TECHNICAL REPORT

ISO/IEC TR 24722

First edition
2007-07-01

Information technology — Biometrics — Multimodal and other multibiometric fusion

Technologies de l'information — Biométrie — Fusion multimodale et autre fusion multibiométrique

Reference number
ISO/IEC TR 24722:2007(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

1 Scope 1

2 Terminology issues 1

3 Terms and definitions 3

4 Overview of multimodal and other multibiometric systems 5
4.1 General 5
4.2 Simultaneous and sequential presentation 7
4.2.1 General multibiometric system model 7
4.2.2 Simultaneous presentation 8
4.2.3 Sequential presentation 8
4.3 Correlation 8

5 Levels of combination 9
5.1 Overview 9
5.2 Decision-level fusion 12
5.2.1 Simple decision-level fusion 12
5.2.2 Advanced decision-level fusion 13
5.2.2.1 General model 13
5.2.2.2 Layered system 15
5.2.2.3 Cascaded system 15
5.3 Score-level fusion 15
5.3.1 Overview 15
5.3.2 Score normalisation 16
5.3.3 Score fusion methods 19
5.4 Feature-level fusion 21

6 Characterisation data for multibiometric systems 21
6.1 Overview 21
6.2 Use of characterisation data in normalisation and fusion 21

7 Scope and options for standardisation 22
7.1 Introduction 22
7.2 Implementation areas 22
7.3 Interoperability requirements 23
7.4 Possible standardisation activity 23
7.4.1 On record format standardisation 23
7.4.2 On framework standardisation 25
7.4.3 On application profile standardisation 27
7.4.4 On compliance standardisation 27
7.4.5 On multimodal testing standardisation 28
7.5 Summary 28

Bibliography 29

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard (“state of the art”, for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 24722, which is a Technical Report of type 2, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

Introduction

Some applications of biometrics require a level of technical performance that is difficult to obtain with a single biometric measure. Such applications include prevention of multiple applications for national identity cards and security checks for air travel. In addition, provision is needed for people who are unable to give a reliable biometric sample for some biometric modalities.

Use of multiple biometric measurements from substantially independent biometric sensors, algorithms or modalities typically gives improved technical performance and reduces risk. This includes an improved level of performance where not all biometric measurements are available such that decisions can be made from any number of biometric measurements within an overall policy on accept/reject thresholds.

Of the various forms of multibiometric systems, the potential for multimodal biometric systems, each using an independent measure, has been discussed in the technical literature since at least 1974 [22, 49]. Advanced methods for combining measures at the score level have been discussed [15, 16]. At the current level of understanding, combining results at the score level typically requires knowledge of both genuine and impostor distributions. All of these measures are highly application-dependent and generally unknown in any real system. Research on the methods not requiring previous knowledge of the score distributions is continuing and research on fusion at both the image and feature levels is still progressing.

Given the current state of research into those questions and the highly application-dependent and generally unavailable data required for proper fusion at the score level, work on multimodal and other multibiometric fusion was considered not sufficiently mature to initiate an International Standard on the subject. Instead, it was considered appropriate to publish a Technical Report on the subject. This Technical Report is meant to provide information for future development of standards on multibiometric systems, in particular regarding the various aspects of fusion. It will also provide a reference on multibiometric fusion for developers of other biometric standards and implementers.

Information technology — Biometrics — Multimodal and other multibiometric fusion

1 Scope

This Technical Report contains descriptions of and analyses of current practices on multimodal and other multibiometric fusion, including (as appropriate) references to more detailed descriptions. It also discusses the need for, and possible routes to, standardisation to support multibiometric systems.

This Technical Report contains descriptions and explanations of high-level multibiometric concepts to aid in the explanation of multibiometric fusion approaches including multimodal, multiinstance, multisensorial, multialgorithmic, decision-level and score-level logic.

2 Terminology issues

The primary motivation in addressing the terms and definitions in Clause 3 is to draw a distinction between "multibiometric" and "multimodal" terms that appeared to be used in the literature interchangeably. To support defining this terminology, the term "modality" is a key, and Table 1 provides a listing of modalities based on CBEFF [30]. The distinction between conventional and unconventional categories is subjective, and based on past and current biometric products.

Table 1 — Terms for biometric modalities or data types

Category	Biometric Type
Other	No Value Available
Multiple	Multiple Biometric Types
Conventional	Face
	Voice
	Finger
	Iris
	Retina
	Hand Geometry
	Signature or Sign
Unconventional	Keystroke
	Lip Movement
	Gait
	Vein
	DNA
	Ear
	Foot
	Scent

(Source: ISO/IEC 19785-1: 2006, *Information technology — Common Biometric Exchange Formats Framework — Part 1: Data element specification,* Table 1 — Abstract values for BDB_biometric_type.)

3 Terms and definitions

Note: Two categories of terms are defined here:

— terms that are specific to multimodal and multibiometric systems;

— terms that are not specific to multimodal and multibiometric systems, but are required to define the terms in the first category and not defined in the latest revision of ISO/IEC JTC 1/SC 37 Standing Document 2 [33].

For definitions of other terms in the subject field of biometrics, refer to ISO/IEC JTC 1/SC 37 Standing Document 2 [33].

For the purposes of this document, the following terms and definitions apply.

3.1
biometric characteristic
biometric (deprecated)
biological and behavioural characteristic of an individual that can be detected and from which distinguishing, repeatable biometric features can be extracted for the purpose of automated recognition of individuals

Note 1: Biological and behavioural characteristics are physical properties of body parts, physiological and behavioural processes created by the body and combinations of any of these.

Note 2: Distinguishing does not necessarily imply individualization.

Examples: Galton ridge structure, face topography, facial skin texture, hand topography, finger topography, iris structure, vein structure of the hand, ridge structure of the palm, and retinal pattern.

3.2
biometric modality
the biometric characteristic which is used in a biometric process

3.3
biometric process
automated process using one or more biometric characteristics of a single individual for the purpose of enrollment, verification or identification

3.4
biometric fusion
combination of information from multiple sources, i.e. sensors, modalities, algorithms, instances or presentations

3.5
cascaded system
system where pass/fail thresholds of biometric samples are used to determine if additional biometric samples are required to reach an overall system decision

3.6
layered system
system where individual biometric scores are used to determine the pass/fail thresholds of other biometric data processing

3.7
multialgorithmic
using multiple algorithms for processing the same biometric sample

3.8
multibiometric
pertaining to multibiometrics

Note: Multibiometric has five distinct subcategories: multimodal, multiinstance, multisensorial, multialgorithmic and multipresentation

3.9
multibiometric process
biometric process involving the use of biometric fusion

3.10
multibiometrics
automated recognition of individuals based on their biological or behavioral characteristics and involving the use of biometric fusion

3.11
multiinstance
using multiple biometric instances within one biometric modality

Examples: Iris (left) + Iris (right), Fingerprint (left index) + Fingerprint (right index).

3.12
multimodal
using multiple different biometric modalities

Example: Fingerprint + Face.

3.13
multipresentation
using either multiple presentation samples of one instance of a biometric characteristic or a single presentation that results in the capture of multiple samples

Example: Several frames from video camera capture of a face image (possibly but not necessarily consecutive).

Note: Multipresentation biometrics is considered a form of multibiometrics, if fusion techniques are employed. Many fusion and normalisation techniques are appropriate to the integration of information from multiple presentations of the same biometric instance.

3.14
multisensorial
using multiple sensors for capturing samples of one biometric instance

Examples: For face: infrared spectrum, visible spectrum, 2-D image and 3-D image. For fingerprint: optical, electrostatic and acoustic sensors.

3.15
sequential presentation
capturing biometric samples in separate capture events to be used for biometric fusion

3.16
simultaneous presentation
capturing biometric samples in a single capture event to be used for biometric fusion

4 Overview of multimodal and other multibiometric systems

4.1 General

In general, the use of the terms multimodal or multibiometric indicates the presence and use of more than one modality, sensor, instance, and/or algorithm in some form of combined use for making a specific biometric identification or verification decision. The methods of combining multiple samples, matching scores or matching decisions can be very simple or mathematically complex. For the purpose of this document, any method of combination will be considered a form of "fusion". Combination techniques will be covered in Clause 5 of this document.

Multimodal biometrics were first proposed, implemented and tested in the 1970s. Combining measures was seen as a necessary future requirement for biometric systems. It was widely thought that combining multiple measures could increase either security by decreasing the false acceptance rate or user convenience by decreasing the false rejection rate. These systems did not seem to advance into practical applications.

