, feature level fusion is relatively unexplored. In previous feature level fusion (FLF) attempts, Kumar et al. [23] fused feature extracted from palmprints and hand-geometry by simple concatenation of feature vectors from each modality. Ross and Govindarajan [24] further studied in detail FLF for face and hand-geometry features. They applied feature normalization and introduced a new thresholded distance metric. In their work, improved performance was demonstrated by using a feature selection techniques and feedback system in conjunction with match score. But in all previous work, FLF has been limited to modalities or algorithms which generate compatible feature vector, for example Euclidian features. This limitation, incompatible feature set, has limited researchers from extending FLF to all biometric traits, such as fingerprints and hand geometry.

2.2 Generation of Multimodal Database

In this chapter we will introduce a generic technique to overcome the limitations imposed on feature level fusion due to different representative schemes used for different traits. In order to focus on fusion, where ever possible, *off-the-shelf* feature extraction algorithms

	Algorithm	Domain	Dimensionality
	РСА	Euclidian	25
Face	ICA	Euclidian	25
	Geometric	Real	27
	PCA	Euclidian	First 24 basis vectors out of 99
Iris	ICA	Euclidian	First 24 basis vectors out of 99
	2D Gabor Filter	Binary	34560
Fingerprint	Minutia Points	Location and orientation	Variable

were leveraged. Table 2.1 details the various algorithms and modalities used in our work.

Table 2.1: Detailed description of the different biometric feature sets used for our experiments

Face: Facial images corresponding to 100 users, 5 samples per user were obtained using a CCD camera at West Virginia University. Principle component Analysis (PCA) and Independent Component Analysis (ICA) methods were implemented and used to extract features from the images [6, 7]. Face Geometry features were manually extracted using a Graphical User Interface (GUI) designed by Arun Anthony at West Virginia University.

Iris: Iris images corresponding to 100 eyes, 5 samples per eye were obtained from the publicly available CASIA dataset [8]. The global PCA and ICA encoding methods described by Dorairaj and Schmid [9] were implemented and used to extract features from the normalized iris images. A 2D Gabor filter based algorithm, our implementation of Daughman's algorithm [10], was implemented to obtain Iris Codes. The details of this

system are described in the later sections.

Hand Geometry: It utilizes the geometric and physical characteristics of the hand for establishing identity (e.g. length of the fingers, width of the fingers, width of the palm, thickness of the hand etc.). Even though individual hand features have been found insufficient for identification, a collection of features can be reliably used for the verification of an identity. The hand geometry database consist of hand images corresponding to 100 users with 5 samples per user, obtained from a commercial off the shelf (COTS) hand geometry system installed at West Virginia University. The feature set consists of 9 features corresponding to different measurements of the hand biometric.

Fingerprints: A data base of minutia based fingerprint representation was obtained from a COTS system (VeriFinger from Neurotechnologija) for 100 fingers, 5 samples per finger. The commercial feature extraction module provided the location and orientation of the detected minutia points. Also, using the same COTS system the matching scores and translation parameters were obtained between all pairs of fingerprints. The matching methodology of minutia sets is explained, in detail, in a later section.

2.3 Fusion Framework

By using multimodal systems we desire to obtain highly accurate matching scores between the candidate template and database template. With these fused matching scores, we can threshold them for verification or rank them for identification. Unfortunately, due to wide diversity in information encoding techniques used for different biometric sources, a match-score obtained from naïve concatenation of multiple feature vectors, whenever possible, does not improve performance beyond match score fusion. For example consider fusion of two information sources with feature vectors in Euclidian space, previous methods have fused features by merely concatenating the component feature vectors and a consolidated matching score is calculated by simple Euclidian distance. This simple feature fusion procedure is diagrammatically shown in the figure 2.1. The performance of this trivial methodology was evaluated on our dataset (see section 2.2). From Receiver operator curves (ROC) in Figures 2.2, 2.3 and 2.4, it appears that this technique is sensitive to numerical range of the feature-set and susceptible to redundant and noisy feature values.



Trivial concatenation for feature level fusion does not result in consistent results. The following challenges are associated with feature level fusion [24]

- Each biometric entity is processed by a principally different algorithm hence the relationship between the feature spaces of different systems may not be known. For example fingerprints are represented by location and orientation of minutia points and iris templates by quantized Gabor phase response.
- There exists no theoretical method describing the non-linear relationship between

feature vectors from multiple sources.

- In most cases the fused template or the feature vector is of higher dimensionality than individual templates leading to the '*curse of dimensionality*' problem [53].
- Most commercial systems do not provide access to the feature vectors which they use in their product.



Figure 2.3: Inter modality ICA and PCA feature fusion between Iris and Face respectively. Performance improves even with simple concatenation.



Figure 2.4: fusion of cross modality and cross algorithm features-sets. The performance is at least as good as the strongest individual source.

The next section describes the proposed methodology to overcome these challenges and show feature level fusion to be a viable technique.

2.4 Data Preprocessing

To exploit the competency of feature level fusion it is critical to bring the information in individual feature set to a common platform before fusion. This section explains, in detail, the proposed techniques for data pre-processing to overcome some of the cited problems.

2.4.1 Feature Transformation

In general, we assume that there exists labeled training data available in the form of a multimodal database and unlabeled incoming data known as the candidate. As mentioned section 2.3, direct concatenation of feature sets in not possible when the features are in different space, e.g. fingerprint minutia information and face eigen co-efficient. To address this feature incompatibility issue, it is required to map the features in a common normalized distance space. We use candidate template and database templates to transform and synthesize a normalized distance vector such that it represents the multidimensional distance between the two templates. The numerical range of these distance vectors needs to be normalized such for fusion. We propose to map features into a common domain by applying *local distance kernel*. For this transformation, in this thesis, we assume that features are independent. The local distance kernels transforms each dimension independently. Feature vectors from multiple sources can be processed through biometric and representation specific kernels and then fused to reside in the common transformed space. Formally, denote a feature vectors, $\overline{X}_{[candidate]} \in \mathbb{R}^m$,

submitted for identification from given information source. Denote by, $\overline{Y}_{\{candidate\}} \in \mathbb{R}^{m}$, a feature vector from the same single source stored in a database. Then the local distance kernel, D, between X and Y is defined as a function φ that maps X and Y onto a non-negative portion of the real line, that is, $d_{i} = \{\delta \in \mathbb{R}^{1}_{+} : \delta = \phi(x_{i}, y_{i}) \text{ for } i = [1, ..., M] \}$

The following local kernels are examples used to obtain a local distances vector $\overline{D} = [d_1, d_2, ..., d_m] \in \mathbb{R}^m_+$, where the individual components d_i are defined as,

 $d_{i} = \phi (x_{i}, y_{i})$ for exam ple

$$\phi_{1}(x, y) = (x - y)^{2}$$

$$\phi_{2}(x, y) = \frac{|x - y|}{m a x (x, y) - m i n (x, y)}$$

$$\phi_{3}(x, y) = e^{-\frac{(x-y)^{2}}{2\sigma^{2}}}$$

2.4.2 Feature Normalization

Normalization is a common technique utilized to scale data into a common numerical range. Jain, Karthik and Ross [64] have used different normalization techniques for fusing match-scores and analyzed them in terms of there robustness and efficiency for rule based fusion. Direct normalization of feature values is not always possible because many feature sets exist in domains other than Euclidian, e.g. fingerprints and binary codes. To combine feature information, normalization is embedded in the transformation function itself which scales the contribution of each dimensionality to comparable numeric range such that the true relationship with the final matching is consistent. The

transformation $\phi_2(\bullet) \& \phi_3(\bullet)$ yields local distance vector of the same dimensionality but in a new common space and within common numerical range. The above transformation functions are directly applicable to features in Euclidian space, but for feature-sets in other domains slightly modified transformations are used which are discussed further in this chapter. Once the features are transformed to a common domain, they can be fused by mere concatenation.

2.4.3 Special Transformation

Biometric entities are represented by various processing algorithms which exploit most distinctive features of a given modality. One expects these features to be approximately similar for different samples of a given individual and not similar for samples obtained from other individual. The feature comparison is mathematically achieved by defining a distance measure between the two templates. This section explains in detail, the techniques used to transform minutia templates and binary iris codes to a normalized distance space. Once these features are transformed, fusion with any other transformed feature-set can be achieved by mere concatenation.

Fingerprint Minutia set transformation: A fingerprint refers to the flow of ridge patterns in the tip of the finger. The ridge flow exhibits anomalies in the local regions of the fingertip, and these are the position and orientation of these anomalies that are used to represent and match fingerprints. Each fingerprint is unique due to complex structure of it ridges and location of anomalous ridge endings known as minutia points [16]. The distribution of minutia points is used to match and establish the similarity between two fingerprints [65]. Minutia-based techniques attempt to align two sets of minutiae points

and determine the total number of matched minutia [66]. The performances of minutia based techniques rely on the accurate detection of minutiae points and the use of sophisticated matching techniques to compare two minutia sets. Generally, automatic fingerprint recognition is achieved with point pattern matching (minutiae matching) instead of appearance based matching or a ridge pattern matching of fingerprint images. To compensate for translation and rotation, simple affine transform is used for which scale and orientation parameters have to be estimated [66]. More advanced processing is applied to compensate for non-linear deformation due to the elastic physiology of fingerprints [67]. Once the alignment is achieved, a match score is generated based on the number of matching minutia.

Minutia based systems have several drawbacks [47].

- Due to the small contact area in new generation fingerprint sensors, sufficient minutia information is not captured. Due to this alternate information from fingerprints has been augmented to improve performance.
- Minutia based systems rely on accurate extraction of minutia location through advanced image processing. However, a fingerprint minutia extraction system cannot work accurately on images from different sensors.
- As the finger needs to be pressed against a flat sensor surface for imaging, nonlinear distortions are introduced with the fingerprint image obtained. Appropriate measures need to taken to compensate for this distortion prior to matching.
- Although relative minutia locations do not change over time, it is possible for minor cuts and bruises to obscure minutia information in the acquired image.

