Chapter 2

INFORMATION FUSION IN BIOMETRICS

2.1 Introduction

Information fusion has a long history and the theory of multiple classifier systems (MCS) has been rigorously studied over the past several years (Ghosh, 2002). In fact information fusion is an integral part of various application domains ranging from automatic target recognition (ATR) and remote sensing to weather forecasting, object tracking and robotics. The concept of fusion has been studied under several different terminologies (Ho, 2002; Kuncheva et al., 2001), including

- stacked generalizations (Wolpert, 1990)

- classifier ensembles (Drucker et al., 1994)

- hybrid methods (Bunke and Kandel, 2002)

- cooperative agents (Tan, 1997)

- dynamic classifier selection (Woods et al., 1997)

- opinion pool (Benediktisson and Swain, 1992)

- sensor fusion (Iyengar et al., 1995)

- mixture of experts (Jacobs et al., 1991)

- consensus aggregation (Benediktisson and Swain, 1992)

- divide-and-conquer classifiers (Chiang and Fu, 1994)

- social choice functions (Arrow, 1963).

Ho, 2002 states that there has been a paradigm shift in the approach to solving pattern recognition problems:

> Instead of looking for the best set of features and the best classifier, now we look for the best set of classifiers and then the best combination method.

The goal of information fusion, therefore, is to determine the best set of experts in a given problem domain and devise an appropriate function that can optimally combine the decisions rendered by the individual experts (Figure 2.1). A similar philosophy has been advocated by several researchers, including Minsky (Minsky, 1991) who states

> To solve really hard problems, we'll have to use several different representations ....

and,

> It is time to stop arguing over which type of pattern classification technique is best because that depends on our context and goal. Instead we should work at a higher level of organization and discover how to build managerial systems to exploit the different virtues and evade the different limitations of each of these ways of comparing things.

We briefly examine the role of data fusion in different applications. The purpose is to indicate to the reader the diversity of scientific fields that rely on information fusion schemes.

1 **Weather forecasting**: An elaborate weather forecasting system relies on the evidence provided by diverse sources of information such as geostationary meteorological satellites, weather balloons/planes, ground stations, radars, automated buoys, etc. in order to compute geophysical parameters of interest. These geophysical parameters are then collectively interpreted by an automated system to facilitate weather forecasting. The system also relies on previous results of weather prediction (temporal information) to continually refine its outputs (Palmer, 2000).

2 **UAV swarms**: A group of unmanned aerial vehicles (UAVs), searching for a mobile evasive target in a potentially hazardous environment, has to determine a flight arrangement that optimizes the integrated sensing capability of component UAVs (Vachtsevanos et al., 2004). In this type of scenario, an optimal flight configuration has to be derived based on the nature of the data acquired by the individual UAVs, constraints on the amount of information that can be transmitted between UAVs and the possibility of losing a UAV (e.g., UAV missing in action). An appropriate fusion architecture is necessary to accommodate the dynamics of the topology as well as the reliability of the sensor data obtained in order to generate efficient actions.

3 **Object detection**: Many applications attempt to detect and establish the trajectories of objects based on the evidence supplied by multiple image modalities. The fusion of visible and non-visible information pertaining to
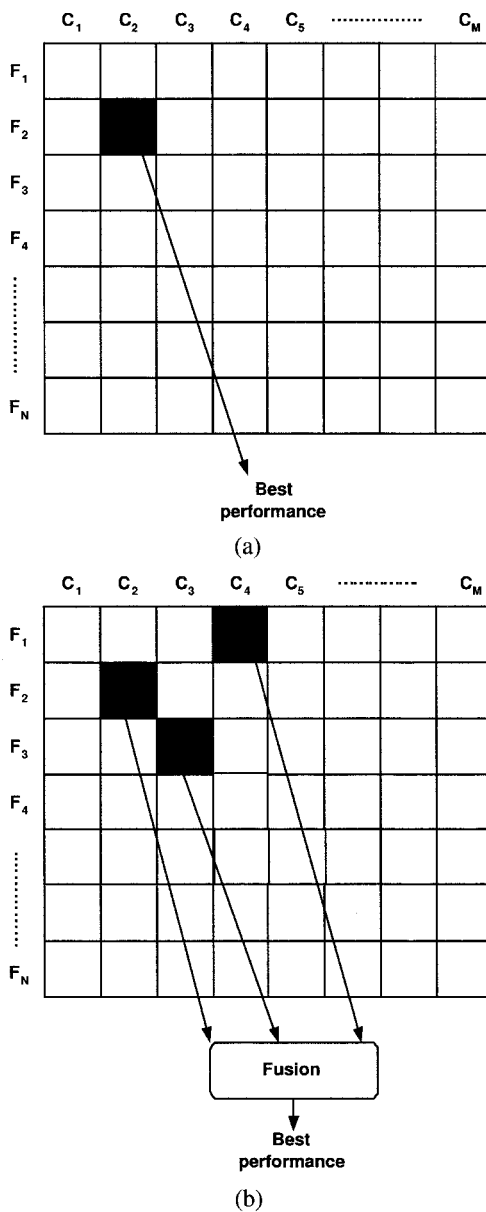
Figure 2.1. Two general approaches to solving a pattern recognition problem. Each cell in this diagram indicates the application of a particular classifier, C_i, to a specific pattern representation (i.e., feature set), F_j. The approach in (a) is to determine the best set of features and the best classifier, while in (b) the goal is to determine the best set of classifiers and an optimal fusion algorithm to integrate these classifiers. The feature sets $F_1, F_2, \ldots, F_N$ do not have to be mutually exclusive.

different wavelengths in the electromagnetic spectrum (e.g., radar and infra-red images, or thermal and visible spectrum images) can assist in estimating the location and kinematic features of objects such as T-72 tanks or a squad of soldiers in a night-time battlefield. These applications rely on image fusion methodologies to combine multiple modalities (Blum and Liu, 2006).

4 **Robot navigation**: A robot is typically fitted with a variety of sound, light, image, range, proximity and force sensors that permit it to record its environment. In order to determine a suitable action (e.g., move right or tilt camera at a certain angle), the data acquired using these multiple sensors are processed simultaneously (Abidi and Gonzalez, 1992). Sensor integration in a modular framework is a challenging task since it entails the reconciliation of non-commensurate data.

5 **Land mine detection**: Several types of sensor technologies are being used to detect buried land mines. These include electromagnetic induction (EMI), ground penetrating radar (GPR), infra-red imaging (IR), quadrupole resonance (QR), chemical detectors and sensors of acoustically induced surface vibrations (Gunatilaka and Baertlein, 2001). In many cases, the data presented by these multiple sensors are concurrently used to improve the accuracy of land mine detection algorithms.