The use of fusion and related methods has been a key tool in the successful implementation of large scale automated fingerprint identification systems (AFISs), starting in the 1980s. Until recently, multiple modalities have not been used in AFIS; however, most methods of fusion discussed elsewhere in this report have been successfully implemented using fingerprints alone. Some of the ways that fusion has been implemented in AFISs include:

— Image (AKA sample) fusion in creating a single "rolled" image from a series of plain impressions on a livescan device;

— Template fusion in the use of multiple feature extraction algorithms on each fingerprint image;

— Multiinstance fusion in the use of fingerprints from all ten fingers;

— Multipresentation fusion in the use of rolled and slap (plain) fingerprints;

— Algorithm fusion for the purpose of efficiency (cost, computational complexity, and throughput rate); generally matchers are used as a series of filters in order of increasing computational complexity. These are generally implemented as a mix of decision and score-level fusion;

— Algorithm fusion for the purpose of accuracy (decreasing false accept rate and/or false reject rate, lessening sensitivity to poor-quality data); matchers are used in parallel, with fusion of resulting scores.

The use of fusion has made AFISs possible, because of fusion's increase in both accuracy and efficiency.

Most work to date on multibiometrics has focused only on improving false acceptance and false rejection error rates. Work at University of Kent, on project IAMBIC (Intelligent Agents for Multimodal Biometric Identification and Control) is notable as it considers the use of multibiometrics to flexibly improve usability, security or accuracy [65].

To further the understanding of the distinction among the multibiometric categories, Table 2 illustrates the basic distinctions among categories of multibiometric implementation. The key aspect of the category that makes it multi-"something" is shown in boldface.

Table 2 — Multibiometric categories illustrated by the simplest case of using 2 of something

Category	Modality	Algorithm	Biometric characteristic (e.g., body part)	Sensor
Multimodal	**2 (always)**	2 (always)	2 (always)	2 (usually)[b]
Multialgorithmic	1 (always)	**2 (always)**	1 (always)	1 (always)
Multiinstance	1 (always)	1 (always)	**2 instances of 1 characteristic (always)**	1 (usually)[c]
Multisensorial	1 (always)	1 (usually)[a]	1 (always, and same instance)	**2 (always)**
Multipresentation	1	1	1	1

a It is possible that two samples from separate sensors could be processed by separate "feature extraction" algorithms, and then through a common comparison algorithm, making this "1.5 algorithms", or two completely different algorithms.

b Exception: a multimodal system with a single sensor used to capture two different modalities. For example a high resolution image used to extract face and iris or face and skin texture.

c Exception may be the use of two individual sensors to each capture one instance, for example possibly a two-finger fingerprint sensor.

Multimodal biometric systems take input from single or multiple sensors that capture two or more different modalities of biometric characteristics. For example, a single system combining face and iris information for biometric recognition would be considered a "multimodal" system regardless of whether face and iris images were captured by different imaging devices or the same device. It is not required that the various measures be mathematically combined in anyway. For example, a system with fingerprint and voice recognition would be considered "multimodal" even if the "OR" rule was being applied, allowing users to be verified using either of the modalities.

Multialgorithmic biometric systems receive a single sample from a single sensor and process that sample with two or more algorithms. This technique could be applied to any modality. Maximum benefit (theoretically) would be derived from algorithms that are based on distinctly different and independent principles (such algorithms may be called "orthogonal").

Multiinstance biometric systems use one (or possibly multiple) sensor(s) to capture samples of two or more different instances of the same biometric characteristic. For example, systems capturing images from multiple fingers are considered to be multiinstance rather than multimodal. However, systems capturing, for example, sequential frames of facial or iris images are considered to be multipresentation rather than multiinstance.

Multisensorial biometric systems sample the same instance of a biometric characteristic with two or more distinctly different sensors. Processing of the multiple samples can be done with one algorithm, or some combination of multiple algorithms. For example, a face recognition application could use both a visible light camera and an infrared camera coupled with a specific frequency (or several frequencies) of infrared illumination.

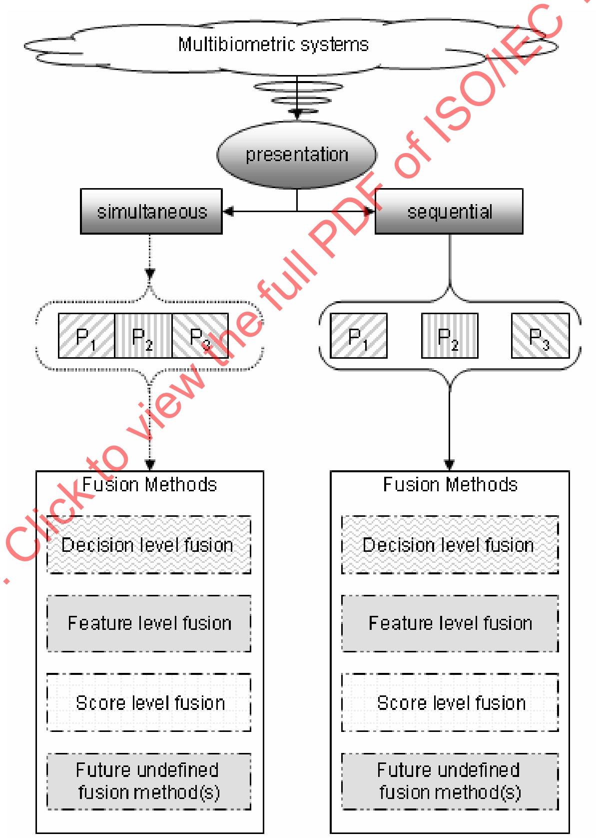
For a specific application in an operational environment, there are numerous system design considerations, and trade-offs that must be made, among factors such as improved performance (e.g., identification or verification accuracy, system speed and throughput, robustness, and resource requirements), acceptability, circumvention, ease of use, operational cost, environmental flexibility, and population flexibility [44].

Especially for a large-scale human identification system, there are additional system design considerations such as operation and maintenance, reliability, system acquisition cost, life cycle cost, and planned system response to identified susceptible means of attack, all of which will affect the overall deployability of the system [44].

4.2 Simultaneous and sequential presentation

4.2.1 General multibiometric system model

A general multibiometric system model is shown in Figure 1. For explanatory purposes, this model uses three biometric samples (P_1, P_2, P_3) from 3 unique biometric modalities, except for where specified differently. At the topmost level a subject presents their biometric characteristic(s) to the system. Dependent upon the system design, there are two methods of presenting characteristics for acquisition by the system: 1) **simultaneous** and 2) **sequential**.



Note: the presentation method (simultaneous or sequential) is distinct from the fusion process itself. The purpose of including this information is to illustrate considerations that may influence multibiometric system design.

Figure 1 — Multibiometric system model

4.2.2 Simultaneous presentation

Simultaneous presentation (with successful capture) provides biometric sample(s) from multiple modalities in a single event (e.g., a face and iris taken from the same camera). System designs that utilize simultaneous acquisition would tend toward high throughput applications at the expense of possible added complexity (to synchronize sample collection) or difficulty of use (dual sensor interaction, user multi-tasking).

4.2.3 Sequential presentation

Sequential capture acquires biometric sample(s) from one or multiple modalities in separate events. Sequential capture may be utilized in three concepts discussed in the literature. The first is multiinstance, which is the use of two or more instances within one modality for a subject, i.e. Fingerprint (left index) + Fingerprint (right index). In this example, one single digit fingerprint reader is used twice in sequence. The second concept is multimodal, which is the use of multiple different biometric modalities captured from one or more sensors for a subject, i.e. Hand + Face in sequence. The third concept is multisensorial, which is the use of two or more distinct sensors for capturing the same biometric feature(s) for a subject, but not at the same time. To avoid confusion with multimodal, which may also capture biometric feature(s) from two or more distinct sensors, multisensorial can be clarified as “unimodal multisensorial”. Examples for face recognition are: infrared spectrum, visible spectrum, 2-D image, and 3-D image; for fingerprint recognition: optical, electrostatic and acoustic sensors.

4.3 Correlation

In multimodal biometric systems the information being fused may be correlated at several different levels [57] as illustrated in the following examples.

— Correlation between modalities: This refers to biometric samples that are *physically related* such as the speech and lip movement of a user.

— Correlation due to identical biometric samples: This is the case in multialgorithmic systems where the *same* biometric sample (e.g., a fingerprint image) or sub-sets of the biometric sample (e.g., voice, where an entire sample may be used by one algorithm and part of the sample by another) is subjected to different feature extraction and matching algorithms (e.g., a minutiae-based matcher and a texture-based matcher).

— Correlation between feature values: A subset of feature values constituting the feature vectors of different modalities may be correlated. For example, the area of a user’s palm (hand geometry) may be correlated with the width of the face.

— Correlation among instances due to common operating procedures (e.g., common capture device and operator training).