 A section of the user population consistently fails to provide good fingerprint images. Minutia based systems will cause inconvenience for such users. Hence, multi biometric capability should be available.

To apply local distance kernels to fingerprint minutiae templates, candidate minutia information is transformed by comparing the input and fingerprint with a database fingerprint. After applying all the required preprocessing for minutia matching, minutia pairs are defined by nearest neighboring minutia point to the reference minutia. A fixed length distance vector is synthesized from best 'n' minutiae pairs, each represented by their corresponding off-set. In figure 2.5, diagrammatic methodology used to generate fingerprint distance vector is shown. In our experiments simple affine transform was used.



The choice of 'n', number of minutia pairs to consider is empirically determined such that minimum probability of error is achieved on the training set.

The emphasis of this transformation scheme is not to propose a minutia based matching scheme but only to solve the template incompatibility challenge. The fundamental principle is to synthesize a distance vector between input information and the database. Current fingerprint systems use more complicated algorithms for minutiae matching, which compensate for nonlinear deformation in fingerprints [67]. We have assumed a simple model for fingerprint matching because the main focus of this work is on feature level fusion and not on individual algorithms.

Iris Recognition, Representation and Transformation: Of the many biometrics that exist, iris recognition is attracting more attention than any other biometric, mainly due to its permanence, accuracy, and increasing acceptance. The iris of the human eye is the annular part between the pupil and the sclera. It is a complex texture that contains a large number of irregular features, including freckles, coronas, stripes, furrows, crypts, etc. Furthermore, the spatial distribution of these features in the iris is also complex and can be viewed as realization of a complex spatial random biological process. Such randomly distributed and irregular features constitute the most distinguishing characteristics of the iris. This section explains in detail,

- a. The techniques used to encode iris features using 2D Gabor filters.
- b. Extracting local hamming distance vector from IrisCodes.

In the early 90's, Daugman extended his theoretical findings from modeling of primate visual cortical cell response to iris recognition [70]. He used 2D Gabor wavelets as carrier waves for iris feature extraction by phase demodulation. Daugman's work was motivated by the earlier finding of Oppenheim and Lim [71] who demonstrated the importance of phase for the perception of visual features. Daugman's iris recognition

system was patented in 1994. As a result, there are no details provided for accurate implementation of the system. In this thesis, IrisCodes are generated using our own implementation of Daugman's algorithm. Figure 2.6 shows a general block diagram that we followed in our implementation. The following steps are involved.



Step 1: Localization of the iris from a close up eye image: Localization is the process of segmenting the iris region from a background of pupil, eyelids, eyelashes etc. Daugman designed a circular edge detector using an integro-differential operator for performing this step [10].

Step 2: Normalization: Normalization is a transformation by which a scale and translation invariant representation of the localized iris is obtained. Daugman employs a doubly dimensionless polar representation by the use of a rubber sheet model. Nearest neighbor interpolation is used to extract information from Cartesian coordinate pixel system. In our implementation iris texture is normalized to a standard size of 60-by-360, implying 60 concentric radial bands at an angular resolution of 1 degree.

Step 3: Encoding: Daugman, in his earlier work, proposed 2D Gabor functions. The 2D

Gabor functions are local spatial bandpass filters that achieve the theoretical limit for conjoint resolution of information in the 2D spatial and 2D Fourier domains. This finding was applied to iris recognition by using 2D Gabor wavelets as the carrier waves and representing iris features using the phase of the response. Encoding is applied to the normalized iris region of the image. The phase information, obtained from demodulation by 2D Gabor wavelet demodulation, is coarsely quantized to four levels using binary codes. This output from the normalized iris region forms the binary template called "IrisCode". Although this general principle has been published, details of the required parameters settings are not made public. In this work, we used the fact that the iris has very fine texture near the pupil and the texture gradually becomes coarser as the radius increases. The iris textural content cannot be sufficiently described by any one scale and frequency choice of Gabor filter scale and frequency. Figure 2.7 shows the difference in texture in three different regions of the iris and the different scales of the Gabor filter applied in order to encode the different textures of the iris region.



Figure 2.7: Diagrammatic representation of the designed iris encoding method

Through empirical evaluation of the iris structure we came up with 3 different scales and

an orientation range which best suited our database. The phasor response from each pixel quantized to 2 bit representation. The outer annular bands were ignored during template generation in order to compensate for the lack of depth-of-focus of the CCD camera. This resulted in a 34560 bit template (48-by-720).

Step 4: Matching: The dissimilarity between the input iris image and the database template can be easily determined by finding the Hamming distance between their IrisCodes. During this stage the response obtained from noisy iris regions such as eyelids, eyelashes and specular reflection is ignored using a binary mask as per the following equation

$$HD = \frac{\|(code1 \oplus code2) \cap mask1 \cap mask2\|}{\|mask1 \cap mask2\|}$$

Where, code 1 and code2 are the IrisCodes of 2 irises. The binary mask (*mask 1 and mask 2*) is a binary vector of the same size as the IrisCodes. Based on the detected occlusions in each template, the corresponding mask is synthesized by allocating ones to iris pixels and zeros to detected occlusion.

Figure 2.8 shows the score distribution and ROC curve when our implementation ofDaugman's recognition algorithm is tested the iris images of the CASIA dataset (100subjects,6samplespersubjects,



Figure 2.8: Score distributions (left) and ROC curves of the designed algorithm on CASIA dataset (right)

Our implementation yielded an equal error rate of 3.15%. The binary templates are organized in a matrix format in order to draw correspondence with its parent pseudo polar iris texture representation. Both, the real and imaginary components of the response contribute to the phase. In our implantation, quantized bits from each component were merely concatenated.

Iris Template Transformation: To map each pair of IrisCodes into a "common feature space", we subdivide each IrisCode into rectangular subblocks of equal size as shown in Figure 2.9. We further treat blocks within each IrisCode as being independent. This is due to a random structure of the iris, where different subregions within the iris are weakly correlated.

The response of non-overlapping blocks, here of size 24 X 24, is considered as independent features of the normalized iris representation. To synthesize a normalized transformation vector, Hamming distance between each block is calculated.


Figure 2.9: schematic representation of the transformation technique applied to the binary iris templates

This gives a 60 dimensional transformed vector. Intuitively, this represents the dissimilarities within different regions of the iris structure.

With methodology founded for fusing different feature representation of various modalities, all further processing can be applied to a common fused vector.

2.5 Feature Selection

Consider a long vector of noisy features. Selecting uncorrelated and discriminative feature in any pattern recognition system often aids the over all recognition accuracy of the system. However, in reality, including new features improves performance initially, achieves the best performance point and then starts to degrade again. This phenomenon is known as "peaking phenomenon" or "curse of dimensionality". Feature selection is closely related to the more general problem of dimensionality reduction and efficient data representation.

Roughly speaking, supervised feature selection methods are applied in one of two

conceptual frameworks: the *filter model* and the *wrapper model* (Kohavi & John, 1997). In the wrapper model the selection method tries to directly optimize the performance of a specific predictor generalization performance for the selected feature set in each step. The main drawback of this method is its computational deficiency. In the filter model the selection is done as a *preprocessing*, without trying to optimize the performance of any specific predictor directly. This is usually achieved though an *evaluation function* using a search method in order to select a set that maximizes this function.

In this work, we assume that extracted features in feature vector are independent. But nonetheless one should expect the presence of noisy features is most datasets. Therefore in this context a dynamic dimensionality reduction scheme based on novel feature evaluation functions and a simple learnable threshold is proposed.

Let \overline{D} be the distance vector computed as per any of the local kernels defined above. We desire to estimate an *indicator function* for feature selection and seek our solution to be in a linear space. An evaluation function determines the quality of the feature and a soft or hard threshold removes a feature whose information level is lower than the threshold. An indicator function $\{\chi(A)\}$ is estimated or learned in order to dynamically select features. $\chi(A)$, for event A is defined as

$$\chi(A) = \begin{cases} 1, & \text{if } A \text{ is true} \\ 0, & \text{Otherwise} \end{cases}$$

From the concatenated distance vector (\overline{D}) , only features those features contribute whose information level is above a learned threshold. The following 2 methods are used to define conditions (A) to select features. Although the general principle behind these two methods is similar, the critical part is in how the information is measured.

<u>Method 1</u>: (Single Parameter Case)

Irrespective of the *type* of distance vector, *{genuine; imposter}*, each feature follows a model in the manner in which it contributes to the final score. Noisy features form anomalies to this model. Hence, their elimination will result in improved recognition performance. We eliminate them by thresholding them out. An indicator function is estimated or learned in order to predict the *usefulness* of a given feature. Let A be defined as

$$A = \left\{ \begin{array}{ccc} \frac{d_{i}}{\sum_{i=1}^{M} d_{i}} < \tau_{i} \\ \sum_{i=1}^{M} d_{i} & \end{array} \right\}$$

where M is that total number of kernel distances in the kernel distance vectors and τ_i is a threshold, a design parameter, such that, $0 \le \tau_i \le 1$, for i = 1, ..., M.

The thresholds are selected such that they minimize the total probability of error for the individual feature when the indicator function is applied. Thresholds are exhaustively selected for each feature in the kernel distance vector.

Method 2: (Two Parameters Case)

The information content of a Kernel distance feature can be measured from its margin with respect to a scalar boundary. In this thesis, the margin is defined as the distance of the datapoint from the optimal scalar boundary (as per minimum classification error). Define event A for this case as follows

$$\mathbf{A} = \left\{ \left| d_i - \kappa_i \right| < \tau_i \right\}$$

Where κ_i is the optimal scalar threshold for feature *i* and τ_i is the optimal feature

selection threshold for the i^{th} kernel distance.