2.2   Fusion in biometrics

Humans recognize one another based on the evidence presented by multiple biometric characteristics (behavioral or physical) in addition to several contextual details associated with the environment. The recognition process itself may be viewed as the reconciliation of evidence pertaining to these multiple modalities. Each modality on its own cannot always be reliably used to perform recognition. However, the consolidation of information presented by these multiple experts can result in the accurate determination or verification of identity.

Biometric systems can also be designed to recognize a person based on information acquired from multiple biometric sources. Such systems, known as *multibiometric* systems, can be expected to be more accurate due to the presence of multiple pieces of evidence (Hong et al., 1999). Multibiometric systems offer several advantages over traditional (uni)biometric systems. Some of these advantages are listed below.

1 Multibiometric systems can offer substantial improvement in the matching accuracy of a biometric system depending upon the information being combined and the fusion methodology adopted. Thus, the FAR and the FRR of the verification system can be reduced simultaneously. Furthermore, the availability of multiple sources of information increases the feature space

available to individuals and, hence, the capacity of an identification system may be increased in order to accommodate more individuals.

2 Multibiometrics addresses the issue of non-universality or insufficient population coverage. If a person's dry fingers prevent him from successfully enrolling into a fingerprint system, then the availability of another biometric trait, say iris, can aid in the inclusion of this individual in the identity management system. A certain degree of flexibility is achieved when a user enrolls into the system using several different traits (e.g., face, voice, fingerprint, iris, hand) while only a subset of these traits (e.g., face and voice) is requested during authentication based on the nature of the application under consideration and the convenience of the user.

3 It becomes increasingly difficult (if not impossible) for an impostor to spoof multiple biometric traits of a legitimately enrolled individual. If each subsystem indicates the probability that a particular trait is a 'spoof', then appropriate fusion schemes can be employed to determine if the user, in fact, is an impostor. Furthermore, by asking the user to present a random subset of traits at the point of acquisition, a multibiometric system facilitates a challenge-response type of mechanism, thereby ensuring that the system is interacting with a *live* user. Note that a challenge-response mechanism can be initiated in unibiometric systems also (e.g., system prompts "Please say 1-2-5-7", "Blink twice and move your eyes to the right", "Change your facial expression by smiling", etc.).

4 Multibiometric systems also effectively address the problem of noisy data. When the biometric signal acquired from a single trait is corrupted with noise, the availability of other (less noisy) traits may aid in the reliable determination of identity. Some systems take into account the *quality* of the individual biometric signals during the fusion process. This is especially important when recognition has to take place in adverse conditions where certain biometric traits cannot be reliably extracted. For example, in the presence of ambient noise, when an individual's voice characteristics cannot be accurately measured, the facial characteristics may be used by the multibiometric system to perform authentication. Estimating the quality of the acquired data is in itself a challenging problem but, when appropriately done, can reap significant benefits in a multibiometric system.

5 These systems also help in the *continuous* monitoring or tracking of an individual in situations when a single trait is not sufficient. For example, a person walking down a crowded aisle can be recognized using his face and gait cues. However, depending upon the distance and pose of the subject with respect to the camera, both these characteristics may not be simultaneously

available. Therefore, either (or both) of these traits can be used depending upon the situation.

6 A multibiometric system may also be viewed as a fault tolerant system which continues to operate even when certain biometric sources become unreliable due to sensor or software malfunction, or deliberate user manipulation. The notion of fault tolerance is especially useful in large-scale authentication systems handling a large number of users (e.g., a border control system).

2.3   Issues in designing a multibiometric system

Multibiometric systems rely on the evidence presented by multiple sources of biometric information. An information fusion scheme in the context of biometrics raises several design questions as we will see shortly. Primary among these is the design of a suitable human computer interface (HCI) that would permit the efficient acquisition of an individual's biometric information. An appropriately designed interface can ensure that multiple pieces of evidence pertaining to an individual's identity are reliably acquired whilst causing minimum inconvenience to the user (Oviatt, 2003). Consider the user interface shown in Figure 2.2 which acquires the face, fingerprint and hand geometry information of an individual. This particular arrangement of the scanners might make it tedious for the person to interact with the system since the hand geometry and fingerprint sensors are spatially separated requiring the individual to explicitly interact with these two sensors. A better arrangement would be to integrate these two sensors into a single device thereby capturing the hand and fingerprint modalities simultaneously with minimum user inconvenience. As one moves from unimodal to multimodal systems, it is imperative that HCIs be carefully designed.

Some of the other factors that impact the design and structure of a multibiometric system are described below.

1 **Cost benefits**: What is the tradeoff between the added cost and the improvement in matching performance? The cost is a function of the number of sensors deployed, the time taken to acquire the biometric data, the storage requirements, the processing time of the algorithm and the perceived (in)convenience experienced by the user.

2 **Determining sources of biometric information**: What are the various sources of biometric information that can be used in a multibiometric system? Which of these sources are relevant to the application at hand?

3 **Acquisition and processing sequence**: Should the data corresponding to multiple information sources (e.g., modalities) be acquired simultaneously or at different time instances, as the need arises, in a serial fashion? Simi-

Figure 2.2. A multimodal interface to acquire face, fingerprint and hand geometry images of a person. A well designed interface can enhance user convenience and ensure that multiple sources of evidence are reliably acquired. In this example, integrating the hand and fingerprint input devices into a single unit may be beneficial as it would reduce the burden on the individual to explicitly interact with two spatially separated devices.

larly, should the information acquired be processed sequentially or simultaneously?

4 **Type of information**: What type of information or attributes (i.e., features, match scores, decisions, etc.) is to be fused? What is the impact of correlation among the sources of information on the performance of the fusion system?

5 **Fusion methodology**: What fusion scheme should be employed to combine the information presented by multiple biometric sources? Is it possible to predict the performance gain obtained using different fusion methodologies in order to determine the optimal one?