— Correlation among instances due to subject behaviour (e.g., coloured contact lenses on both eyes).

However, in order to determine the *extent* of correlation it is necessary to examine the match scores (or the ACCEPT/REJECT decision) pertaining to the matchers involved in the fusion scheme. In the multiple classifier system literature, it has been demonstrated that fusing uncorrelated classifiers leads to a significant improvement in matching performance [57].

For two classifiers of reasonable accuracy involved in a fusion scheme, score outputs from inputs that come from the same subject may, but need not, be correlated. Therefore it is more appropriate to consider the correlation of classifier errors as described by Goebel, Yan, and Cheetham [20]. The correlation ρ_{nc} is given by:

$$\rho_{n_c} = \frac{nN_c^f}{N - N_c^t - N_c^f + nN_c^f}$$

Where *n* is the number of classifiers under test, *N* is the total number of sequences, N_C^f is the number of sequences where all classifiers have an incorrect output at threshold *C*, and N_C^t is the number of sequences where all classifiers have a correct output at a threshold *C*. (NOTE: This expression is relevant for computing the correlation of errors at the *decision* level.)

5 Levels of combination

5.1 Overview

As a basis for the definition of levels of combination in multibiometric systems, we first introduce the single-biometric process and its building blocks, using the example of an authentication system. Figure 2 shows the block diagram of a single-biometric process.

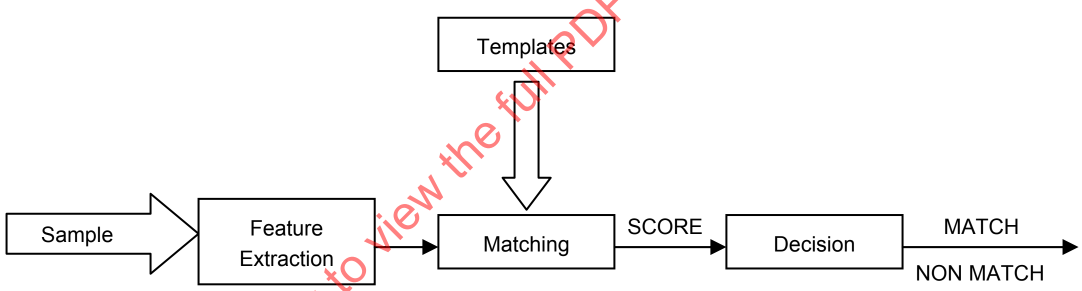


Figure 2 — Single biometric process (generic)

A biometric Sample captured by a biometric sensor (e.g., a fingerprint image) is fed into the Feature Extraction module. Using signal processing methods, the feature extraction module converts a sample into Features (e.g., fingerprint minutiae), which form a representation apt for matching. Usually, multiple features are collected into a feature vector. The Matching module takes the feature vector as input and compares it to a stored Template. The result is a match Score, which is used by the Decision module to decide (e.g., by applying a threshold) whether the presented sample matches with the stored template. The outcome of this decision is a binary *match* or *non-match*.

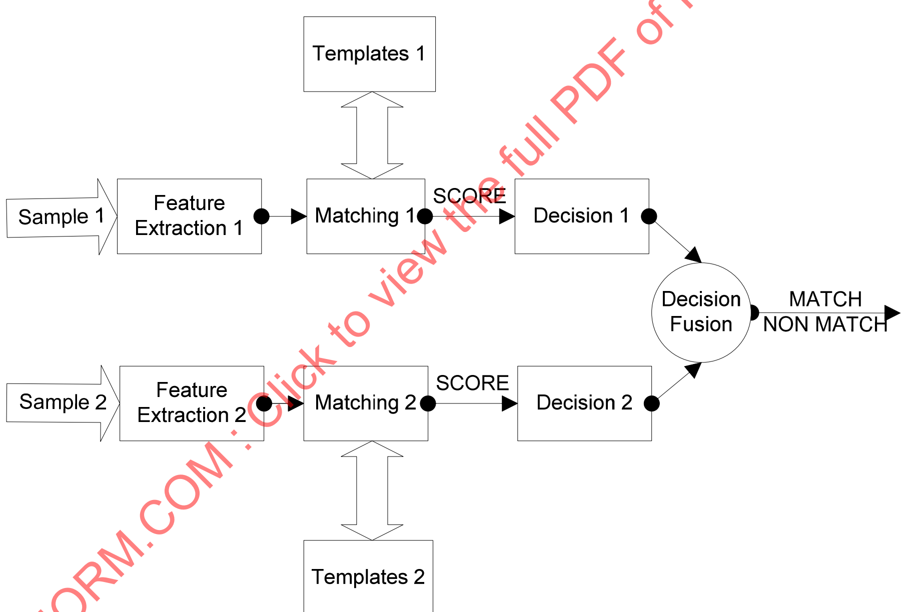
Generalizing the above process to multiple biometrics, there are several levels at which fusion can take place. These include consolidating information at the (i) decision level, (ii) match score level, (iii) feature level, and (iv) sample level. Note that fusion at levels (i) and (ii) occur after the matching module is invoked, while levels (iii) and (iv) occur before the matcher. Although integration is possible at these different levels, fusion at the feature set level, the match score level and the decision level are the most commonly used. Figure 3 illustrates the following different levels of fusion for the case of a multimodal system [7, 45].

a) **Decision level:** Each individual biometric process outputs its own Boolean result. The fusion process fuses them together by a combination algorithm such as AND and OR, possibly taking further parameters such as sample quality scores as input.

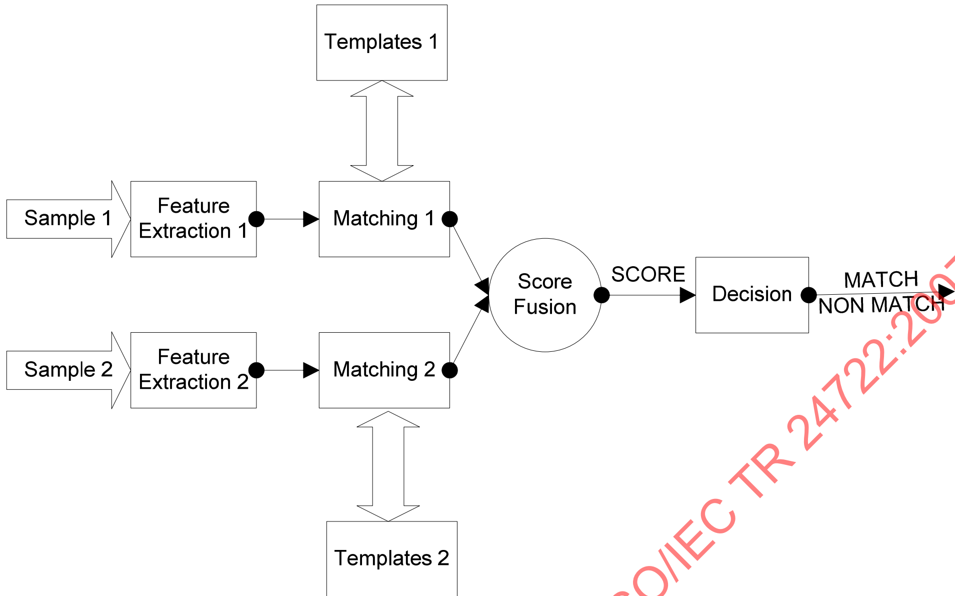
b) **Score level:** Each individual biometric process typically outputs a single match score but possibly multiple scores. The fusion process fuses these into a single score or decision, which is then compared to the system acceptance threshold.

c) **Feature level:** Each individual biometric process outputs a collection of features. The fusion process fuses these collections of features into a single feature set or vector.