Thus, a simple match-score can be calculated with feature selection incorporated within the match-score generation from the following formulation

$$MatchScore = \sum_{i=1}^{M} d_i \cdot \chi(A)$$

2.6 Training and Testing Procedure

Figure 2.10 shows the over all testing and training design deployed in our experiments. Distance vectors are generated by comparing all 500 samples with each other. The proposed fusion and feature selection scheme are empirically evaluated by using the 'Leave-one-out' design. As per this design, each feature sample is considered to be a candidate while the remainder of the data is treated as database entries. This process is repeated for all samples in the database. This procedure generates 1,000 genuine distance vectors 247,500 imposter vector. For estimating our system design parameters, we use all 1000 genuine distance vectors and randomly sample 10,000 imposter vectors. The same numbers of training samples are used in chapter 3 and 4 also. For testing, match scores are generated form all 1000 genuine fused distance vectors and all 247,500 imposter distance vectors. The results of this process are summarized in the form of a Match Score *Matrix*, where each element in the matrix represents the matching score generated by the system between elements indicated by the indices. Verification and identification analysis performed using the Match Score Matrix. are



2.7 Recognition Performance

During this cyclic process, matching scores are generated by the methodology explained in section 2.7.1. For both feature selection techniques the function parameters (required thresholds) are calculated independently for each feature using the training data (1000 genuine vectors and 10,000 imposter transformed distance vectors). Vectors are randomly sampled from the exhaustive set generated by leave-one-out combination of database samples.

2.7.1 Implementation Methodology

To demonstrate the effect of the transformation function and feature selection on real world features and compare performance with normalized match score fusion the following procedure are implemented

For Feature level fusion

Step 1. Feature-sets of 2 input biometric sources, $\overline{P} \in \mathbb{R}^{m_1}$ and $\overline{Q} \in \mathbb{R}^{m_2}$, are transformed

using local distance kernel

$$\overline{D}_{\text{Source 1}} = \phi(\overline{P}^{\text{Candidate}}, \overline{P}^{\text{Database}}) \text{ and } \overline{D}_{\text{Source 2}} = \phi(\overline{Q}^{\text{Candidate}}, \overline{Q}^{\text{Database}})$$

:

Step 2. Fusion is achieved by concatenating transformed distance vectors

$$\overline{D}_{Fused} = \left[\overline{D}_{Source 1}; \overline{D}_{Source 2}\right] \in R^{m1+m2}$$

Step 3. The final matching score, with feature selection, is calculated as $Scr_{Fused} = \sum_{i=1}^{m_{1}+m_{2}} d_{i} \cdot \chi(A) \text{ where } \chi(A) \text{ is learned from training data.}$

For Match-score Level fusion

Step 1. Matching Scores for each of the two biometric sources is calculated using the Euclidian distance measure

$$Scr_{\text{Source 1}} = \sum_{i=1}^{m1} \left(p_i^{\text{Candidate}} - p_i^{\text{Database}} \right) \text{ and } Scr_{\text{Source 2}} = \sum_{i=1}^{m2} \left(q_i^{\text{Candidate}} - q_i^{\text{Database}} \right)$$

Step 2. Individual matching scores are normalized using min-max normalization

$$s_i' = \frac{s_i - \min(s)}{\max(s) - \min(s)}$$

Step 3. The fused matching score is obtained by the sum rule

$$Scr_{Fused} = Scr'_{Source 1} + Scr'_{Source 2}$$

The verification and identification performance of the above scheme is calculated based on the match scores generated as per the above formulation. In this work only bi-modal systems are evaluated, i.e. two sources of information. For verification scenario we plot ROC curves and calculate the equal error rate (EER) value. Identification performance is based on the first best match or also known as the 1-Nearest Neighbor classification. **Intra-Modal fusion**: Fusion of this category involves combining independent feature sets, extracted by applying different encoding algorithms on the same image. This fusion scheme aims at improving performance of a biometric system which is based on a single physiological trait like face or iris.

Face: Three face recognition feature sets are extracted for all our experiments. ICA and PCA analysis or the better known as "eigen face" method are the most popular algorithms for face recognition and widely used in commercial systems. are Over a dataset of 500 samples, the individual EERs for PCA and ICA features are 6.7% and 10.4% respectively. Additionally, a third feature set is manually collected by extracting the inter-pixel distance between predefined landmark points on the same frontal face images. We have defined this feature-set as geometric face features. These geometric features are principally independent from ICA and PCA features, and exhibit a poor equal error rate of 29.63%. In our experiments we investigate the fusion performance obtained when all three combinations of feature-sets are combined.



Figure 2.11: Fusion of face ICA and PCA feature. No significant improvement in performance is observed.

	Equal Error Rate (%)	Identification Error Rate (%)
face PCA	6.9892	4.8
Face ICA	10.395	6.6
Fusion + Thresholding-1	10.046	8.8
Fusion + Thresholding-2	10.69	9.4
Match Score Fusion	8.0615	4.4

Table2.2: Performance of Identification and Verification systems



Figure2.12: Fusion of face Geometry and Face ICA features. The proposed fusion framework outperforms match-score fusion

	Equal Error Rate	Identification Error Rate
	(%)	(%)
Face Geometry	29.63	84
Face ICA	10.4	6.6
Fusion + Thresholding-1	6.5	11
Fusion + Thresholding-2	6.7	4.2
Match Score Fusion	9.3	9.2

Table2.3: Performance enumerated for Identification and Verification case



False Accept Rate(%) Figure 2.13: Fusion of face Geometry and Face PCA features. The proposed fusion framework outperforms match-score fusion

	Equal Error Rate	Identification Error Rate
	(%)	(%)
Face Geometry	29.63	84
Face PCA	6.9	4.8
Fusion + Thresholding-1	6.12	7.4
Fusion + Thresholding-2	7.11	4
Match Score Fusion	9.11	14.4

Table 2.4: Performance enumerated for Identification and Verification case

From the ROC curves shown in Figure 2.11, 2.12 and 2.14, we see that feature level fusion after normalization and feature selection performs better than match score level fusion scheme. In case of combining PCA and Geometric features, feature level fusion outperformed match level fusion, but could only marginally improve performance than that of the PCA features (best individual biometric). Whereas, adding face geometry information to ICA features at the feature level yielded better results for identification and verification mode of operation. There was no clear evidence of performance improvement when PCA and ICA feature are fused (see Figure 2.11). Table 2.2, 2.3 and 2.4 show the EER for verification systems and classification error for identification system.

Intra-Modal fusion For Iris: Three feature-sets are extracted for iris recognition. As explained before, binary iris codes are generated using a domestic 'Daugman like' algorithm. These binary codes represent the local iris texture captured by the phase response of two dimensional Gabor filters. Also, 2 feature sets are extracted by applying the ICA and PCA principles. It is worthwhile to keep in mind that all feature sets are obtained from the same original source, i.e. the image, but by applying different encoding techniques. We analyze the performance achieved by fusing these IrisCodes with the feature vectors obtained from ICA and PCA algorithms. The raw features cannot be directly concatenated as they exist in incompatible domains, i.e. binary and Euclidian. We use the proposed special local distance kernels to transform them into common arbitrary distance space. The fusion is achieved by concatenating the transformed distance vectors.



Figure 2.14: Fusion of IrisCodes and PCA feature. Feature fusion performs slightly lower than IrisCodes

	Equal Error Rate	Identification Error Rate
	(%)	(%)
Iris Codes	5.94	1.5
Iris PCA	17.3	22
Fusion + Thresholding-1	7.17	8.6
Fusion + Thresholding-2	16.9	31.0
Match Score Fusion	9.4	5.4

Table 2.5: Performance enumerated for Verification and Identification case



Figure 2.15: Fusion of IrisCodes and ICA feature. Feature fusion performs slightly lower than IrisCodes

	Equal Error Rate (%)	Identification Error Rate (%)
Iris Codes	5.94	1.5
Iris ICA	17.73	22
Fusion + Thresholding-1	6.79	8.0
Fusion + Thresholding-2	15.63	31.2
Match Score Fusion	9.96	5.4

Table 2.6: Performance enumerated for Verification and Identification case



performance is observed.

	Equal Error Rate (%)	Identification Error Rate (%)
iris PCA	17.349	22
iris ICA	17.731	22.8
Fusion + Thresholding-1	19.395	44.6
Fusion + Thresholding-2	19.039	54.8
Match Score Fusion	17.584	22.6

Table 2.7: Performance enumerated for Verification and Identification case

As shown in figure 2.14, 2.15 and 2.16, the ROC curves for intra-modal fusion of Iris codes with ICA and PCA information shows that, performance of the match-score fusion degrades. Applying the proposed fusion and thresholding techniques, at the feature level, does not improve performance. The resulting ROC curve is as good as the ROC curve for the strongest biometric (IrisCodes).

A similar trend as that of face ICA and PCA features is observed when similar type of fusion is performed for Iris. There was no evident improvement in performance when these two feature sets were fused. Learned feature selection by thresholding led to lack of generalization and resulted in mildly degraded performance. Match-score fusion also did not make any significant improvement in performance over the best individual source.

Table 2.5, 2.6 and 2.7, show the equal error rate, for intra modal fusion of Iris, in a verification system and classification error rate in a identification system.

Inter-Modal Fusion: This type of fusion refers to combining information from different modalities. In this case, the information sources, i.e., raw image, are drastically different and are considered independent. Although the processing algorithm can be based on the same principle, e.g. ICA or PCA features, but as the originating image is different the feature-sets contain more independence.

Face and Iris: The combination from face and iris is very lucrative. As the iris is an inherent part of the face, with a high resolution imaging setup and appropriate eye finding software, we can acquire both face and iris images at the same time. We analyze the performance when different Face feature-sets are fused with iris codes.

From the obtained results (Figure 2.17, 2.18 and 2.19), it is evident that fusion with thresholding improves performance over individual feature-sets and also over match-



score fusion. Match-level fusion of Iris codes with ICA and Geometric feature-sets from

Figure 2.17: Fusion of IrisCodes and face ICA feature. Feature Level fusion outperforms match-score fusion.

	Equal Error Rate	Identification Error Rate
Iris Codes	5.94%	1.4%
Face ICA	10.39%	6.6%
Fusion + Thresholding-1	3.94%	1.6%
Fusion + Thresholding-2	11.21%	12.6%
Match Score Fusion	5.25%	1.6%

Table 2.7: Performance enumerated for Verification and Identification case



Figure2.18: Fusion of IrisCodes and face PCA feature. Match-score fusion outperforms feature level fusion.