To make a business case for multibiometric systems, it is necessary to measure the performance gain as a function of the cost incurred in deploying such a system. The addition of multiple sensors, for example, would increase the cost of the system significantly especially if the user interface has to be altered in order to accommodate new devices. Furthermore, the throughput of the system can potentially decrease if the time taken to acquire the biometric data corresponding to multiple traits is high. While it is possible to quantify the additional cost of sensors and the increased authentication time, it is substantially difficult to quantify the system's ability to deter potential impostors from launching a

spoof attack (if multiple traits are used). Similarly, it may not be possible to quantify the time needed (number of authentication attempts) for user habituation and the potential inconvenience as perceived by the user. In light of this, the benefit of a multibiometric system is often evaluated based on its matching accuracy, the number of users that can be accommodated in the system, the cost of adding new sensors and the additional time required for acquiring and processing multiple traits both during enrollment and authentication.
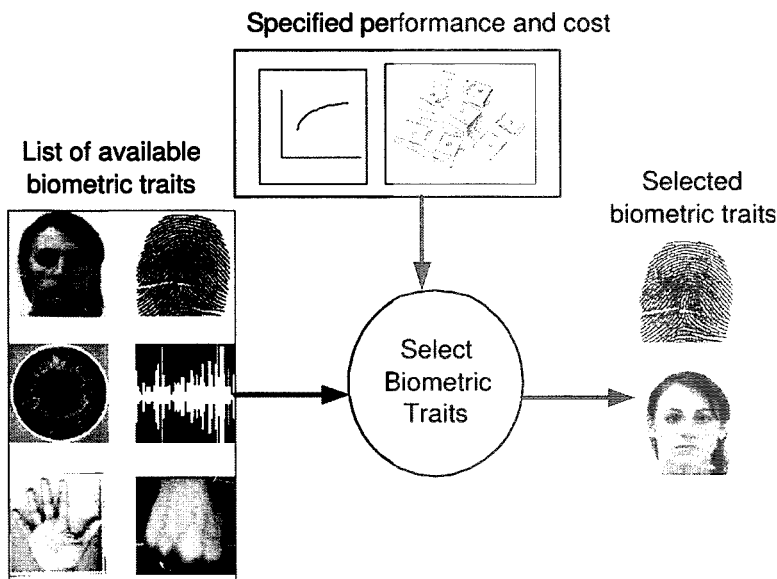

Figure 2.3. Multimodal biometric systems utilize different body traits to establish identity. In principle, a large number of traits can be used to improve the identification accuracy. In practice, factors such as cost of deployment, finite training sample size, throughput time and user training will limit the number of traits used in a particular application.

2.4   Sources of multiple evidence

What are the *sources* of information that can be considered in a multibiometric system? We address this question by introducing some terminology to describe the various scenarios that are possible to obtain multiple sources of evidence (see Figure 2.4). In the first four scenarios described below, information fusion is accomplished using a single trait, while in the fifth scenario multiple traits are used.

1 **Multi-sensor systems**: In these systems, a single biometric trait is imaged using multiple sensors in order to extract diverse information from

Left Profile   Frontal   Right Profile

Multi-sample

Fingerprint

Optical

Multi-sensor   **Multi-modal**

Face

Solid-state

Multibiometric
Sources

Iris

Multi-algorithm   **Multi-instance**

Right Eye   Left Eye
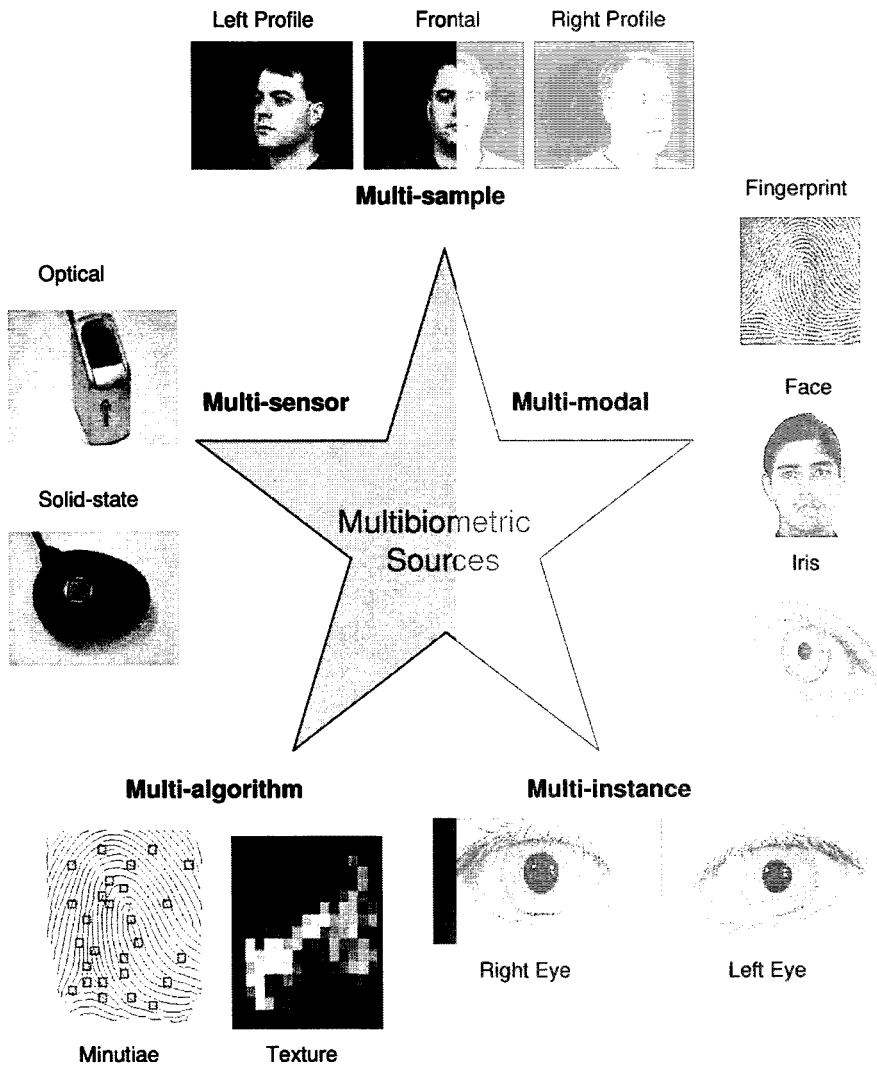
Minutiae   Texture

Figure 2.4. The various sources of information in a multibiometric system: multi-sensor, multi-algorithm, multi-instance, multi-sample and multimodal. In the first four scenarios, a single biometric trait provides multiple sources of evidence. In the fifth scenario, different biometric traits are used to obtain evidence.

(spatially) registered images. For example, a system may record the two-dimensional texture content of a person's face using a CCD camera and the three-dimensional surface shape of the face using a range sensor in order to perform authentication. The introduction of a new sensor (in this case, the

range sensor) to measure the facial surface variation increases the cost of the multibiometric system. However, the availability of multi-sensor data pertaining to a single trait can assist the *segmentation* and *registration* procedures also (Bendjebbour et al., 2001) besides improving matching accuracy.