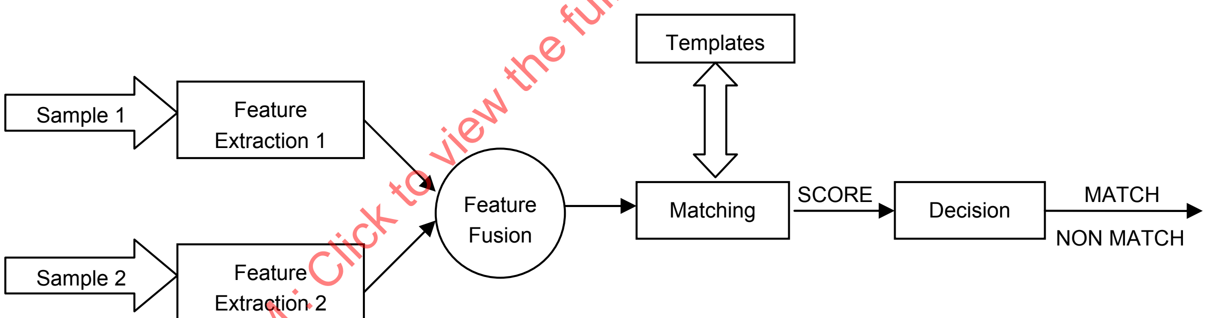
d) **Sample level:** Each individual biometric process outputs a collection of samples. The fusion process fuses these collections of samples into a single sample.



a) Decision-level fusion

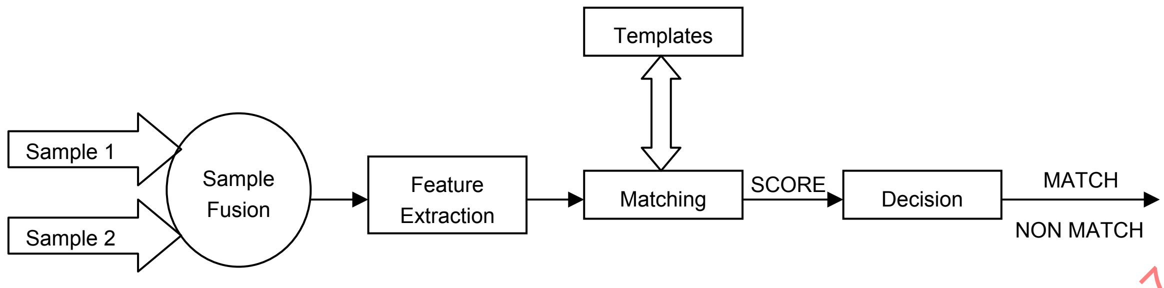


b) Score-level fusion



(Note: Sample 1 and Sample 2 may be the same sample.)

c) Feature-level fusion



d) Sample-level fusion

Figure 3 — Levels of fusion for a multimodal system

For simultaneous or sequential biometric sample acquisition, features are extracted, and are compared against the template. P_1, P_2, and P_3 from Figure 1 refer to the match score from the comparison against the template. How the match scores are determined is system dependent and outside the scope of this technical report. The match scores of P_1, P_2, and P_3 are then sent to the fusion module for a final result. In multibiometric systems the fusion may occur at the decision or score level.

5.2 Decision-level fusion

5.2.1 Simple decision-level fusion

Decision-level fusion occurs after a match decision has been made for each biometric component. It is based on the binary result values *match* and *non-match* output by the decision modules (see Figure 3 a), Decision-level fusion).

For biometric systems composed of a small number of components, it is convenient to assign logical values to match outcomes so that fusion rules can be formulated as logical functions. The behaviour of the two most widely used functions, AND and OR, are listed in Table 3, assuming a pair of decision-level outputs.

Table 3 — AND & OR fusion of decisions for a case of two biometric modalities

Decision Biometrics 1	Decision Biometrics 2		AND-fused decision	OR-fused decision
X	X		X	X
X	●		X	●
●	X		X	●
●	●		●	●

X*: Non-Match, •: Match*

For biometric systems using many components, voting schemes have been established as fusion rules, the most common of which is majority voting rule. The AND and OR are specific examples of voting schemes.

5.2.2 Advanced decision-level fusion

5.2.2.1 General model

Decision-level fusion is based upon individual accept/reject decisions for each sample. The two sub groups of advanced decision-level fusion are 1) **layered** and 2) **cascaded**. A layered system uses individual biometric scores to determine the pass/fail thresholds for other biometric data processing. Cascaded systems use pass/fail thresholds of modality-specific biometric samples to determine if additional biometric samples from other modalities are required to reach an overall system decision. Decision-level fusion for the two subgroups are shown in Figure 4.

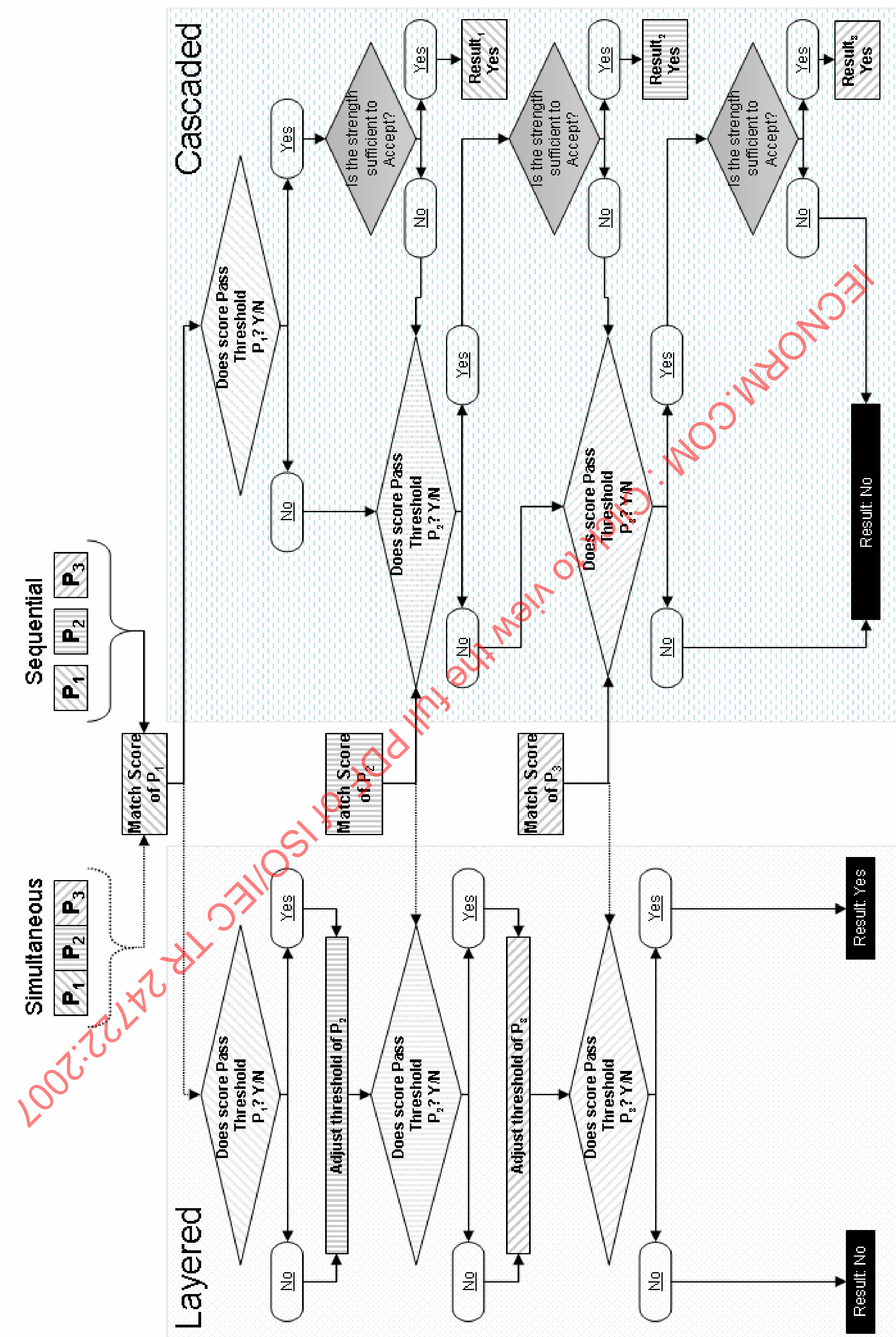


Figure 4 — Advanced decision-level fusion

5.2.2.2 Layered system

Independent of whether the presentation was simultaneous or sequential, the match score of P_1 enters the layered system. The system processes the score against the system defined threshold. If it passes the criteria/threshold for modality P_1 the output would adjust (raise or lower) the threshold needed to pass for modality P_2. If P_1 fails to meet the criteria/threshold for modality P_1 then the output most likely would increase the threshold required for modality P_2. Upon completion of processing P_1 and resetting the thresholds requirements for modality P_2, the match score of P_2 enters the system. The process iterates as discussed above for P_2 and P_3. Once the modality P_3 process is completed, a final accept/reject decision is made.

5.2.2.3 Cascaded system

Independent of simultaneous or sequential presentation, cascaded systems rely on at least one biometric sample. If the first sample does not meet the requirements, additional samples are matched. Using Figure 4 as the model for this discussion, match score P_1 enters the system and is matched against the threshold for sample P_1. If the score exceeds the criteria/threshold required for P_1 a subsequent decision is made on the strength of the result (which could also include sample quality measures). If this strength is sufficient, the subject is accepted. If the score of P_1 fails the initial threshold test or passes the initial threshold test, but fails the strength decision, cascaded systems require the use of the score of P_2. This process is repeated for scores P_2 and P_3. Note that cascaded systems may not require P_2 or P_3 to be captured if P_1 passes the threshold and strength test.