	Equal Error Rate	Identification Error Rate
Iris Codes	5.94%	1.4%
Face PCA	6.9%	4.8%
Fusion + Thresholding-1	3.6%	1.0%
Fusion + Thresholding-2	10.31%	12.8%
Match Score Fusion	2.3%	0.4%

Table 2.8: Performance enumerated for Verification and Identification case



False Accept Rate(%) Figure 2.19: Fusion of IrisCodes and face Geometric feature. Feature level fusion outperforms match-score fusion.

	Equal Error Rate	Identification Error Rate
Iris Codes	5.94%	1.4%
Face Geometry	29.63%	84%
Fusion + Thresholding-1	4.99%	5.6%
Fusion + Thresholding-2	10.31%	21.8%
Match Score Fusion	15.57%	8.4%

Table 2.9: Performance enumerated for Verification and Identification case



False Accept Rate(%) Figure 2.20: Fusion of IrisCodes and fingerprint minutia features. Feature level fusion performs poorly.

	Equal Error Rate	Identification Error Rate
Iris Codes	5.94%	1.4%
Fingerprints	5.37%	2.6%
Fusion + Thresholding-1	13.37%	32.0%
Fusion + Thresholding-2	20.0%	40.8%
Match Score Fusion	1.02%	0.2%

Table 2.10: Performance enumerated for Identification and Verification case

face, results in performance worse than the best individual modality. In the same case, feature fusion with thresholding gives the best results. Table 2.8, 2.9 and 2.10, enumerate the equal error rate for verification system and classification error rate for identification when face feature are fused with IrisCodes.

Fingerprint Minutiae and Iris codes: Fusion of Iris and Fingerprint is much desired as they exhibit exceptional accuracy. Fingerprint recognition technology has existed for over three decades, resulting in numerous matching techniques. The matching scores for out fingerprint data were obtained using a commercial system and hence show excellent performance. Not knowing the involved matching algorithm, we worked directly with the minutiae information given by the same system. For each minutia-set pair, the alignment was achieved using a simple affine transform. The Euclidian distance of the best 15 nearest neighbor pairs constituted the fingerprint distance vector (described in section 2.4.3). All procedures for Iris data-set were unchanged.

The above results are contrary to expectations with fusing two good biometric (Iris and Fingerprints). Match-level fusion shows exceptional performance as the generated iris codes exhibit good matching scores and the fingerprint matching score were obtained from a commercial system (who's processing techniques are not known). Feature level fusion heavily degrades performance because of the simplistic representation of fingerprint distance.

2.8 Summary

Feature level fusion has greater potential than match-score or decision level fusion and with good fusion schemes it can significantly improve performance. We have examined different multiple modalities and proposed representation, transformation and matching schemes for different feature representations. We have demonstrated that with our fusion scheme, most current feature representations can be combined to give performance superior to the performance of individual information sources. We have also proposed dynamic feature selection or noise removal algorithms which demonstrate superior performance when compared with match-score fusion in most cases. We have shown that match-score fusion of a weak biometric with a strong biometric, sum rule fusion, leads to degradation of overall fused performance. The use of thresholding scheme helps to remove noisy features and hence is most successful fusion procedure.

Specific to Iris recognition, we have developed a complete system from open source literature and demonstrated performance on CASIA datasets. Intra-modal fusion of IrisCodes with ICA and PCA features, lead to only limited success. This is mainly due to the fact that the participating feature set, although in a common distance domain, have drastically different relationships. Due to this a simple local distance summation does not give better matching scores.

Feature Level fusion of geometric face features with PCA and ICA face features improves performance, even though geometric feature exhibit extremely poor performance independently. Fusion of face features (PCA, ICA and Geometric) with IrisCodes, at the feature level, resulted in performance better than match-score fusion if not equal.

In spite of promising results, the simple match-score calculation from fused features can be improved. As observed in some cases, direct match-score calculation from the transformed vectors does more harm than good. In the next two chapters we describe, two approaches for performance improvement by introducing a *learnable null hypothesis*. The new match-score is then calculated by measuring the confidence of a transformed distance vector with respect to the null hypothesis.

Chapter 3: Boosting Biometric Features

3.1 Motivation and Challenges

Most biometric systems are designed as a derivative of more general pattern recognition and machine learning concepts. Machine learning deals with adaptive techniques capable of making accurate prediction based on a finite amount of partially characterized data. In our context, biometric systems are trained on labeled training data, with the goal to find a strategy which can generate accurate matching scores on the new test samples are trained on labeled training with the to make accurate predictions based on past observations. In our context, biometric systems have available labeled training data from which we need to design a strategy which can generate accurate matching scores on new test samples.

Building a highly accurate prediction rule is a difficult task, especially when most biometric features exist in high dimensional space and are not known to follow any particular family of distribution. With respect to feature level fusion this problem is further complicated by different feature representation schemes. In order to overcome the challenge of combining information across different feature representation scheme, we follow the local distance transformation detailed in Chapter 2. After the original features are transformed in to a filtered form by local distance kernel representation, the matchscore or confidence measure needs to be calculated by combining useful information and ignoring noise. This chapter details our "boosting" approach to feature selection and confidence measure generation.

3.2 Learning Boosted Function for Match-Score Generation

Boosting, a machine-learning method, is utilized to learn "match-score generation rules".

To apply the boosting approach, we start with the transformed and fused multidimensional local distance vectors. Consider two feature-sets, having m1 and m2 features respectively. After transformation and concatenation the fused-set has M = [m1 + m2] features. Our hypothesis, supported by experimental observations, is that a sub-set of these features can be linearly combined to give an effective matching score. The challenge is to find the sub-set of features and their respective combination rules.

In general, the boosting algorithm repeatedly calls different feature subsets from the original data, each time feeding it a different distribution or weighting over the training examples. Each time it is called, the base learning algorithm generates a new weak prediction rule, and after many rounds, the boosting algorithm combines the weak rules into a single prediction. To make this approach work, there are two fundamental questions that must be answered: first, how should each distribution be chosen on each round, and second, how should the weak rules be combined into a single rule?

In our system a variant of the AdaBoost (used by Viola and Jones [63]) is used both to select the features and to train the classifier [72]. In its original form, the AdaBoost learning algorithm is used to boost the classification performance of a simple learning algorithm (e.g. it might be used to boost the performance of a simple perceptron). It does this by combining a collection of weak classification function to form a stronger classifier. So, for example, the perceptron learning algorithm searches over the set of possible perceptron and returns the perceptron with the lowest classification error. The learner is called weak learner because no one expects even the best classification function to classify the training data well. In order for the weak learner to be boosted, it is called upon to solve a sequence of learning problems. After the first round of learning, the

examples are re-weighted in order to emphasize those which were incorrectly classified by the previous weak classifier. The final strong classifier takes the form of a perceptron, a weighted combination of weak classifiers followed by a threshold.

Several researchers have applied boosting in different machine learning applications and have reported excellent results [76]. Freund and Schapire [76] proved, using theory of margins, that the proposed boosting algorithm does not cause over-fitting when more simple classifiers are added to the final strong classification rule. They derived the training error to exponentially approach zero after each round. More importantly a number of results were later proved about generalization performance [73]. The generalization performance has a strong relationship with the margin of the samples and the AdaBoost algorithm is known to rapidly achieve large margins with each round of boosting.

With respect to our proposed multi-modal feature fusion framework, the challenge is to associate a large weight with each good classification function and a smaller weight with poor functions. One practical method for completing this analogy is to restrict the weak learner to the set of classification functions each of which depends on a single feature. In support of this goal, the weak learning algorithm is designed to select the single local distance feature which best separated the genuine and imposter examples. For each distance dimension, the weak learner determines the optimal threshold classification function, such that minimum examples are misclassified. A weak classifier $(h_j(d))$ thus consists of a distance dimension (d_j) , a threshold (θ_j) and a parity (p_j) indicating the direction of the inequality sign: $h_j(d) = \begin{cases} 1 & \text{if } p_j d_j < p_j \theta_j \\ 0 & \text{otherwise} \end{cases}$

57

- Given training data $(\overline{D}_1, y_1), ..., (\overline{D}_n, y_n)$ where $\overline{D} \in \mathbb{R}^m$ and $y_i = 0, 1$ for genuine and imposter distance vectors respectively.
- Initialize weights $w_{1,i} = \frac{1}{2p}, \frac{1}{2q}$ for $y_i = 0, 1$ respectively, where *p* and *q*

are the number of imposter and genuine training vectors respectively.

1. Normalize the weights,
$$w_{t,i} = \frac{W_{t,i}}{\sum_{l=1}^{n} w_{t,l}}$$
, so that w_t is a

probability distribution

2. For each feature, j, train the classifier which is restricted to using a single feature. The error is evaluated with respect to w_t

$$\varepsilon_{j} = \sum_{i=1}^{n} w_{i} \left| h_{j} \left(x_{i} \right) - y_{i} \right|.$$

- 3. Choose the classifier, h_t , with the lowest error ε_t .
- 4. Update the weights:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e}$$

Where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and

$$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}.$$

• The Final Matching Score is : $MatchScore = \sum_{t=1}^{T} \alpha_t h_t(d)$, where

 $\alpha_t = \log \frac{1}{\beta t}$

3.3 Training and Testing Architecture

For improving recognition performance via feature fusion, we apply the above described AdaBoost learning algorithm to select feature and calculate weights to be associated with each feature. Randomly sampled training data is used to formulate match-score generation rules. Each feature of the input vector independently contributes, according to the transformation function and the database entries, to a local distance feature. During training, the weak learner calculates the optimal scalar threshold for each distance feature. Each distance feature, d, when associated with a threshold (θ) and the respective polarity (p) constitutes a weak classifier. During each round, t, of boosting the best classifier (feature, threshold and polarity) is selected based on minimum classification error for the current distribution (W_t) . A weight is associated with the selected classifier, which is inversely proportional the empirical error rate. α_t to



In our experiments, 1000 genuine distance vector and 10,000 imposter distance vectors were used for learning. In the Testing stage, all-to-all comparisons are computed based on the learned rules. We have restricted our-selves to bi-modal fusion, i.e. information from two sources only. Within this paradigm, results have been demonstrated for Intra-modal and Inter-modal fusion. Datasets from the Iris, Face and Fingerprint domain were combined and there verification and identification performance is reported. Identification is based on the 1-Nearest Neighbor classification procedure.