Marcialis and Roli, 2004a discuss a scheme to fuse the fingerprint information of a user obtained using an optical and a capacitive fingerprint sensor (spatial registration between the two sensors is not necessary in this case). The authors, in their work, indicate that the two sensors provide complementary information thereby resulting in better matching accuracy. They also suggest the possibility of employing a dynamic sensor selection scheme (Woods et al., 1997; Giacinto and Roli, 2001) wherein, based on the nature of the input data obtained from the two sensors, the information from only one of the sensors may be used to perform recognition. Chen et al., 2005a examine the face images of an individual obtained using a thermal infrared camera and a visible light camera. They demonstrate that integrating the evidence supplied by these two images (both at the score-level and rank-level) improves matching performance. Socolinsky and Selinger, 2004 and Heo et al., 2004 also demonstrate the benefits of using thermal infrared and visible light imagery for face recognition.

2 **Multi-algorithm systems**: In these systems, the same biometric data is processed using multiple algorithms. For example, a texture-based algorithm and a minutiae-based algorithm can operate on the same fingerprint image in order to extract diverse feature sets that can improve the performance of the system (Ross et al., 2003). This does not require the use of new sensors and, hence, is cost-effective. Furthermore, the user is not required to interact with multiple sensors thereby enhancing user convenience. However, it does require the introduction of new feature extractor and/or matcher modules which may increase the computational requirements of the system (Figure 2.5).

A multi-algorithm system can use multiple feature sets (i.e., multiple representations) extracted from the same biometric data or multiple matching schemes operating on a single feature set. Lu et al., 2003 discuss a face recognition system that employs three different feature extraction schemes (Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discriminant Analysis (LDA)) to encode (i.e., represent) a single face image. The authors postulate that the use of different feature sets makes the system robust to a variety of intra-class variations normally associated with the face biometric. Experimental results indicate that combining multiple face classifiers can enhance the identification rate of the biometric system. Han and Bhanu, 2005 present a context-based gait recognition system which invokes and combines two gait recognition classifiers based

on the walking surface. A probabilistic approach is used to combine the participating classifiers. The authors demonstrate that using context information in a fusion framework has the potential to improve the identification rate of the system. Jain et al., 1999c fuse the evidence of three different fingerprint matchers to determine the similarity between two minutiae sets. The three minutiae matchers considered in their system are based on the Hough transform, one-dimensional string matching and two-dimensional dynamic programming. They observe that the matching performance obtained by combining two of the three matchers is comparable to combining all the three matchers. Factors such as the correlation between component algorithms, the disparity in their matching accuracies, and the fusion methodology adopted significantly impact the performance obtained after fusion.
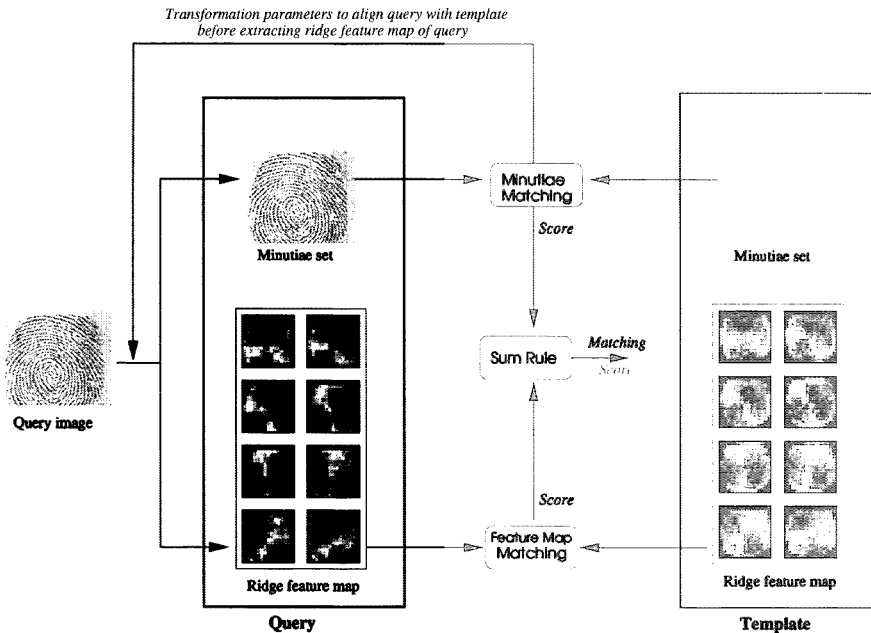

Figure 2.5. The multi-algorithm fingerprint matcher designed by Ross et al., 2003. The system utilizes both minutiae and texture information to represent and match two fingerprint images (query and template). The minutiae matching module provides the transformation parameters necessary to align the query image with the template before extracting the texture information from the former. The texture information is represented using ridge feature maps.

3 **Multi-instance systems**: These systems use multiple instances of the same body trait and are also referred to as multi-unit systems in the literature. For example, the left and right index fingers, or the left and right irises of

an individual may be used to verify an individual's identity. These systems
generally do not necessitate the introduction of new sensors nor do they en-
tail the development of new feature extraction and matching algorithms and
are, therefore, cost efficient. However, in some cases, a new sensor arrange-
ment might be necessary in order to facilitate the simultaneous capture of
the various units/instances. Automated Fingerprint Identification Systems
(AFIS), that obtain ten-print information from a subject, can benefit from
sensors that are able to rapidly acquire impressions of all ten fingers. Multi-
instance systems are especially beneficial for users whose biometric traits
cannot be reliably captured due to inherent problems. For example, a single
finger may not be a sufficient discriminator for a person having dry skin.
However, the integration of evidence across multiple fingers may serve as
a good discriminator in this case. Similarly, an iris system may not be able
to image significant portions of a person's iris due to drooping eyelids. The
consideration of both the irides will result in the availability of more texture
information that can be used to establish the individual's identity in a more
reliable manner. Multi-instance systems are often necessary in applications
where the size of the system database (i.e., the number of enrolled individu-
als) is very large (FBI's database currently has ~ 50 million ten-print images
and multiple fingers provide additional discriminatory information).

4 **Multi-sample systems**: A single sensor may be used to acquire multiple
  samples of the same biometric trait in order to account for the variations
  that can occur in the trait, or to obtain a more complete representation of
  the underlying trait. A face system, for example, may capture (and store)
  the frontal profile of a person's face along with the left and right profiles
  in order to account for variations in the facial pose. Similarly, a fingerprint
  system equipped with a small size sensor may acquire multiple dab prints
  of an individual's finger in order to obtain images of various regions of the
  fingerprint. A mosaicing scheme may then be used to stitch the multiple
  impressions and create a composite image. One of the key issues in a multi-
  sample system is determining the *number* of samples that have to be acquired
  from an individual. It is important that the procured samples represent the
  variability as well as the *typicality* of the individual's biometric data. To
  this end, the desired relationship between the samples has to be established
  before-hand in order to optimize the benefits of the integration strategy.
  For example, a face recognition system utilizing both the frontal- and side-
  profile images of an individual may stipulate that the side-profile image
  should be a three-quarter view of the face (Hill et al., 1997; O'Toole et al.,
  1995). Alternately, given a set of biometric samples, the system should be
  able to automatically select the "optimal" subset that would best represent

the individual's variability. Uludag et al., 2004 discuss two such schemes in the context of fingerprint recognition.