5.3 Score-level fusion

5.3.1 Overview

In score-level fusion, each system provides matching scores indicating the proximity of the feature vector with the template vector. These scores can then be combined to improve the matching performance.

From a theoretical point of view, biometric processes can be combined reliably to give a guaranteed improvement in matching performance. Any number of suitably characterized biometric processes can have their matching scores combined in such a way that the multibiometric combination is guaranteed (on average) to be no worse than the best of the individual biometric devices. The key is to identify correctly the method which will combine these matching scores reliably and maximize the improvement in matching performance.

The mechanism (for this sort of good combination of scores within a multibiometric system) must follow at least two guidelines. Firstly, each biometric process must produce a score, rather than a hard accept/reject decision, and make it available to the multibiometric combiner. Secondly, in advance of operational use, each biometric process must make available to the multibiometric combiner, its technical performance (such as score distributions) in the appropriate form (and with sufficient accuracy of characterisation).

Both verification (1:1) and identification (1:N) systems can support fusion at the match score level. However, identification systems can also integrate information available at the rank level (which is a form of score level with multiple scores or indices based on scores). In identification systems a template from a biometric sample is compared against templates from a subset of identities present in the database and, therefore, a sequence of ordered match scores pertaining to these identities is available. Ho et al. [23] describe three methods to combine the ranks assigned by the different matchers. In the *highest rank method*, each possible match is assigned the highest (minimum) rank as computed by different matchers. Ties are broken randomly to arrive at a strict ranking order and the final decision is made based on the combined ranks. The *Borda count* method uses the sum of the ranks assigned by the individual matchers to calculate the combined ranks. The *logistic regression* method is a generalization of the Borda count method where the weighted sum of the individual ranks is calculated and the weights are determined by logistic regression.

5.3.2 Score normalisation

Score normalisation methods attempt to map the scores of each biometric process to a common domain. Some approaches are based on the Neyman-Pearson lemma, with simplifying assumptions. For example, mapping scores to likelihood ratios allows them to be combined by multiplying under an independence assumption. Other approaches may be based on modifying other statistical measures of the match score distributions.

The parameters used for normalisation can be determined using a fixed training set or adaptively based on the current feature vector. (Note: The computed characteristic may represent only "estimates" of the underlying population characteristics.) Score normalisation is closely related to score-level fusion since it affects how scores are combined and interpreted in terms of biometric performance. As in [36]:

a) The matching scores at the output of the individual matchers *may not be homogeneous*. For example, one matcher may output a distance (dissimilarity) measure while another may output a proximity (similarity) measure.

b) Further, the outputs of the individual matchers *need not be on the same numerical scale* (range).

c) Finally, the matching scores at the output of the matchers *may follow different statistical distributions*.

Due to these reasons, scores are generally normalized prior to fusion into a common domain. Figure 5 depicts a score-level fusion framework for processing two biometric samples, taking normalisation into account.

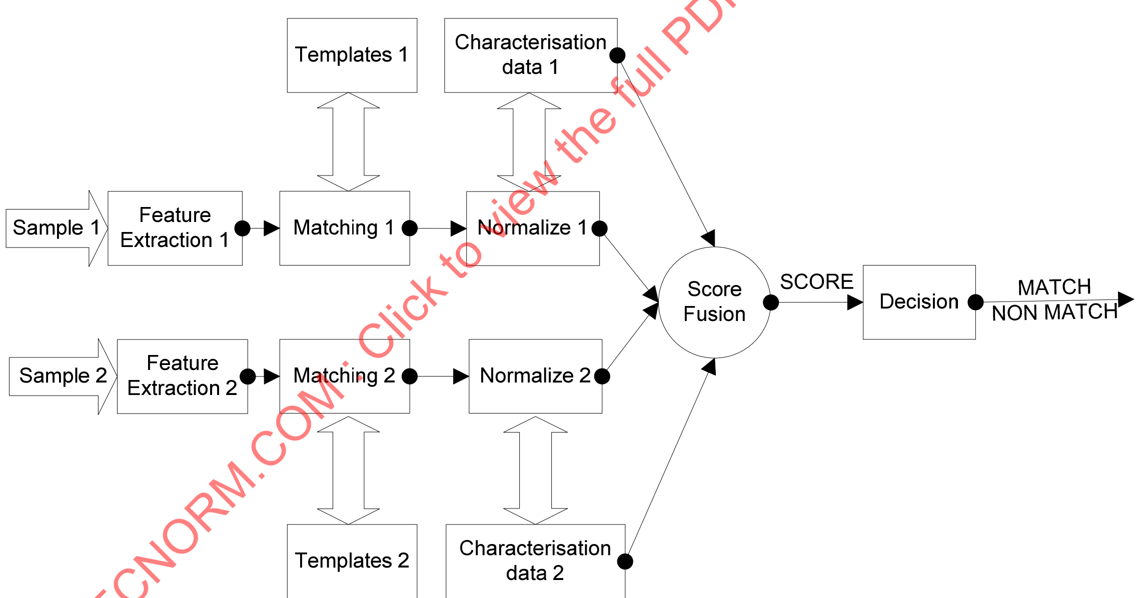


Figure 5 — A framework for score-level fusion

Table 5 lists, under the framework of Figure 5, several commonly used score normalisation methods. Note that some fusion methods use probability density functions (PDFs) directly and do not require normalisation methods. Table 4 defines the symbols used in Table 5. In some cases, PDFs are used to convert raw/native scores directly into Probability of False Accept and thus to a decision without need to have native scores brought to a common reference range using normalization.

Table 4 — Symbols used for score normalisation formulas

Statistical measures	Characterisation data		
	Genuine distribution	Impostor distribution	Both genuine and impostor distributions
Minimum score	S^{G}_{Min}	S^{I}_{Min}	S^{B}_{Min}
Maximum score	S^{G}_{Max}	S^{I}_{Max}	S^{B}_{Max}
Mean score	S^{G}_{Mean}	S^{I}_{Mean}	S^{B}_{Mean}
Median score	S^{G}_{Med}	S^{I}_{Med}	S^{B}_{Med}
Score standard deviation	S^{G}_{SD}	S^{I}_{SD}	S^{B}_{SD}
Constant	C	C	C
Probability density function	PDF^{G}	PDF^{I}	N.A.
Centre of PDF crossover	S_{center}		
Width of PDF crossover	S_{width}		
NOTE S – represents Similarity score; Subscript G stands for Genuine; Subscript I stands for Impostor ; Subscript B stands for Both.			

Table 5 — Examples of score normalisation methods

Method	Formula	Data elements	Comment
Min-max (MM)	$S' = (S - S^B_{Min})/(S^B_{Max} - S^B_{Min})$	S^B_{Min} S^B_{Max}	• Uses empirical data (or theoretical limit or vendor provided) • No accounting for non-linearity
Z-score	$S' = (S - S^I_{Mean}) / S^I_{SD}$	S^I_{Mean} S^I_{SD}	• Assumes normal distribution • Symmetric about mean • Assumes stability of both distributions across populations
Median absolute deviation (MAD)	$S' = (S - S^B_{Med}) / (C \cdot \text{median}\,\lvert S - S^B_{Med}\rvert)$	S^B_{Med} C	• Assumes stability of both distributions across populations
Hyperbolic tangent (Tanh)	$S' = 0.5(\tanh(C(S - S^G_{Mean})/S^G_{SD})+1)$	S^G_{Mean} S^G_{SD}	• Mean and variance of transformed data distribution • Assumes stability of both distributions across populations
Adaptive (AD) [a] a) Two-quadrics (QQ) b) Logistic c) Quadric-line-quadric (QLQ)	$n_{AD} = \begin{cases} \frac{1}{c} n_{MM}^2, & n_{MM} \le c \\ c + \sqrt{(1-c)(n_{MM} - c)}, & \text{otherwise} \end{cases}$ $n_{AD} = \frac{1}{1 + A \cdot e^{-B \cdot n_{MM}}}$ $n_{AD} = \begin{cases} \frac{1}{(c-\frac{w}{2})} n_{MM}^2, & n_{MM} \le (c - \frac{w}{2}) \\ n_{MM}, & (c-\frac{w}{2}) < n_{MM} \le (c+\frac{w}{2}) \\ (c+\frac{w}{2}) + \sqrt{(1-c-\frac{w}{2})(n_{MM} - c - \frac{w}{2})}, & \text{otherwise.} \end{cases}$	c w Δ $A = \frac{1}{\Delta} - 1$ $B = \frac{\ln A}{c}$	• Assumes non-linearity • 3 modeling methods • Assumes stability of both distributions across populations • n_{AD} = adaptive normalisation score; n_{MM} = min-max normalized score; c = center of overlap of genuine and impostor score distributions; w = width of the overlap; Δ = a small value (0.01 in [63])
Biometric Gain against Impostors (BGI)	$P_{Si\|I} / P_{Si\|G}$, $P_{Si\|G}$ = Value of PDF^G at score S_i $P_{Si\|I}$ = Value of PDF^I at score S_i	PDF^G PDF^I	• Assumes stability of both distributions across populations
BioAPI	$S' = FAR_{(threshold\ =\ score)}$	PDF^I	• Assumes stability of imposter distribution
Borda count	N - Rank(S) (Where *N* is the number of alternatives).	Rank	• Applicable only to 1:N matching

NOTE This table lists two types of normalisation schemes: (i) schemes that modify the location and scale parameters of the score distribution; and (ii) schemes that consider only the overlap region of the genuine and impostor scores. Thus, the min-max, z-score, MAD and tanh techniques fall under category (i), while QQ and QLQ fall under category (ii). Typically, category (ii) techniques are used *after* having applied one of the category (i) schemes.