3.3.1 Intramodal Fusion

Intra-modal fusion is a very attractive means of improving the performance of a biometric system based on a single trait. This process involves combining independent features-sets, extracted by using different encoding algorithms. In our experiments we perform bi-modal fusion among iris and face feature-sets. **Face:** Three different processing algorithms were used for feature extraction, i.e. PCA, ICA and Geometric features. PCA and ICA features are correlated due to similar fundamental principles of feature extraction. The geometric features represent, manually extracted, inter pixel distances between predefined landmark points on a frontal face image. Geometric features are not strongly correlated with ICA/PCA features, but exhibit poor performance individually. Figure 3.2 and Table 3.2 shows the ROC curve and enumerates verification and Identification performance respectively, for feature level fusion of face ICA and PCA feature. There was no performance improvement observed. AdaBoost based feature selection and classification led to mixed results when compared to match-score fusion. Although an improvement over the direct fusion strategy described in Chapter 2, when



Figure 3.2: Fusion of face PCA and ICA features. Match Level Fusion and Feature Level Fusion fail to improve performance.

	Equal Error Rate (%)	Identification Error Rate (%)
face PCA	6.9892	4.8
face ICA	10.395	6.6
Boost Fusion	10.883	15.2
Match Score Fusion	8.0615	4.4

Table 3.2: Performance enumerated for Verification and Identification case



Figure 3.3: Fusion of face Geometric and ICA features. Feature level fusion outperforms Match score fusion

	Equal Error Rate (%)	Identification Error Rate (%)
face GEO	29.638	84
face ICA	10.395	6.6
Boost Fusion	7.3573	10.4
Match Score Fusion	9.3322	9.2

Table 3.3: Performance enumerated for Verification and Identification case



Figure 3.4: Fusion of face Geometric and ICA features. Feature level fusion outperforms Match score fusion

	Equal Error Rate (%)	Identification Error Rate (%)
face GEO	29.638	84
face PCA	6.9892	4.8
Boost Fusion	6.156	8.4
Match Score Fusion	9.8188	14.4

Table 3.4: Performance enumerated for Verification and Identification case

compared with SVM based fusion proposed in Chapter 4, boosting does not yield similar outstanding results. SVM based feature fusion has better recognition performance, with every combination of feature-sets. Figure 3.3 shows the ROC curve and Table 3.3 shows the equal error rate for verification system and classification error rate for identification system when Geometric and ICA features are fused. Feature level fusion, using boosting, out performs match level fusion. Figure 3.4 shows the ROC curve and Table 3.4 shows the equal error rate for verification system and classification error rate for identification system when Geometric and PCA features are fused. Feature level fusion, using boosting, out performs match level fusion.

As seen from Figure 3.3 and 3.4, boosting performs well only when weak or poor feature sets are combined with strong or good feature sets. The fusion of ICA and PCA face features results in degraded fused performance. This degradation is possibly due to the lack of generalization of the selected features and inherent correlation among features. As AdaBoost is designed to "boost" the performance by combining weak classifiers, the fusion with geometric face features resulted in good overall performance.

Iris: Three processing algorithms were used to extract features from pseudo-polar iris images, i.e. ICA, PCA and Gabor Phase. Different transformation techniques were used to bring Iris codes and ICA/PCA features to common arbitrary space, after which fusion was achieved by simple concatenation. In our experiments, AdaBoost learning selected a small subset of features and there corresponding weights. But we observed, it lacked generalization capabilities when subjected to test data. The intra-modal iris fusion performance encountered only limited success. Figure 3.5, shows the ROC curve obtained when ICA and PCA feature of iris are fused. Table 3.5 shows the equal error

rate of a verification system and classification error rate of an identification system. The feature level fusion performance in this case was poorer than the both ICA and PCA features. As ICA and PCA encoding schemes are inherently based on dimensionality reduction principles, further feature selection by AdaBoost results in poor generalization capability.

Figure 3.6 shows the ROC curve and Table 3.6 shows the equal error rate for verification system and classification error rate for identification system when IrisCodes and ICA features are fused. Feature level fusion, using boosting, performs only performs as well as IrisCodes.

Figure 3.7 shows the ROC curve and Table 3.7 shows the equal error rate for verification system and classification error rate for identification system when IrisCodes and PCA features are fused. Feature level fusion, using boosting, only performs as well as IrisCodes. Individually, Iris codes perform better than ICA/PCA features. When fused with ICA and PCA features, the subset of selected features consisted mostly of transformed features from IrisCodes due to the high performance of transformed IrisCodes. Match-score fusion of Iris codes and ICA/PCA features showed similar degradation as in the case of face. Typical to this work, when ICA and PCA (iris or face) features are fused there is no significant performance improvement at feature level and match-score level.



Figure 3.5: Fusion of IrisCodes and ICA features. Feature level fusion performs only as well as IrisCodes.

	Equal Error Rate (%)	Identification Error Rate (%)
Iris gabor	5.9407	1.4
iris ICA	17.731	22.8
Boost Fusion	7.759	5.8
Match Score Fusion	9.9673	5.4



Figure 3.6: Fusion of IrisCodes and PCA features. Feature level fusion performs only as well as IrisCodes.

	Equal Error Rate (%)	Identification Error Rate (%)
iris gabor	5.9407	1.4
iris PCA	17.349	22
Boost Fusion	7.7876	5.6
Match Score Fusion	9.4528	5.4

Table 3.6: Performance enumerated for Verification and Identification case



Figure 3.7: Fusion of iris ICA and PCA features. Match score fusion and Feature level fusion do not improve performance

	Equal Error Rate (%)	Identification Error Rate (%)
iris ICA	17.731	22.8
iris PCA	17.349	22
Boost Fusion	20.061	40.4
Match Score Fusion	17.584	22.6

Table 3.7: Performance enumerated for Verification and Identification case
After the transformation into arbitrary distance domain, the features are concatenated. The transformed and fused features are randomly sampled for training. AdaBoost treats the fused features as a pool of classifiers, from which a subset of features is selected and their weighting coefficients are formulated. IrisCodes are combined with different feature sets from face and with minutia information from finger. Performance results of other fusion combinations are shown in the Appendix.

Iris Codes and Face Features: All face feature sets (PCA, ICA and Geometric feature) exist in Euclidian domain and the IrisCodes are in binary domain. Figure 3.8, shows the ROC curves for the fusion IrisCodes and face PCA features. Table 3.9 shows the equal error rate of a verification system and classification error rate of an identification system. There is no significant improvement observed. In Figure 3.9, similar results are observed when face ICA features are fused with IrisCodes. Table 3.9, shows the verification and identification error rates. Applying AdaBoost for fusion of PCA and ICA face feature with IrisCodes, no significant increase in performance was achieved. Once again, like in intra-modal fusion, boosting helps improve performance when weak feature set, like geometric face, is fused with a strong feature set. Figure 3.10, further confirms this hypothesis as fusion of poor geometric features with the strong IrisCodes results in improved performance for feature level fusion.

Iris Codes and Fingerprints: Minutia information and iris binary bits are transformed by applying their corresponding distance kernels. From Chapter 2 we know that trivial match score calculation does not perform well. But, boosting gives comparable results to match-score fusion. As a note, the fingerprint matching scores were generated from a COTS system. Figure 3.11, shows the ROC curves for the fusion of fingerprints and



	Equal Error Rate (%)	Identification Error Rate (%)
iris gabor	5.9407	1.4
face PCA	6.9892	4.8
Boost Fusion	5.4801	4.4
Match Score Fusion	2.3859	0.4

Table 3.8: Performance enumerated for Verification and Identification case



Figure 3.9: Fusion of IrisCodes and face ICA features. No significant improvement in performance is observed

	Equal Error Rate (%)	Identification Error Rate (%)
iris gabor	5.9407	1.4
face ICA	10.395	6.6
Boost Fusion	5.6965	5.8
Match Score Fusion	5.253	1.6

Table 3.9: Performance enumerated for Verification and Identification case



Figure 3.10: Fusion of IrisCodes and face Geometric features. Feature level fusion outperforms match-score fusion

	Equal Error Rate (%)	Identification Error Rate (%)
iris gabor	5.9407	1.4
face GEO	29.638	84
Boost Fusion	5.221	4.8
Match Score Fusion	15.573	8.4

Table 3.10: Performance enumerated for Verification and Identification case



Figure 3.11: Fusion of IrisCodes and fingerprint minutia features. Match score fusion outperforms feature level fusion.

	Equal Error Rate (%)	Identification Error Rate (%)
finger minutia	5.5327	2.6
iris gabor	5.9407	1.4
Boost Fusion	5.9866	7.6
Match Score Fusion	1.021	0.2

Table 3.11: Performance enumerated for Verification and Identification case

Chapter 4: Support Vector Machines for Fusion

4.1 Motivation and Challenges

The major challenge with feature level fusion is the lack of knowledge of the underlying multidimensional distributions of the feature-sets from different sources. The transformed vectors exist in a multidimensional space. The best match score can be calculated by projecting these vectors on an optimal separating hyperplane. In this chapter we justify this phenomenon by using both "toy data" and real world biometric data. The optimal hyperplane is learned by using support vector (SV) theory. Finally, a complete training and testing system architecture is proposed and results are shown on multiple datasets.

4.2 Learning with Support Vector Machines

In the previous chapters, techniques to transform feature vectors to a common distance domain have been explained. After the transformation, feature information is combined to obtain a fused matching score. Until now a match score is calculated from direct summation of these distance vector after eliminating noisy contributions. Match score thus calculated are inefficient as the summation is equivalent to simplistically projecting the distance vectors to a constant diagonal hyperplane in order to generate a matching score.

Match Score = $\overline{W}'\overline{D}$, where W = [1, 1, ..., 1]

Consider a two class problem, Class 1 and Class 2.