5 **Multimodal systems**: These systems combine the evidence presented by different body traits for establishing identity. For example, some of the earliest multimodal biometric systems utilized face and voice features to establish the identity of an individual (Brunelli and Falavigna, 1995). Physically uncorrelated traits (e.g., fingerprint and iris) are expected to result in better *improvement* in performance than correlated traits (e.g., voice and lip movement). The cost of deploying these systems is substantially more due to the requirement of new sensors and, consequently, the development of appropriate user interfaces. The identification accuracy can be significantly improved by utilizing an increasing number of traits although the *curse-of-dimensionality* phenomenon would impose a bound on this number. The curse-of-dimensionality limits the number of attributes (or features) used in a pattern classification system when only a small number of training samples is available (Jain and Chandrasekaran, 1982). The number of traits used in a specific application will also be restricted by practical considerations such as the cost of deployment, enrollment time, throughput time, expected error rate, user habituation issues, etc.

6 **Hybrid systems**: Chang et al., 2005 use the term *hybrid* to refer to systems that integrate a subset of the five scenarios discussed above. For example, Brunelli and Falavigna, 1995 describe an arrangement in which two speaker recognition algorithms are combined with three face recognition algorithms at the match score and rank levels via a HyperBF network. Thus, the system is multi-algorithmic as well as multimodal in its design. Similarly, the NIST BSSR1 dataset (National Institute of Standards and Technology, 2004) has match scores pertaining to two different face matchers operating on the frontal face image of an individual (multi-algorithm), and a fingerprint matcher operating on the left- and right-index fingers of the same individual (multi-instance). Hybrid systems attempt to extract as much information as possible from the various biometric modalities.

Besides the above scenarios, it is also possible to use biometric traits in conjunction with non-biometric identity tokens in order to enhance the authentication performance. For example, Jin et al., 2004 discuss a dual factor authenticator that combines a pseudo random number (present in a token) with a facial feature set in order to produce a set of user-specific compact codes known as BioCode. The pseudo random number and the facial feature sets are fixed in length and an iterated inner product is used to generate the BioCode. When an individual's biometric information is suspected to be compromised, then the token containing the random data is replaced, thereby revoking the previous authenticator. The use of biometric and non-biometric authenticators in tandem

is a powerful way of enhancing security. However, some of the inconveniences associated with traditional authenticators remain (such as "Where did I leave my token?").

Beattie et al., 2005 discuss a scenario in which biometric sensors are placed at various locations in a building in order to impart security to individual facilities/rooms (Figure 2.6). The building is partitioned into various zones based on access privileges assigned to different users of the building. The authentication decision rendered at a particular zone (for a specific user) may depend on the decisions made previously in other zones (for the same user). Furthermore, in very sensitive zones, a combination of biometric evidences may be used to validate an individual's identity, while in less sensitive zones, a single biometric evidence may be sufficient to establish identity. The fusion scheme used to combine the decisions of multiple sensors can also vary depending upon the zone that a user intends to enter. For example, the AND decision rule may be used in high security areas - a user can enter such a zone only when *all* the sensors successfully confirm the individual's identity (see Varshney et al., 2002). Therefore, the scenario described by Beattie et al., 2005 permits the inclusion of multiple fusion rules involving multiple sensors in a dynamic architecture. The presence of biometric sensors in various zones can also aid in determining an individual's location within the building.
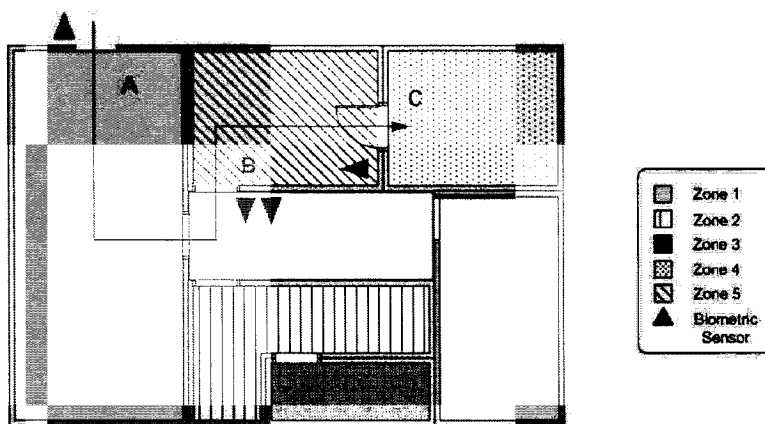

Figure 2.6. The scenario envisioned by Beattie et al., 2005 in which biometric sensors are installed at various locations within a building that is partitioned into various zones. The authentication decision rendered at a particular location for a specific user, is a function of the decisions generated at other locations previously visited by the same user. Thus, there is an integration of evidence across space and time. Moreover, the fusion rule employed at a particular site can vary depending upon the security level of the associated zone. For example, in the above illustration, a user entering site **B** has to be verified using two biometric sensors whose decisions may be combined using the AND decision rule.

2.5   Acquisition and processing architecture

As indicated earlier, the nature of the human computer interface adopted by a multibiometric system impacts its usability. Specifically, the order or sequence of biometric data acquisition has a bearing on the convenience imparted to the user. The enrollment time and the failure to enroll (FTE) rate can be substantially reduced by designing an acquisition protocol that enhances user convenience while ensuring that good quality biometric data is obtained from the user. Also, the sequence in which the procured biometric data is processed can significantly impact the throughput time in large-scale identification systems (involving millions of enrolled users) since it may be possible to arrive at an identification decision rapidly. The various types of acquisition and processing architectures are discussed below.