[a] Refer to document [63] in Bibliography.

5.3.3 Score fusion methods

When individual biometric matchers output a set of possible matches along with the quality of each match (match score), integration can be done at the match score level. This is also known as fusion at the measurement level or confidence level. The match score output by a matcher contains the richest information about the input biometric sample in the absence of feature-level or sensor-level information. Furthermore, it is relatively easy to access and combine the scores generated by several different matchers. Consequently, integration of information at the match score level is the most common approach in multimodal biometric systems. Table 6 provides an outline of several score fusion methods and their associated needs for data that characterise the matcher performance.

In the context of verification, there are two distinct approaches to score-level fusion. One approach is to formulate it as a classification problem, while the other approach is to treat it as a combination problem [36, 39]. In the classification approach, a feature vector is constructed using the matching scores output by the individual matchers; this feature vector is then classified into one of two classes: “Accept” (genuine user) or “Reject” (impostor). Generally, the classifier used for this purpose (e.g., decision tree, neural network, support vector machine, K-nearest neighbour, random forest, etc.) is capable of learning the decision boundary irrespective of how the feature vector is generated [6, 65, 66]. Hence, the output scores of the different modalities can be non-homogeneous (distance or similarity metric, different numerical ranges, etc.) and no processing is required prior to presenting them to the classifier. In the combination approach, the individual matching scores are combined to generate a single scalar score, which is then used to make the final decision [42]. To ensure a meaningful combination of the scores from the different modalities, if necessary, the scores may be first transformed to a common domain prior to combining them. This is known as score normalisation (as discussed in 5.3.2) [27].

As part of a pattern classification problem, in the classification approach, the fusion module design aims at finding an optimal two-class classifier for genuine and impostor classes. The classifier uses the vector of match scores provided by the matchers and assigns one of the two classes to it. For this purpose the classifier defines two decision regions in the feature vector space, one for genuine class and one for impostor class. These regions are separated by decision boundaries, which need to be optimized during the design of the fusion module. These decision boundaries can have various forms depending upon the complexity and the nature of the distributions of the two classes. They can be as simple as a line as in linear discriminant functions or more complex as in multilayer neural networks and support vector machines. The boundaries can also be determined from statistics such as the Neyman-Pearson likelihood ratio. Regardless of the chosen technique, the ultimate goal is to find decision boundaries that improve classification performance to fit the application.

Combination approaches are some of the simplest and most effective methods for biometric fusion, provided scores are homogeneous or can be normalised to make them so. Because of this simplicity and effectiveness they are some of the most common methods for use in multibiometric systems. Kittler’s theoretical framework for combining classifiers [42] describes some of the most popular techniques, these being the product, sum, max, min and median rules. Each of these techniques uses simple arithmetic or rule operations to combine scores from multiple sources. These techniques were extended by Benediktsson and Swain [1] to allow weighting of the match scores based on performance. If more information on the distribution of match scores is available then one may use Bayesian statistics in combining the scores of different biometric matchers as demonstrated by Bigun et al. [3]. Their technique takes into account the estimated accuracy of the individual classifiers during the fusion process. In general, fusion can be accomplished using a Bayesian classifier when sufficient training data is available. Let $P_i(S|G)$ and $P_i(S|I)$ denote the probability densities of score S (corresponding to the i^{th} modality) under the genuine and impostor hypothesis, respectively. A simple Bayesian classifier (SBC) would make a MATCH/NO-MATCH decision based on the posterior densities $P(G|S_1, S_2, \ldots S_N)$ and $P(I\ |S_1, S_2, \ldots S_N)$. In the absence of sufficient training data (i.e., genuine and impostor match scores) it is not possible to reliably estimate the *joint density* involving multiple modalities. Thus, the posterior probability could be estimated by the *product* of individual densities, i.e., $P(G|S_1, S_2, \ldots S_N) = \prod P_i(S_i|G)$ and $P(I|S_1, S_2, \ldots S_N) = \prod P_i(S_i|I)$.

Table 6 — Examples of score fusion methods

Method	Score fusion equation	Characterisation data required					
		None	PDF_G	PDF_I	EER	V_G, V_I	Personal
Simple sum	$\sum$ (i=1 to N) S_i'	O					
Minimum score	min (i=1 to N) S_i'	O					
Maximum score	max (i=1 to N) S_i'	O					
Matcher weighting	$\sum$ (i=1 to N) $W_i \cdot S_i'$				O		
Matcher weighting with PDF fusion for decision [a]	$\sum$ (i=1 to N) $W_i' \cdot S_i'$		O	O			
User weighting	$\sum$ (i=1 to N) $W_i^* \cdot S_i'$						O
Weighted product	$\prod$ (i=1 to N) $W_i \cdot S_i'$				O		
Sum of probabilities Genuine	$\sum$ (i=1 to N) $P_{G\mid Si}$		O				
Sum of probabilities Impostor	$\sum$ (i=1 to N) $P_{I\mid Si}$			O			
Product of probabilities Genuine	$\prod$ (i=1 to N) $P_{G\mid Si}$		O				
Product of probabilities Impostor	$\prod$ (i=1 to N) $P_{I\mid Si}$			O			
BGI [b]	$\prod$ (i=1 to N) BGI_i		O	O			
Likelihood ratio[c]	PDF_G/PDF_I		O	O			
K-nearest neighbour	-					O	
Decision trees	-					O	
Support vector machines	-					O	
Discriminant analysis	-					O	
Neural network	-					O	

NOTE The following symbols and abbreviations are used in the table.

i = i-th biometric score

N = Number of fusion inputs

S_i' = i-th normalized match score

W_i = i-th matcher weight factor

W_i^* = i-th user weight factor

Wi' = i-th matcher weight factor in case of PDF fusion

BGI = Biometric gain against impostors

PDF_G = Probability density functions of scores from genuine users for each dimension

PDF_I = Probability density functions of scores from impostors for each dimension

EER = Equal error rate

V_G = N-dimensional genuine score vector; N is the number of modalities

V_I = N-dimensional impostor score vector; N is the number of modalities

$P_{G\mid Si}$ = Value of PDF_G at score S_i

$P_{I\mid Si}$ = Value of PDF_I at score S_i

[a] Refer to document [64] in Bibliography.

[b] Refer to documents [60, 61] in Bibliography.

[c] Refer to documents [51] in Bibliography.

5.4 Feature-level fusion

In feature-level combination biometric information is fused after feature extraction but before matching [see Figure 3 c)]. There are several ways features can be combined. The simplest form is to integrate the feature vectors (or sets if there is no implicit correspondence) of component biometrics and to apply feature classification methods to the combined feature vector. Where features from contributing multibiometrics are not independent, good feature-level combination should, in some circumstances, allow dependencies to be more fully exploited than by solely using score-level combination. This should give better overall performance. However, fusion at this level is difficult to achieve in practice because of the following reasons: (i) the feature vectors of multiple modalities may be incompatible (e.g., minutiae set of fingerprints and Eigen-coefficients of face); (ii) the relationship between the feature spaces of different biometric systems may not be known; (iii) concatenating two feature vectors may result in a feature vector with very large dimensionality leading to the 'curse of dimensionality'; and (iv) a significantly more complex matcher might be required in order to operate on the concatenated feature vector [56].

Notwithstanding these challenges, fusion at the feature level has been attempted in several contexts. Chang et al. [5] demonstrate feature-level fusion of face and ear modalities showing significant improvements in performance. Kumar et al. [45] integrate the palm-print and hand geometry features of an individual in order to enhance matching performance. In their experiments, fusion at the match score level was observed to be superior to fusion at the feature level. However, Ross and Govindarajan [56] combine the hand and face modalities of a user (multibiometrics) as well as the R, G, B channels of the face image of a user (multisensorial) at the feature level and demonstrate that a feature selection scheme may be necessary to improve matching performance at this level. Thus, it is imperative that an appropriate feature selection scheme is used when combining information at the feature level.