Figure 3.1: A two class, linearly separable classification problem. Infinite decision boundaries can generate. All decision boundaries are equally good.

As seen from the figure 4.1, when the data is linearly separable, there can be infinite optimal decision boundaries with zero classification error. But not all boundaries will have good generalization capabilities.

We believe that since the distance vectors reside in a multi-dimensional space, one can extract a better matching score by measuring the confidence of each vector with respect to a separating hyperplane. The theory of learning a separating hyperplane via Support Vector learning is applied to generate confidence scores.

The principle of SVMs relies on a linear separation in a high dimensional feature space where the data have been previously mapped, in order to take into account the eventual non-linearties of the problem. In order to achieve a good level of generalization capability, the margin between the separator hyperplane and the data is maximized.

The margin can be loosely defined as the width that the boundary could be increased by before hitting a datapoint. The decision boundary should be as far away from the data of both classes as possible. The choice of best decision boundary is mostly determined by datapoints that lie closer to the boundary itself. These datapoints are called support vectors.



Figure 4.2: Geometric representation of the large margin principle.

Formally, the training set $S = \{x_i\}_{i=1}^l \in \mathbb{R}^m$, where *l* is the number of training vectors, is labeled with two-class targets $\{y_i\}_{i=1}^l$, where $y_i \in \{-1,1\}$. $\Phi: \mathbb{R}^m \to F$ maps the data into a feature space *F*. Vapnik [62] has proved that maximizing the minimum distance in space *F* between $\Phi(S)$ and the separating hyperplane $H(w,b) = \{f \in F | \langle w, f \rangle_F + b = 0\}$, (where $\langle \cdot, \cdot \rangle_F$ denotes inner product in space *F*), is a good means of reducing a bound on the generalization risk. Vapnik also proved that the optimal hyperplane can be obtained solving the convex quadratic programming (QP) problem:

Minimize
$$\frac{1}{2} \|w\|^{2} + C \sum_{i=1}^{l} \xi_{i}$$
with the constraints
$$y_{i} \left(\left\langle w, \Phi\left(x_{i}\right) \right\rangle_{F} + b \right) \geq 1 - \xi_{i} \quad i = 1, ..., l$$

$$\xi_{i} \geq 0 \qquad \qquad i = 1, ..., l \qquad (1)$$

Where constant *C* and the slack variable ξ_i are introduced because of the eventual nonseparability of $\Phi(S)$ in space *F*. Applying the Karush-Kuhn-Tucker conditions to the problem in (1), the following sparse expression is obtained for the optimal hyperplane $H(w^*, b^*)$:

$$w^* = \sum_{i \in SV} \alpha_i y_i \Phi(x_i)$$
 (2)

Where, $SV = \{i \mid \alpha_i > 0\}$ is the set of support vectors. Taking into account that the decision function S that classifies a test pattern x_T is:

$$D(x_T) = sign\left\{ < w^*, \Phi(x_T) >_F + b^* \right\}$$
(3)

defining $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle_F$ as the kernel function and using (2) leads to

$$D(x_T) = sign\left\{\sum_{i \in SV} \alpha_i y_i K(x_i, x_T) + b^*\right\}$$
(4)

Problem (1) is solved for $(a_i)_{i=1}^l$ and b^* in its dual form with a standard quadratic programming solver which, together with decision function (4), avoids manipulating directly the elements of *F* and starting the design of the SVM for classification directly from the kernel function. The choice for K has been in this case a Radial Basis Function (RBF):

$$K\left(x_{i}, x_{j}\right) = \exp\left(-\left\|x_{i} - x_{j}\right\|^{2} / 2\sigma^{2}\right)$$
(5)

where σ is the variance of the kernel

To obtain the final classifier score, the proximity of the test pattern to the separating surface is proposed here. The combined score $Scr \in R$ of the input test pattern $p \in R^m$ with respect to the database entry $q \in R^m$ is calculated as:

$$Scr(p,q) = \sum_{i \in SV} \alpha_i y_i K(d_i, \phi(p,q)) + b^*$$
(6)

Following this approach, the verification threshold parameter can be adjusted to reach different operational points. The before described feature selection scheme can be easily incorporated into the SVM based score calculation

$$Scr(x,\mu) = \sum_{i \in SV} \alpha_i y_i K(d_i, \phi_{feature \ selection}(x,\mu)) + b^*$$
(7)

4.3 Architecture of SVM Based Information Fusion

Information is represented in the form of feature-sets for each biometric. These feature sets are independently extracted and transformed to common arbitrary distance space, so that information can be meaningfully fused. Chapter 2 explains various techniques used for such transformation. From this point the task is to calculate an efficient matching score utilizing fused information. As explained above the theory of SVM's is justified to learn a hyperplane on which the distance vectors are projected to get matching scores.

This system design involves two stages: training and testing. During the training stage we use a random subset of transformed and fused vectors to learn required parameters. In the testing stage, each sample is classified assuming remaining samples to be the database.

The high level design of the system for training and testing is shown in Figure 4.3. All available biometric databases have 100 unique classes, 5 samples per class. Corresponding features from each modality are calculated by different algorithms as explained before. All-to-All combination of the database templates generates 1000 genuine comparisons and 247,500 imposter comparisons. To evaluate the performance of a SVM based fusion scheme, a SVM based hyperplane is learned using 1000 genuine and

10,000 imposter distance vectors. OSU-SVM software package was used in this work to estimate the hyperplane and evaluate fusion performance [74].



4.3 Performance of SVM Fusion Scheme

Bi-modal fusion was performed according the above stated procedure. The recognition performance was evaluated under verification and identification scenarios. Based on the participating information sources the results have been reported for Intra modal fusion and Inter modal fusion. Intra modal fusion refers to combining information obtained from different algorithms but from the same trait (e.g. PCA and ICA vector for face images). Inter modal fusion refers to combining information from different biometric traits irrespective of the processing algorithm.



Figure 4.4: Fusion of face PCA and face ICA features. Feature level fusion clearly outperforms match-score fusion

	Equal Error Rate	Identification Error Rate
	(%)	(%)
face PCA	6.9892	4.8
Face ICA	10.395	6.6
SVM Fusion	2.1348	1
Match Score Fusion	8.0615	4.4



Figure 4.5: Fusion of face Geometry and face ICA features. Feature level fusion clearly outperforms match-score fusion

	Equal Error Rate	Identification Error Rate
	(%)	(%)
face GEO	29.638	84
Face ICA	10.395	6.6
SVM Fusion	1.2474	0.2
Match Score Fusion	9.3322	9.2

Table 4.2: Performance enumerated for Identification and Verification case



Figure 4.6: Fusion of face Geometry and face ICA features. Feature level fusion clearly outperforms match-score fusion

	Equal Error Rate	Identification Error Rate
	(%)	(%)
face GEO	29.638	84
	6.0002	4.0
face PCA	6.9892	4.8
SVM Fusion	0.99545	0
Match Score Fusion	9.8188	14.4

Table 4.3: Performance enumerated for Identification and Verification case

4.3.1 Intramodal Fusion

Face: As per the results from the Chapter 2 and 3, fusing information from different facial feature-sets resulted in improvement over match-score fusion but did not improve performance beyond the best individual information source. Figure 4.4, shows the ROC curve when face ICA and PCA features are fused. Table 4.1 enumerates the equal error rate and classification error rate. Feature level fusion gives extremely good results and performs better than match score fusion. When poor geometric feature are combined with ICA and PCA fetures, the improvement in overall performance, see Figure 4.5 and 4.6, is extremely significant when compared to match score fusion. In all three bimodal fusion attempts, the proposed fusion scheme led to gain in performance much higher than match-score fusion. Match-score fusion was ineffective, if not derogatory, when a very poor (error > 25%) feature-set of geometric face information was fused with PCA and ICA features. SVM based feature fusion performed extremely well even when very poor features are combined with strong features.

Iris: In the chapters 2 we discussed the fusion of iris codes and with PCA and ICA iris features. Iris codes individually, demonstrated good performance (~95%) and any direct fusion, at match-score and feature level, met with limited success. When feature level fusion is carried out using SVM, intra modal fusion of any two feature sets leads to a dramatic improvement in performance. Figure 4.7 and 4.8, shows the fusion ROC curves when iris ICA and PCA feature combined with IrisCodes. There is a sharp improvement in performance with feature level fusion when compared to match score fusion. From

Table 4.6 and Figure 4.9, it is seen that when PCA and ICA iris information is combined it leads to an improvement in EER by 7%. But the identification performance did not improve, this can be explained by the inherent correlation between ICA and PCA feature sets.

By applying the proposed SVM based fusion scheme, performance of an iris based system can be improved immensely by combining information from any two feature sets extracted by different algorithms but from the same image set.

Iris codes are very discriminative local features. When they are combined with ICA and PCA information using the before explained transformation scheme, almost perfect (\sim 0%) recognition performance is achieved (Table 4.4 and 4.5).

Intra modal fusion, for iris and face, shows high potential of improving a system which is based only on one modality. An interesting observation is the exceptional improvement in performance observed when local features, like Iris codes or Face geometry, are combined with ICA and PCA feature of each trait respectively. In spite of the learned optimal hyperplane, the inherent correlation between PCA and ICA feature hampers the over fused performance.

The results show that even if one utilized the same image sample (iris or face) and if 2 or more sets of independent feature vectors can be extracted from it, then the proposed scheme of feature level fusion yields improved performance. This implies capability of achieving much higher performance from existing systems at no additional hardware costs.



Figure 4.7: Fusion of IrisCodes and PCA features. Feature level fusion outperforms match score fusion

	Equal Error Rate	Identification Error Rate
	(%)	(%)
Iris PCA	17.349	22
Iris ICA	17.731	22.8
SVM Fusion	10.061	22.6
Match Score Fusion	17.584	22.6

Table 4.4: Performance enumerated for Identification and Verification case



Figure 4.8: Fusion of IrisCodes and ICA features. Feature level fusion outperforms match score fusion.