2.5.1   Acquisition sequence

The acquisition sequence in a multibiometric system refers to the order in which the various sources of evidence are acquired from an individual (in the case of multi-algorithm systems, only a single biometric sample is required and, therefore, the acquisition methodology is not an issue). Typically, the evidence is gathered sequentially, i.e., each source is independently obtained with a short time interval between successive acquisitions. In some cases, the evidence may be acquired simultaneously. For example, the face and iris information of a user may be obtained nearly simultaneously by utilizing two cameras housed in the same unit. Similarly, the face, voice and lip movements of a user may be acquired simultaneously by using a video camera (Frischholz and Dieckmann, 2000). Simultaneous procurement of information presents the possibility of (spatially) registering the information gleaned from multiple sources. In a multimodal face and iris system, the face image may be used to estimate the gaze direction which can then assist in localizing the iris image (in several instances, eye localization precedes face detection; therefore, the system might first detect the eyes of the subject before attempting to locate the face). Socolinsky et al., 2003 discuss a face acquisition setup that is capable of obtaining face images pertaining to the visible as well as the longwave infrared (LWIR) spectrum. The sensor captures video sequences of an individual's face by employing a CCD array and a LWIR microbolometer. The procured image pair (each of size 240x320) is co-registered to sub-pixel accuracy. This makes it possible to have a one-to-one correspondence between salient facial features present in both the images. Kumar et al., 2003 present a setup that acquires the palmprint and hand geometry details of an individual using a single camera. Simultaneously procuring multiple modalities can decrease enrollment time in multibiometric systems.

2.5.2   Processing sequence

The processing sequence adopted by a multibiometric system refers to the order in which the acquired information is used in order to render a decision. Here, the focus is not on the order of acquisition, but on the order in which the information is processed. Thus, information may be *acquired sequentially* but *processed simultaneously*.

In the serial or cascade mode, the processing of information takes place sequentially. In Figure 2.7, the fingerprint information of the user is first processed; if the fingerprint sub-system is unable to determine the identity, then the data corresponding to the face biometric is processed. In such an arrangement, the processing time can be effectively reduced if a decision is made before going through all the biometric subsystems. In the parallel mode, on the other hand, each sub-system processes its information independently at the same time and the processed information is combined using an appropriate fusion scheme (see Figure 2.8).

The cascading scheme can improve user convenience as well as allow fast and efficient searches in large scale identification tasks. For example, when a cascaded biometric system has sufficient confidence on the identity of the user after processing the first modality, the user may not be required to provide the other traits. The system can also allow the user to decide which modality he/she would present first. Finally, if the system is faced with the task of identifying the user from a large database, it can utilize the outcome of each modality to successively prune the database, thereby making the search faster and more efficient. Thus, a cascaded system may be more convenient to the user and it generally requires a shorter recognition time compared to its parallel counterpart. However, robust algorithms are essential to efficiently handle the various sequence of events that are possible. Hong and Jain, 1998 propose a cascaded system in which face recognition is used to retrieve the top n matching identities while fingerprint recognition is used to determine the final identity based on the retrieved identities only. This is significant because (i) face matching using fixed length feature vectors is generally faster than fingerprint matching; (ii) fingerprint identification is more accurate than face identification. Thus, the advantages of both modalities are exploited in this scheme (Figure 2.9).

A multibiometric system designed to operate in the parallel mode generally has a higher accuracy because it utilizes more evidence about the user for recognition. Of course, in the cascade mode, as information from multiple sources is progressively accumulated, the system is also expected to have a higher accuracy. Most multibiometric systems proposed in the literature have a parallel architecture because the primary goal of system designers has been to reduce the error rates of biometric systems (see Ross and Jain, 2003, Snelick et al., 2005 and the references therein) and not necessarily the throughput and/or processing time.
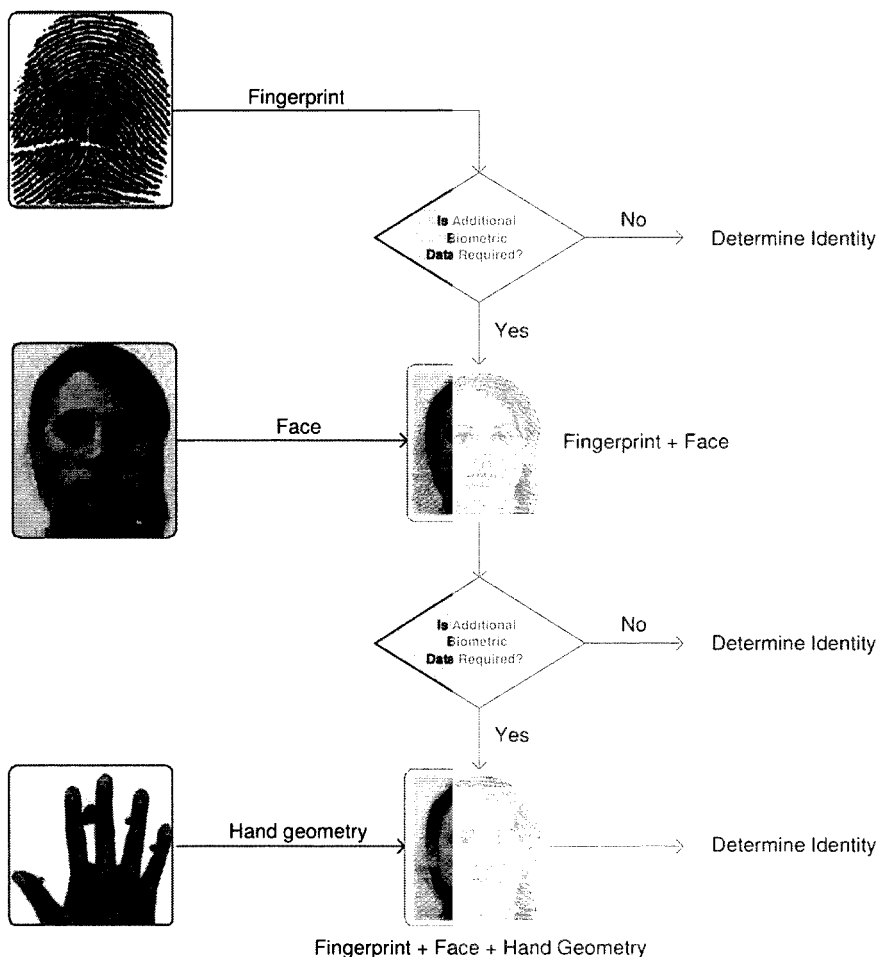
Figure 2.7. In the cascade (or serial) mode of operation, evidence is incrementally processed in order to establish identity. This scheme is also known as sequential pattern recognition. It enhances user convenience while reducing the average processing time since a decision can be made without having to acquire all the biometric traits.