Features can also be combined in a more complex way on an algorithmic level through co-registration. Most feature extraction algorithms require the localization of landmarks in order to establish a common coordinate frame between samples for feature extraction. In multibiometric systems individual components can exchange landmarks or mutually support their extraction. This technique, called co-registration, is considered a form of feature-level combination. For example, a face recognition algorithm may provide eye locations for an iris recognition algorithm, or depth landmarks in a 3D face recognition system may be used to correct the pose of faces in texture images.

6 Characterisation data for multibiometric systems

6.1 Overview

One of the most important aspects of normalisation and combination for multibiometric systems is the origin of parameters for such normalisation and/or combination. In the case of score-level combination using statistical pattern matching theory, the PDFs of genuine and impostor score distributions are required. In other score-level combination and in feature-level and decision-level combination, there are usually important parameters that, in many cases, are required to be derived from characterisation data. Thus this issue is all-pervading.

This clause is allocated to analysis and discussion of characterisation data, its expected origin(s), extent of its validity (e.g., through small sample sizes or other limitations on characterisation sample populations) and how such data would be disseminated or otherwise made available for use.

6.2 Use of characterisation data in normalisation and fusion

Score-level fusion combines the similarity scores from one or more matchers. In the multimodal and multialgorithmic case there will generally be two or more such matching systems. In the multisensor, multiinstance, and multipresentation cases only one matcher will usually be in use, but in any case, multiple scores will be available to a fusion module. The distribution of matcher scores will depend on the matching system and the statistics of these variables will not usually be on any common range. Thus the normalisation process of clause 5.3.2 is a necessary precursor of the fusion process.

The characterisation data, discussed in this section, is needed to support normalisation and fusion. At its most simple this may be just the location and shape parameters of each score's "natural" distribution. For example a face and fingerprint fusion scheme would use some prior estimates of the median and median absolute deviation (see Table 5) to effect normalisation of two scores. More usefully a full specification of the distribution of the scores would be used, and such a description would be provided for both the genuine and impostor distributions.

Thus a biometric system's characterisation data is just some representative summary of the statistics of its output scores. One powerful and simple characterization is the cumulative distribution function (cdf), which may be expressed as N pairs of (S_i, cdf(S_i)) or some functional fit of the data (see [18, 40]).

7 Scope and options for standardisation

7.1 Introduction

This Technical Report lists many ways of combining multibiometric processing and performing biometric fusion. Due to their complexity and number it should be clear that not all of these options can be made part of a biometric fusion standard. To decide which options deserve further study in the standardisation process, one should focus on the interoperability requirements, and through the process of creating the standard, determine which methods best meet those requirements, and in addition, which of those methods have industry consensus for implementation.

7.2 Implementation areas

A scope statement can include use case scenarios indicating where the standard typically will be applied. In some cases, standards themselves define a use-case scenario; SC 37 defines “application profiles” standards in this sense. It is likely that future fusion standard activity may be of four types.

a) Record Formats. The definition and standardisation of data to be exchanged between processes and stored on various media. Biometric Record Formats defined in SC 37/WG 3 are examples of this type of standard.

b) Framework. Definition of standard APIs for processes, the Record Formats used by the processes, and the initialization procedure of the processes in the system. The BioAPI framework [29] defined in SC 37/WG 2 is an example of this type of standard.

c) Application Profile. A list of clauses in either a) or b), and possibly other standards, that are mandatory for a particular use case scenario. The SC 37/WG 4 project on ILO (International Labour Organization) Seafarer ID profile [31] is an example of this type of standard.

d) Conformance Criteria. A description of performance criteria and test data that allows for the assurance that systems have complied with the standards. These types of standards are under development in SC 37 for the biometric record formats.

The use of multibiometric systems has been considered for two major and differing use cases. The first is high-security biometric use where the combination of biometrics provides a stronger assurance of impostor rejection for a relatively small, trained population. The second is in the context of large-scale ID systems, such as travel document systems, where the multibiometric combination provides for the reduction of rejection rates and easier system usage for a very large, untrained population.

In the context of the large-scale ID systems, there can be many solution providers providing components to the overall system. For example, the creator of the electronic biometric document may not be the same vendor that creates the physical document, and neither may be the vendor that performs the biometric test(s) (verification or identification) during the document's usage. This situation can clearly benefit from a biometric fusion standard when the document contains multiple biometrics. Therefore, one would expect a mature standards process to yield two or more application profiles for biometric fusion applications.

7.3 Interoperability requirements

In the context of biometric fusion, one can propose the following interoperability requirements for standardisation on multibiometric systems.

a) **Standard multibiometric systems may be required to be designed and certified (or evaluated) based on common performance requirements.** These performance requirements should be independent of the biometric modalities in use. This includes performance measures such as failure to enroll, failure to acquire, false rejection rate, false acceptance rate, system throughput, and the resistance to active impostor attacks.

b) **Standard multibiometric systems may be required to be designed so that a single biometric subsystem can be separately upgraded.** All biometric device characteristics change over time as research and development improves accuracy and lowers cost. The development of each biometric system however, proceeds on its own timeline. Therefore only if separate upgrading is possible will it be convenient to upgrade a multibiometric system in the field.

c) **A standard multibiometric system may be required to be able to accept historical information for a given user, such as scores and processing times.** With this information, the system can be optimized in both security and throughput to take advantage of the type of biometric modality that is favored by the particular user.

d) **Standard multibiometric systems may be required to be compatible with existing single biometric standard systems.** In particular, SC 37/WG 2 is defining BioAPI for single biometric outputs. Future multibiometric standards or amendments/revisions to existing standards should allow for the use of BioAPI Biometric Service Providers (BSPs) and take into consideration the BioAPI framework [29].

7.4 Possible standardisation activity

7.4.1 On record format standardisation

There are two types of items to define in a framework. Data Records carry information from one Process to another, and Processes convert one set of Data Records into another set. This section will discuss Data Records with the goal of listing a relatively small number of record types to consider. As it is not possible to discuss Data Formats without listing Processes, this section begins with a listing of the basic processes involved with biometric fusion. In Clause 5, Figures 3 a), 3 b), 3 c) and 3 d), the following Processes were defined – Feature Extraction, Matching, Decision Making, Sample Fusion, Feature Fusion, Score Fusion and Decision Fusion. These are shown in Figure 6 below. Each Process has Data Records as inputs and outputs.

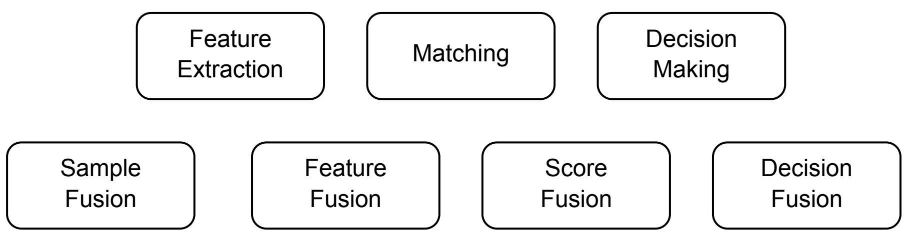


Figure 6 — Multibiometric Processes discussed in Clause 5

The nine Data Record inputs and outputs discussed in Clause 5 are: Biometric Samples (inputs), Biometric Templates (inputs), Feature Extraction Features, Matching Scores, Sample Fusion Process Samples, Feature Fusion Process Features, Score Fusion Process Scores, Decision Fusion Process Decisions, and Decision Making Process Decisions. In addition, each of the seven processes given in Figure 6 above can or must have *additional inputs* associated with initialization, optimization, or a use-case scenario. In summary, there are seven processes and nine records to consider when creating the framework for a biometric fusion standard.

To reduce the set of Records to consider and subsequently define, one would want not to redefine the data types already used by the single biometric standard, BioAPI and that are also unlikely to change when considering multibiometric usage. This excludes from the list of records Biometric Samples (the BioAPI raw or processed data types), and Biometric templates (the BioAPI template type). The BioAPI score is an exception as it is likely to be reconsidered in the context of a multibiometric standard and should therefore be included in the list of multibiometric records.

In addition, each type of fusion process can likely be supported by a single fusion input record with information that denotes the appropriate use case. Therefore, a good starting point for a list of Data Records for a multibiometric framework standard is given by the ten records in Figure 7. These records would be used in conjunction with existing single biometric records to create interoperable data.