	Equal Error Rate	Identification Error Rate
	(%)	(%)
iris gabor	5.9407	1.4
Iris ICA	17.731	22.8
SVM Fusion	1.8267	0.2
Match Score Fusion	9.9673	5.4

Table 4.5: Performance enumerated for Identification and Verification case

4.3.2 Intermodal Fusion

In chapter 2, face and fingerprint feature sets were fused with iris codes. The match scores generated from direct summation of transformed and fused feature vectors demonstrated limited success. We will now demonstrate the effectiveness of the proposed feature fusion scheme when feature-sets from different biometric traits are combined.

Figures 4.10, 4.11 and 4.12 demonstrate the ROC curves when face PCA, ICA and Geometric features are, respectively, fused with IrisCodes. Table 4.7, 4.8 and 4.9, enumerate the equal error rate and classification error rates.

Using the proposed techniques when Iris codes and any face feature-set (ICA, PCA and GEOMETRIC extracted from facial images) leads to verification error (equal error rate) of less than 1% and 0% identification error. It is encouraging to see that the even the combination of very poor geometric face feature set and IrisCodes yielded high performance but it should be noted that our system was not trained on mutually exclusive training and testing data.

Figures 4.13 demonstrate the ROC curves when fingerprint minutia information is fused with IrisCodes. Table 4.10; enumerate the equal error rate (Verification Error Rate) and classification error rates.

Fingerprint minutia information is transformed and is combined with transformed iris codes as explained in the previous chapter. The fused match score is calculated as per the learned SVM hyperplane.



Figure 4.10: Fusion of IrisCodes with face PCA features. Feature level fusion outperforms match score fusion.

	Equal Error Rate	Identification Error Rate
	(%)	(%)
iris gabor	5.9407	1.4
face PCA	6.9892	4.8
SVM Fusion	0.65505	0
Match Score Fusion	2.3859	0.4

Table 4.7: Performance enumerated for Identification and Verification case



Figure 4.11: Fusion of IrisCodes with face ICA features. Feature level fusion outperforms match score fusion.

	Equal Error Rate	Identification Error Rate
	(%)	(%)
iris gabor	5.9407	1.4
face ICA	10.395	6.6
SVM Fusion	0.85313	0
Match Score Fusion	5.253	1.6

Table 4.8: Performance enumerated for Identification and Verification case



Figure 4.12: Fusion of IrisCodes with face Geometric features. Feature level fusion outperforms match score fusion.

	Equal Error Rate	Identification Error Rate
	(%)	(%)
Iris gabor	5.9407	1.4
face GEO	29.638	84
SVM Fusion	1.2082	0.2
Match Score Fusion	15.573	8.4

Table 4.9: Performance enumerated for Identification and Verification case



Figure 4.13: Fusion of IrisCodes with fingerprint minutia features. Feature level fusion and match score fusion perform equally well.

	Equal Error Rate	Identification Error Rate
	(%)	(%)
finger minutia	5.5327	2.6
iris gabor	5.9407	1.4
SVM Fusion	1.8369	0.4
Match Score Fusion	1.021	0.2

Table 4.10: Performance enumerated for Identification and Verification case

4.4 Summary

The objective of this chapter is to demonstrate that feature level fusion can be consistently more effective than match-score fusion. To achieve this goal we have developed a Support Vector Machine based fusion framework that can combine information from any two biometric sources. We have developed further on our concept of local distance kernels and proposed the use of support vector theory to learn an optimal separating hyperplane. An effective and simple method to compute fused matchscore is introduced by projecting the distance vector on the learned hyperplane.

We conducted extensive experiments and demonstrated the effectiveness of our proposed method for identification and verification scenario. Feature sets from various modalities and different algorithms were combined using techniques detailed in chapter 2.

The results showed drastic improvement in performance when any combinations of features were selected for fusion. It is very worthwhile to mention about the fusion results achieved for the following two generic cases

- Fusion of 2 strong biometric feature-sets.
- Fusion of weak features with strong features.

We assume a strong biometric feature sets to be the one less than 5% equal error rate. Two such combinations are the fusion of Iris codes with face PCA features and Minutia based fingerprints. While face PCA features are extracted from using domestically developed algorithm, fingerprint images were completely processed using the VeriFinger SDK. In both cases, feature level fusion performed extremely well and clearly outperformed the match-score fusion.

When a poor feature set, one with equal error rate more than 25%, is combined with

92

strong feature set more often than not the over all fusion performance degrades. This phenomenon was consistently observed for match-score fusion but not for the proposed feature fusion framework. Geometric face features form a classical example for observing this in our experiments. When geometric face features were combined with any other biometric source we saw match-score fusion degraded performance, but when fusion was performed at the feature level there was always a gain in performance.

Chapter 5: Summary and Future Work

Biometrics is becoming an integral part of our networked society but current systems do not have 100% accuracy. The research community is putting a lot of efforts in investigation the field of multimodal biometrics. Extensive research work has been done to identify efficient methods to combine the information obtained from multiple sources. Match level fusion and decision level fusion have been extensively studied in the literature. Feature level fusion is more complex but is capable of delivering better results. The following contributions are made by this thesis:

Generic feature fusion framework: In the past feature fusion was carried out by direct concatenation of homogenous feature vector. We have established a generic feature transformation scheme using local distance kernels. The transformed data reside in a common space where they can be fused.

Local Distance Transformation kernels: In this thesis, a new concept of local distance vectors is introduced and a methodology to achieve this mapping for different feature representation is developed.

Feature Selection: Another important contribution of this thesis is a new feature selection techniques based on thresholding. Two feature evaluation functions are defined which operate on the transformed local distance kernel vectors. These feature selection scheme, dynamically remove noise from the input feature vector hence making feature level fusion very robust.

AdaBoost: A novel application of boosting is proposed to learn a linear function to formulate reliable matching scores. It was shown by experiments that AdaBoost gives good performance, especially when poor biometric feature sets are involved.

94

SVM: Another important contribution of this thesis is the use of SVM to learn an optimal separating hyperplane. The match scores are synthesized by projecting the transformed local distance kernels vectors onto this separating hyperplane. It is important to note that due to the lack of sufficient, same data was used for training and testing and that may have resulted in feature level fused matching scores outperforming ruled based match score fusion.

All the proposed methods require a larger set of training data. As mentioned before, in our experiments, all genuine distance vectors were used for training and testing was also performed using the same data. The performance of our framework is still to be established on larger dataset where training and testing data are mutually exclusive. Al last, we conclude this thesis with a short list of research questions for the future work that may help further enhance the system. The research questions are as follows

- 1. Verify the potential of the proposed methods to larger dataset. Training and testing data should be mutually exclusive to verify generalization capability.
- Compare the obtained system with all existing match score fusion methods, where the match scores are obtained from commercial vendors.
- Extend the current parallel architecture of multimodal fusion to a cascade or hierarchical architecture.
- 4. Design more efficient local distance transformation scheme.
- 5. Investigate better feature evaluation schemes in order to remove noise.
- 6. The experiments must be performed using a larger data set.

BIBLIOGRAPHY

[1] A. K. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics*, 14(1):4–20, January 2004.

[2] S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric Recognition: Security and Privacy Concerns. *IEEE Security and Privacy Magazine*, 1(2):33–42, March-April 2003.

[3] A. K. Jain and A. Ross. Multibiometric Systems. *Communications of the ACM, Special Issue on Multimodal Interfaces*, 47(1):34–40, January 2004.

[4] Maltoni, D., Dario, M., Jain, A., & Prabhakar, S. (2003). *Handbook of Fingerprint Recognition*. London: Springer.

[5] Mansfield, A. J., & Wayman, J. L. (August 2002). *Best Practices in Testing and Reporting Performances of Biometric Devices, Version 2.01* (NPL Report CMSC 14/02). UK: National Physical Laboratory.

[6] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face Recognition by Independent Component Analysis," *IEEE Transaction on Neural Networks*, Vol 13, pp. 1450-1464, 2002.

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

[8] CASIA Iris Image Database(ver.1.0) ,http://www.sinobiometrics.com/casiairis.htm.

[9] V.Dorairaj, N.A. Schmid, and G. Fahmy, "Performance Evaluation of Iris Based Recognition System Implementing PCA and ICA Encoding Techniques", *Proc. of SPIE Conference on Biometric Technology for Human Identification*, Orlando, USA, Mar-Apr, 2005.

[10] J. Daugman, "How iris recognition works," in ICIP02, 2002, pp. I: 33-36.

[11] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. FVC2004: Third Fingerprint Verification Competition. In *Proceedings of International Conference on Biometric Authentication*, pages 1–7, Hong Kong, China, July 2004.

[12] C. Wilson, A. R. Hicklin, H. Korves, B. Ulery, M. Zoepfl, M. Bone, P. Grother, R. J. Micheals, S. Otto, and C. Watson. Fingerprint Vendor Technology Evaluation 2003: Summary of Results and Analysis Report. NIST Internal Report 7123; available at http://fpvte.nist.gov/report/ir_7123_summary.pdf, June 2004.

[13] P. J. Philips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone. FRVT2002: Overview and Summary. Available at http://www.frvt.

org/FRVT2002/documents.htm.

[14] D. A. Reynolds, W. Campbell, T. Gleason, C. Quillen, D. Sturim, P. Torres Carrasquillo, and A. Adami. The 2004 MIT Lincoln Laboratory Speaker Recognition System. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, March 2005.

[15] L. Hong, A. K. Jain, and S. Pankanti. Can Multibiometrics Improve Performance? In *Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies*, pages 59–64, New Jersey, U.S.A., October 1999.

[16] A. Ross and A. K. Jain. Information Fusion in Biometrics. *Pattern Recognition Letters, Special Issue on Multimodal Biometrics*, 24(13):2115–2125, 2003.

[17] L. Hong and A. K. Jain. Integrating Faces and Fingerprints for Personal Identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1295–1307, December 1998.

[18] R. Snelick, U. Uludag, A. Mink, M. Indovina, and A. K. Jain. Large Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):450–455, March 2005.

[19] C. Sanderson and K. K. Paliwal. Information Fusion and Person Verification using speech and face information. Research Paper IDIAP-RR 02-33, IDIAP, September 2002.