Besides the two modes of operation discussed above, it is also possible to have a hierarchical (tree-like) architecture to combine the advantages of both cascade and parallel architectures (Maltoni et al., 2003). In such a scheme, a subset of the acquired modalities may be combined in parallel, while the remaining modalities may be combined in a serial fashion. Such an architecture can be dynamically determined based on the quality of the individual biometric samples as well as the possibility of encountering missing biometric data.
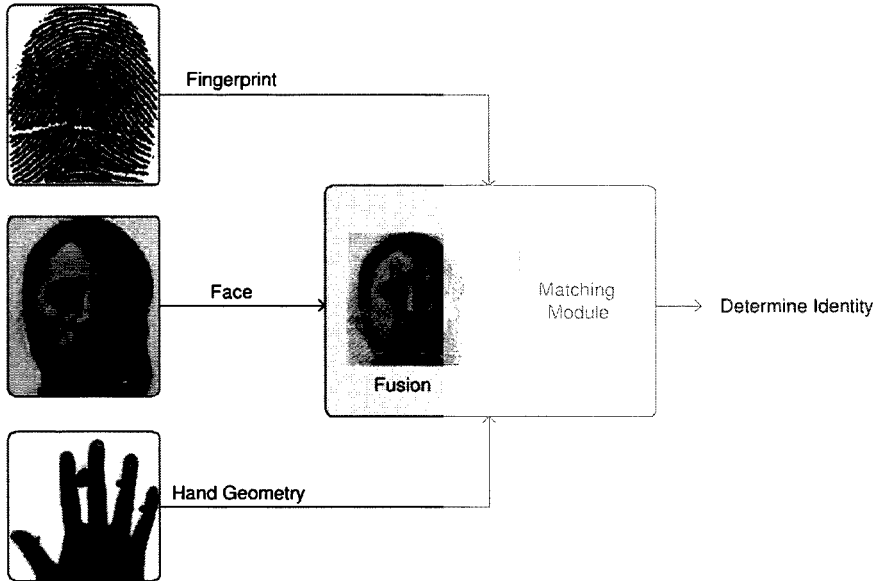
Figure 2.8. In the parallel mode of operation, the evidence acquired from multiple sources is simultaneously processed in order to establish identity. Note that the evidence pertaining to the multiple sources may be acquired in a sequential fashion.
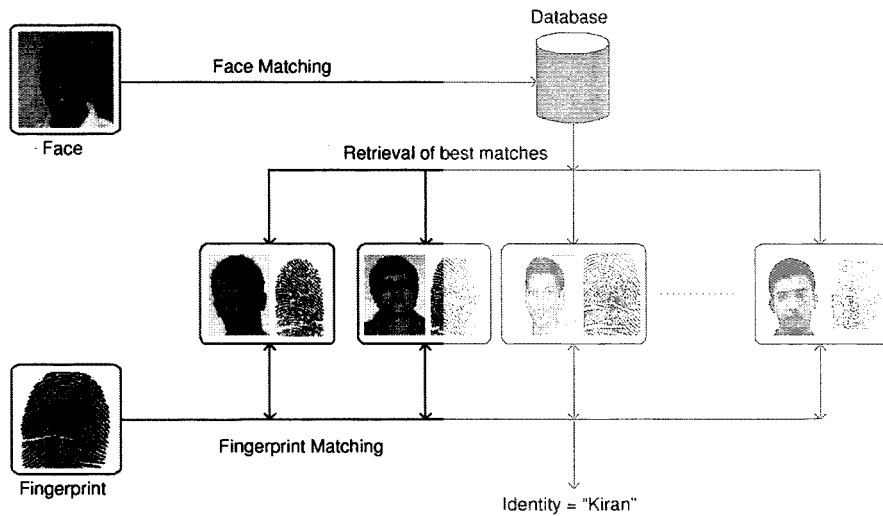

Figure 2.9. The cascade mode of processing permits database indexing where one modality can be used to retrieve a subset of identities while the second modality determines the best match. In this example, the face system is employed to recover the top n matches while the fingerprint system decides the identity of the user based on the n retrieved matches.

However, the design of a hierarchical multibiometric system has not received much attention from researchers.

2.6 Levels of fusion

In a typical pattern recognition system, the amount of information available to the system gets compressed as one proceeds from the sensor module to the decision module (see Figure 3.1). In a multibiometric system, fusion can be accomplished by utilizing the information available in any of these modules. Figure 2.10 indicates the various levels of fusion that are possible in the context of a biometric system. These levels can be broadly classified as (i) fusion prior to matching, and (ii) fusion after matching (Sanderson and Paliwal, 2002). This distinction is made because once the matcher (of a biometric system) is invoked, the amount of information available to the system drastically decreases. In this section we briefly introduce the various levels of fusion. In the next chapter, a more detailed description is provided.

2.6.1 Fusion prior to matching

Prior to matching, integration of information from multiple biometric sources can take place either at the sensor level or at the feature level. The raw data from the sensor(s) are combined in *sensor level fusion* (Iyengar et al., 1995). Sensor level fusion is applicable only if the multiple sources represent samples of the same biometric trait obtained either using a single sensor or different compatible sensors. For example, 2D face images of an individual obtained from several cameras can be combined to form a 3D model of the face. Another example of sensor level fusion is the mosaicing of multiple fingerprint impressions of a subject in order to construct a more elaborate fingerprint image (Jain and Ross, 2002a; Moon et al., 2004). In sensor level fusion, the multiple cues must be compatible and the correspondences between points in the raw data must be either known in advance or reliably estimated.