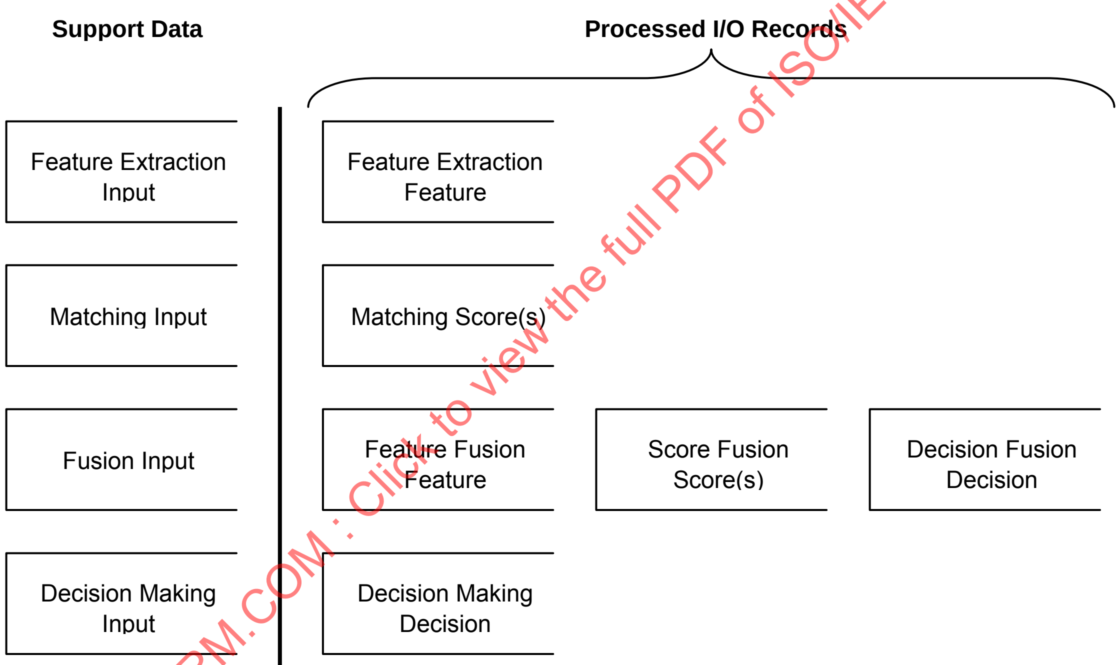


Figure 7 — All records appropriate for definition in a multibiometric standards framework

Examination of the Records listed in Figure 7 indicates that this list can be reduced by unifying the definitions of features, scores, and decisions across processes that use them as inputs and outputs. For example, the Decision Fusion Decision and the Decision Making Decision could be serviced by a single Decision Record. This analysis leaves just seven basic record types to be defined for a simplified multibiometric system as shown in Figure 8. These records would be used in conjunction with existing single biometric records (raw, processed, and template data) and provide data between processes.

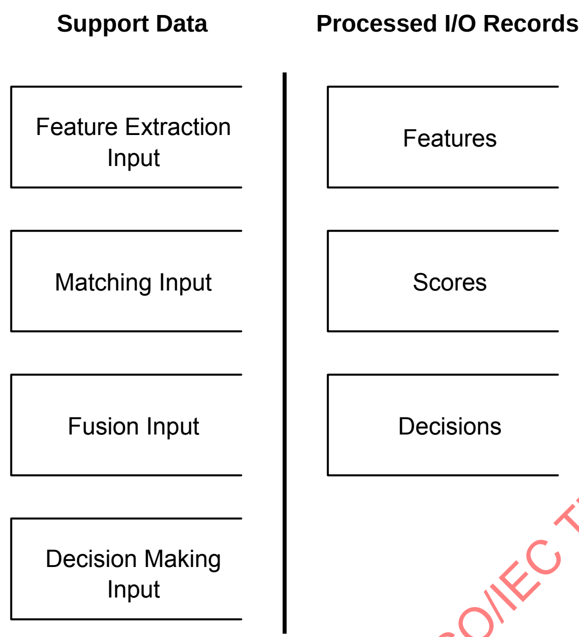


Figure 8 — A simplified set of records appropriate for definition in a multibiometric standards framework

The simplified set of Data Records and Processes for biometric fusion would have seven Process elements and seven Data Records. The definition and use of these elements would be dictated by the interoperability requirements discussed in 7.3.

In particular, the input records can be complex and could contain the following information.

a) The **Feature Extraction Input Record** would contain the user, application profile, and other information required to optimize the extraction of a feature for a given situation.

b) The **Match Input Record** is the non-template data used to optimize or execute the match process – including demographics specific to a given user and also application profile information.

c) The **Fusion Input Record** is the device or user specific information deemed advantageous for optimizing fusion for a particular person, application profile, or biometric configuration (likely in the case of feature-level fusion). It also contains the a-priori information required to perform the mathematical operation of the particular type of normalisation and/or fusion specified by the framework standard or application profile.

d) The **Decision Input Record** is the information deemed advantageous for optimizing decision making for a particular person or application profile. It would also contain the required biometric security level specified in mathematical or statistical terms that the system must try to accomplish.

Note that all newly defined Data Records would be expected to be CBEFF [30] compliant. That is, they would allow for the identification of the creator (vendor), the standards body associated with its definitions, and for encryption of the record for purposes of transmission or storage.

7.4.2 On framework standardisation

The BioAPI specification [29] in SC 37/WG 2 is a good example of a single biometric (usage) framework. It provides nomenclature, data records, programmer APIs, and use-cases in the context of an overall application framework. In the context of the discussion of this Technical Report, it provides a standard score output for each biometric, as well as a biometric decision based upon a requested level of security.

A framework would specify use-cases that combine the records and processes. There may be different frameworks for different levels of fusion. A diagram denoting score-level fusion is shown in Figure 9. Note that only the data flow from one biometric system is shown for simplicity. Note also that there is the possibility of direct decision outputs from Score Fusion. There is also the possibility of feedback between processes.

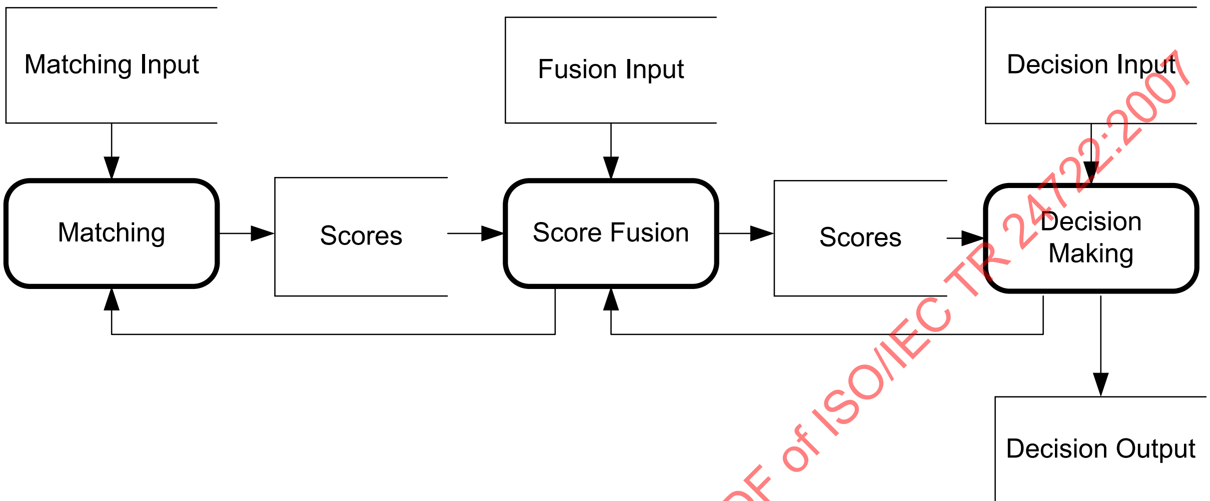


Figure 9 — A framework diagram denoting the use of the newly defined multibiometric records in the context of score-level fusion

Based upon a framework for each fusion type, the standard Records and Process APIs would be determined by consensus in a way that best optimizes the performance and interoperability. Note that for the score-level fusion framework above, there is feedback between processes. The concept is that processes need to communicate to initialize correctly as well as to function appropriately. For example, the Decision Making Process might need to feed the target False Acceptance Rate (FAR) to both the Score Fusion and the Matching Processes to control the operation of the system.

A similar framework diagram for decision-level fusion is shown in Figure 10. Decision Making is still required because Decision Fusion could produce outputs other than decisions, for example rankings or ‘soft’ decisions. Note the similarity of the two approaches from the framework point of view. Note also the possibility of feedback between processes.