[20] S. S. Iyengar, L. Prasad, and H. Min. *Advances in Distributed Sensor Technology*. Prentice Hall, 1995.

[21] A. Ross and A. K. Jain. Fingerprint Mosaicking. In *Proceedings of International Conference on Acoustic Speech and Signal Processing (ICASSP)*, pages 4064–4067, Florida, U.S.A., May 2002.

[22] Y. S. Moon, H.W. Yeung, K. C. Chan, and S. O. Chan. Template Synthesis and Image Mosaicking for Fingerprint Registration: An Experimental Study. In *Proceedings* of 89 International Conference on Acoustic Speech and Signal Processing (ICASSP), volume 5, pages 409–412, Quebec, Canada, May 2004.

[23] A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain. Personal Verification Using Palmprint and Hand Geometry Biometric. In *Proceedings of Fourth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 668–678, Guildford, U.K., June 2003.

[24] A. Ross and R. Govindarajan. Feature Level Fusion Using Hand and Face Biometrics. In *Proceedings of of SPIE Conference on Biometric Technology for Human Identification*, volume 5779, pages 196–204, Florida, U.S.A., March 2005.

[25] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley & Sons, 97

2001.

[26] K. Woods, K. Bowyer, and W. P. Kegelmeyer. Combination of Multiple Classifiers using Local Accuracy Estimates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):405–410, April 1997.

[27] K. Chen, L. Wang, and H. Chi. Methods of Combining Multiple Classifiers with Different Features and Their Applications to Text-Independent Speaker Identification. *International Journal of Pattern Recognition and Artificial Intelligence*, 11(3):417–445, 1997.

[28] L. Lam and C. Y. Suen. Application of Majority Voting to Pattern Recognition: An Analysis of Its Behavior and Performance. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 27(5):553–568, 1997.

[29] L. Lam and C. Y. Suen. Optimal Combination of Pattern Classifiers. *Pattern Recognition Letters*, 16:945–954, 1995.

[30] L. Xu, A. Krzyzak, and C. Y. Suen. Methods for Combining Multiple Classifiers and their Applications to Handwriting Recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(3):418–435, 1992.

[31] J. Daugman. Combining Multiple Biometrics. Available at http://www.cl. cam.ac.uk/users/jgd1000/combine/combine.html.

[32] T. K. Ho, J. J. Hull, and S. N. Srihari. Decision Combination in Multiple Classifier Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1):66–75, January 1994.

[33] Y. Wang, T. Tan, and A. K. Jain. Combining Face and Iris Biometrics for Identity Verification. In *Proceedings of Fourth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 805–813, Guildford, U.K., June 2003.

[34] P. Verlinde and G. Cholet. Comparing Decision Fusion Paradigms using k-NN based Classifiers, Decision Trees and Logistic Regression in a Multi-modal Identity Verifcation Application. In *Proceedings of Second International Conference on Audioand Video-Based Biometric Person Authentication (AVBPA)*, pages 188–193, Washington D.C., U.S.A., March 1999.

[35] V. Chatzis, A. G. Bors, and I. Pitas. Multimodal Decision-level Fusion for Person Authentication. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 29(6):674–681, November 1999.

[36] J. Kittler, M. Hatef, R. P. Duin, and J. G. Matas. On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.

[37] S. Prabhakar and A. K. Jain. Decision-level Fusion in Fingerprint Verification. *PatternRecognition*, 35(4):861–874, 2002.

[38] E. S. Bigun, J. Bigun, B. Duc, and S. Fischer. Expert Conciliation for Multimodal Person Authentication Systems using Bayesian Statistics. In *Proceedings of First International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 291–300, Crans-Montana, Switzerland, March 1997.

[39] A. K. Jain and A. Ross. Learning User-specific Parameters in a Multibiometric System. In *Proceedings of International Conference on Image Processing*, pages 57–60, New York, USA, September 2002.

[40] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The Extended M2VTS Database. In *Proceedings of Second International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 72–77, Washington D.C., U.S.A., March 1999.

[41] M. Indovina, U. Uludag, R. Snelick, A. Mink, and A. K. Jain. Multimodal Biometric Authentication Methods: A COTS Approach. In *Proceedings of Workshop on Multimodal User Authentication*, pages 99–106, Santa Barbara, USA, December 2003.

[42] National Institute of Standards The Image Group of the Information Access Division and Technology. Biometric Scores Set - Release 1. Available at http://www.itl.nist.gov/iad/894.03/biometricscores, September 2004.

[43] P. Verlinde, P. Druyts, G. Cholet, and M. Acheroy. Applying Bayes based Classifiers for Decision Fusion in a Multi-modal Identity Verification System. In *Proceedings of International Symposium on Pattern Recognition "In Memoriam Pierre Devijver*", Brussels, Belgium, February 1999.

[44] R. Snelick, M. Indovina, J. Yen, and A. Mink. Multimodal Biometrics: Issues in Design and Testing. In *Proceedings of Fifth International Conference on Multimodal Interfaces*, pages 68–72, Vancouver, Canada, November 2003.

[45] R. Brunelli and D. Falavigna. Person Identification Using Multiple Cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):955–966, October 1995.

[46] Rohin Govinddarajan, "Feature Level Fusion in Multimodal Biometrics", M.S. Thesis, 2004

[47] Arun Ross, "Information Fusion in Fingerprint Authentication", Ph.D. Thesis, 2003.

[48] P. J. Huber. Robust Statistics. John Wiley & Sons, 1981.

[49] R. Cappelli, D. Maio, and D. Maltoni. Combining Fingerprint Classifiers. In 99

Proceedings of First International Workshop on Multiple Classifier Systems, pages 351–361, June 2000.

[50] F. R. Hampel, P. J. Rousseeuw, E. M. Ronchetti, and W. A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 1986.

[51] F. Mosteller and J. W. Tukey. *Data Analysis and Regression: A Second Course in Statistics*. Addison-Wesley, 1977.

[52] A. K. Jain, L. Hong, S. Pankanti, and R. Bolle. An Identity Authentication System Using Fingerprints. *Proceedings of the IEEE*, 85(9):1365–1388, 1997.

[53] G.V. Trunk, "A Problem of Dimensionality: A Simple Example," *IEEE Trans. PatternAnalysis and Machine Intelligence*, Vol. 1, No. 3, pp. 306–307, 1979.

[54] A. K. Jain, A. Ross, and S. Pankanti. A Prototype Hand Geometry-based Verification System. In *Proceedings of Second International Conference on Audio- and Videobased Biometric Person Authentication (AVBPA)*, pages 166–171, Washington D.C.,USA, March 1999.

[55] A. Bertillon. Signaletic Instructions including the theory and practice of Anthropometrical Identification, R.W. McClaughry Translation. The Werner Company, 1896.

[56] D. D. Heckathorn, R. S. Broadhead, and B. Sergeyev. A Methodology for Reducing Respondent Duplication and Impersonation in Samples of Hidden Populations. In *Annual Meeting of the American Sociological Association*, Toronto, Canada, August 1997.

[57] H. Aillisto, M. Lindholm, S. M. Makela, and E. Vildjiounaite. Unobtrusive User Identification with Light Biometrics. In *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, pages 327–330, Tampere, Finland, October 2004.

[58] J. L. Wayman. Large-scale Civilian Biometric Systems - Issues and Feasibility. In *Proceedings of Card Tech / Secur Tech ID*, 1997.

[59] G. Givens, J. R. Beveridge, B. A. Draper, and D. Bolme. A Statistical Assessment of Subject Factors in the PCA Recognition of Human Subjects. In *Proceedings of CVPR Workshop: Statistical Analysis in Computer Vision*, June 2003.

[60] E. Newham. The Biometrics Report. SJB Services, 1995.

[61] S. Gutta, J. R. J. Huang, P. Jonathon, and H. Wechsler. Mixture of Experts for Classification of Gender, Ethnic Origin, and Pose of Human Faces. *IEEE Transactions on Neural Networks*, 11(4):948–960, July 2000.

[62] Vladimir N. Vapnik, The Nature of Statistical Learning Theory. Springer, 1995.

[63] G. Shakhnarovich, P. Viola, and B Moghaddam. A Unified Learning Framework for 100

Real Time Face Detection and Classification. In *Proceedings of International Conference* on Automatic Face and Gesture Recognition, Washington D.C., USA, May 2002.

[64] Karthik Nandakumar, "Integration of Multiple Cues in Biometric Systems", M.S. Thesis, 2005.

[65] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition, Springer Verlag, 2003.

[66] A.K. Jain, L. Hong and R. Bolle, "On-line Fingerprint Verification", *IEEE Transactions on PAMI*, Vol. 19, No. 4, pp. 302-314, 1997.

[67] A. Ross, S. Dass and A. K. Jain, "A Deformable Model for Fingerprint Matching", *Pattern Recognition*, Vol. 38, No. 1, pp. 95-103, 2005.

[68] Y. H. Kwon and N. V. Lobo. Age Classification from Facial Images. In *Proceedings* of *IEEE Conference on Computer Vision and Pattern Recognition*, pages 762–767, April 1994.

[69] A. Lanitis, C. Draganova, and C. Christodoulou. Comparing Different Classifiers for Automatic Age Estimation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(1):621–628, February 2004.

[70] J.G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," Vision Research, vol. 20, pp. 847–856, 1980.

[71] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, pp. 529–541, 1981.

[72] Y. Freund. An adaptive version of the boost by majority algorithm. *Machine Learning*, 43(3):293-318, June 2001.

[73] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory: Eurocolt* '95, pages 23–37. Springer-Verlag, 1995.

[74] OSU SVM toolbox for MATLAB. Available at http://sourceforge.net/projects/svm.

[75] S. C. Dass, K. Nandakumar, and A. K. Jain. A Principled Approach to Score Level Fusion in Multimodal Biometric Systems. In *Proceedings of Fifth International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA) (To appear)*, New York, U.S.A., July 2005.

[76] Schapire, Y. Freund, P. Bartlett, and W. S. Lee. Boosting the margin: A new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5):1651-1686, October 1998.