Feature level fusion refers to combining different feature sets extracted from multiple biometric sources. When the feature sets are homogeneous (e.g., multiple measurements of a person's hand geometry), a single resultant feature vector can be calculated as a weighted average of the individual feature vectors. When the feature sets are non-homogeneous (e.g., features of different biometric modalities like face and hand geometry), we can concatenate them to form a single feature vector. Feature selection schemes are employed to reduce the dimensionality of the ensuing feature vector (Ross and Govindarajan, 2005). Concatenation is not possible when the feature sets are incompatible (e.g., fingerprint minutiae and eigen-face coefficients).
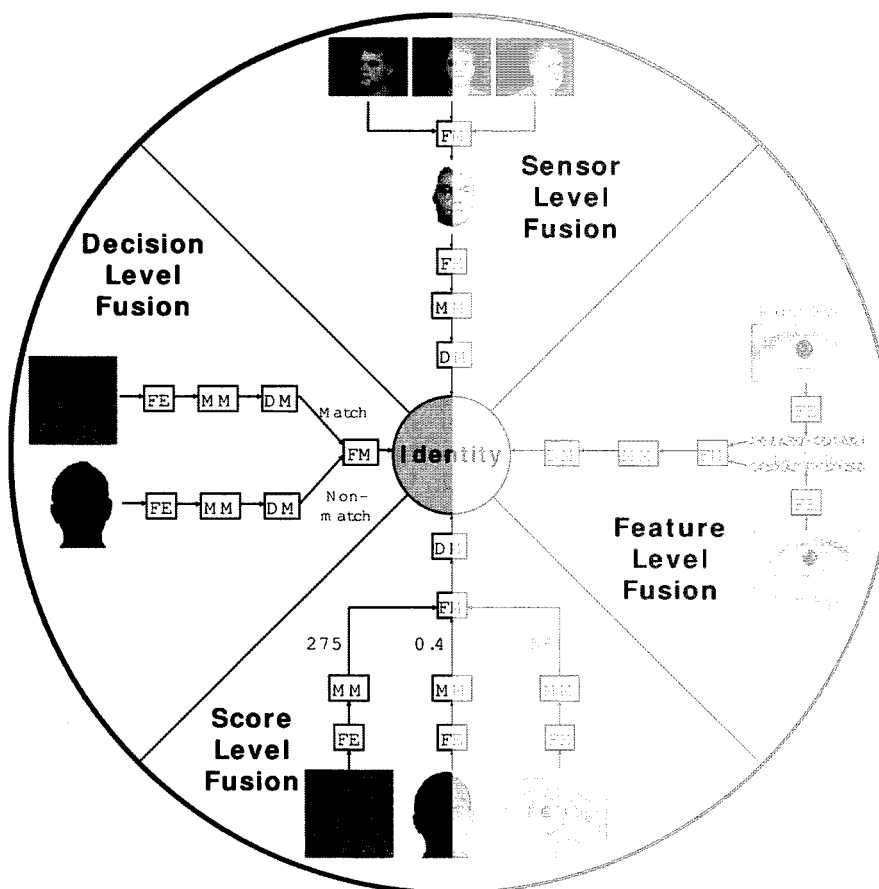
Figure 2.10. Fusion can be accomplished at various levels in a biometric system. Most multi-biometric systems fuse information at the match score level or the decision level. More recently researchers have begun to fuse information at the sensor and feature levels. In biometric systems operating in the identification mode, fusion can be done at the rank level (not shown here). FE: feature extraction module; MM: matching module; DM: decision-making module; FM: fusion module.

2.6.2   Fusion after matching

Schemes for integration of information after the classification/matcher stage can be divided into four categories: dynamic classifier selection, fusion at the decision level, fusion at the rank level and fusion at the match score level. A *dynamic classifier selection* scheme chooses the results of that biometric source which is most likely to give the correct decision for the specific input pattern (Woods et al., 1997). This is also known as the winner-take-all approach and

the module that performs this selection is known as an associative switch (Chen et al., 1997).

When each biometric system outputs a match score indicating the proximity of the input data to a template, integration can be done at the *match score level*. This is also known as fusion at the *measurement level* or *confidence level*. Next to the feature vectors, the match scores output by biometric matchers contain the richest information about the input pattern. Also, it is relatively easy to access and combine the scores generated by the different matchers. Consequently, integration of information at the match score level is the most common approach in multibiometric systems.

Integration of information at the *abstract* or *decision level* can take place when each biometric system independently makes a decision about the identity of the user (in an identification system) or determines if the claimed identity is true or not (in a verification system). Methods like majority voting (Lam and Suen, 1997), behavior knowledge space (Lam and Suen, 1995), weighted voting based on the Dempster-Shafer theory of evidence (Xu et al., 1992), AND/OR rules (Daugman, 2000), etc. can be used to consolidate the decisions rendered by individual systems. Since most commercial biometric systems provide access to only the final decision output by the system, fusion at the decision level is often the only viable option.

When the output of each biometric system is a subset of possible matches (i.e., identities) sorted in decreasing order of confidence, the fusion can be done at the *rank level*. This is relevant in an identification system where a rank may be assigned to the top matching identities. Ho et al., 1994 describe three methods to combine the ranks assigned by different matchers. In the highest rank method, each possible identity is assigned the best (minimum) of all ranks computed by different systems. Ties are broken randomly to arrive at a strict ranking order and the final decision is made based on the consolidated ranks. The Borda count method uses the sum of the ranks assigned by the individual systems to a particular identity in order to calculate the fused rank. The logistic regression method is a generalization of the Borda count method where a weighted sum of the individual ranks is used. The weights are determined using logistic regression.

2.7   Summary

Information and data fusion is an active research area spanning numerous fields and there are several applications that rely on effective evidence reconciliation schemes (Rao et al., 1996). In some applications, fusion may be viewed as a *problem* to be solved (e.g., robotics (Abidi and Gonzalez, 1992)) while in other applications, it may be viewed as a *solution* to a problem (e.g., forecasting (Clemen, 1989)). The role of multiple classifier systems in solving several pattern recognition problems has long been established (for an early example, see

Dasarathy and Sheela, 1979). Multiple classifier systems exploit the complementary strengths of participating experts (viz., classifiers) in order to enhance the performance of a pattern recognition application. In the context of multibiometrics, these experts represent different biometric sources (e.g., multiple biometric sensors, multiple traits, etc.) providing information at multiple levels (e.g., score-level, decision-level, etc.).

The design of a multibiometric system is governed by several different factors including the sources of information to be used, the acquisition and processing sequence to be adopted, the type of information to be combined and the fusion strategy to be employed. The development of robust human computer interfaces (HCIs) is necessary to permit the efficient acquisition of multibiometric data from individuals (see Sharma et al., 1998 and the references therein). A HCI that is easy to use can result in rapid user habituation and promote the acquisition of high quality biometric data. Indeed, the *user* is one of the key components in any biometric system and it is necessary that system designers take into account user-centric issues of the target population (such as age, gender and cultural considerations) whilst designing the HCI (Ashbourn, 2003). Acquiring and processing multibiometric information in a sequential fashion (i.e., cascaded logic) helps curtail the time required for generating a decision. The use of multiple modalities in the cascaded mode facilitates database indexing, where one modality can be used to narrow down the number of possible identities before invoking the next.

Information fusion in biometrics presents an elegant way to enhance the matching accuracy of a biometric system without resorting to non-biometric alternatives. Determining the sources of biometric information that would result in the best matching performance is not an easy task. Chang et al., 2005 describe a multibiometric system that utilizes the 2D and 3D face images of a user for recognition. In their experiments involving 198 subjects, they observe that multi-sensor fusion of 2D and 3D images results in better recognition performance compared to multi-sample fusion of 2D images alone (fusion was accomplished at the match score level in both cases). However, they state that increasing the number of 2D images in multi-sample fusion *may* result in the same recognition performance as multi-sensor fusion. Furthermore, employing alternate fusion strategies at other levels (besides the match score level) can lead to different conclusions. In view of this, it is difficult to predict the optimal sources of biometric information relevant for a particular application based on recognition performance alone. Factors such as cost, throughput time, user convenience, scalability, etc. play a large role in selecting the sources of biometric information and adopting a particular fusion strategy.