



Technical Report

ISO/IEC TR 24722

Information technology — Biometrics — Multimodal and other multibiometric fusion

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

**Third edition
2024-12**

STANDARDSISO.COM : Click to view the full PDF of ISO/IEC TR 24722:2024

STANDARDSISO.COM : Click to view the full PDF of ISO/IEC TR 24722:2024



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2024

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Overview of multimodal and other multibiometric systems	2
4.1 General	2
4.2 Simultaneous and sequential presentation	4
4.2.1 Overview	4
4.2.2 Simultaneous presentation	5
4.2.3 Sequential presentation	5
4.3 Correlation	6
5 Levels of combination	7
5.1 Overview	7
5.2 Decision-level fusion	9
5.2.1 Simple decision-level fusion	9
5.2.2 Advanced decision-level fusion	10
5.3 Score-level fusion	11
5.3.1 Overview	11
5.3.2 Rank-level fusion	11
5.3.3 Score normalization	11
5.3.4 Score fusion methods	15
5.4 Feature-level fusion	17
6 Characterisation data for multibiometric systems	17
6.1 Overview	17
6.2 Use of characterization data in normalization and fusion	18
Bibliography	19

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and <https://patents.iec.ch>. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

This third edition cancels and replaces the second edition (ISO/IEC TR 24722:2015), which has been technically revised.

The main changes are as follows:

- the content of Clause 3 has been removed and ISO/IEC 2382-37 has been listed as a normative reference;
- to enhance information accessibility, symbol descriptors have been paired with clear descriptions;
- the structure of the document has been updated, and various editorial modifications have been made, in order to improve technical accuracy and bring the document in line with the most recent edition of the ISO/IEC Directives Part 2.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

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

Use of multiple biometric measurements from substantially independent biometric sensors, algorithms or characteristic types 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 technical literature since at least 1974.^{[22], [45]} Advanced methods for combining measures at the score level have been discussed in References [15] and [16]. At the current level of understanding, combining results at the score level typically requires knowledge of both mated and non-mated score 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 these questions and the highly application-dependent and generally unavailable data required for proper fusion at the score level, work on multibiometric fusion can in the meantime be considered mature. By intention, this document is not issued as International Standard, in order not to force industrial solutions to conform to the methodology described herein. Rather, the present edition of this document provides a mature technical description for developments of multibiometric systems. It also provides a reference on multibiometric fusion for developers of other biometric standards and implementers.

STANDARDSISO.COM : Click to view the full PDF of ISO/IEC TR 24722:2024

Information technology — Biometrics — Multimodal and other multibiometric fusion

1 Scope

This document provides descriptions and analyses of current practices on multimodal and other multibiometric fusion, including (as appropriate) references to more detailed descriptions.

This document contains descriptions and explanations of high-level multibiometric concepts to aid in the explanation of multibiometric fusion approaches including: multi-characteristic-type, multi-instance, multi-sensorial, multialgorithmic, decision-level and score-level logic.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 2382-37, *Information technology — Vocabulary — Part 37: Biometrics*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 2382-37 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1

multialgorithmic

using multiple algorithms for processing the same biometric sample

3.2

multibiometrics

automated recognition of individuals based on their biological or behavioural characteristic and involving the use of biometric fusion

3.3

multi-characteristic-type

multi-type

using information from multiple types of biometric characteristic

EXAMPLE Biometric characteristic types include: face, voice, finger, iris, retina, hand geometry, signature/sign, keystroke, lip movement, gait, scent, vein, DNA, ear, foot, etc.

3.4

multi-instance

requiring two or more instances of a biometric characteristic

EXAMPLE Iris (left) + Iris (right), Fingerprint (left index) + Fingerprint (right index).

Note 1 to entry: [SOURCE: ISO/IEC 2382-37:2022, 37.03.47, modified — Note 1 to entry has been removed and an Example has been added.]

3.5 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 1 to entry: Multipresentation biometrics is considered a form of multibiometrics if fusion techniques are employed. Many fusion and normalization techniques are appropriate to the integration of information from multiple presentations of the same biometric instance.

3.6 sequential presentation

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

3.7 simultaneous presentation

capturing biometrics 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 characteristic type, 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, comparison scores or comparison 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.^[22,45] 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 data subject 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 largescale automated fingerprint identification systems (AFISs), starting in the 1980s, and was further expanded upon with the introduction of automated biometric identification systems (ABISs) in the 1990s. Most methods of fusion discussed elsewhere in this document have been successfully implemented using fingerprints alone. Some of the ways that fusion has been implemented in AFISs include:

- image fusion (also known as sample fusion), where a single “rolled” image is created from a series of plain impressions on a liveness device;
- template fusion, where features extracted from several presentations are combined into a single template;
- multi-instance fusion, which uses fingerprints from all ten fingers;
- multipresentation fusion, which uses rolled and slap (plain) fingerprints;
- algorithm fusion for the purpose of efficiency (cost, computational complexity, and throughput rate), where comparators are generally 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), where comparators are used in parallel, with fusion of resulting scores.

The use of fusion has made AFIS and ABIS possible because of fusion's potential in improving both accuracy and efficiency.

Most work to date on multibiometrics has focused only on improving false acceptance and false rejection error rates. Some research work considers the use of multibiometrics to flexibly improve usability, security or accuracy.^[61] Further, multibiometrics also aims at decreasing the overall failure-to-enrol rate (FTE) especially in biometric systems where data subject cooperation is not expected (e.g. video surveillance systems). Multibiometrics is an effort to produce a biometric decision even if only a subset of the expected biometric characteristic were captured.^[63]

To further develop the understanding of the distinction among the multibiometric categories, [Table 1](#) illustrates the basic distinctions among categories of multibiometric implementation. The key aspect of the category that makes it multi-“something” is explained below the table.

Table 1 — Multibiometric categories illustrated by the simplest case of using 2 elements

Category	Characteristic type	Algorithm	Instance	Sensor	Presentation
Multi-characteristic-type	2 (always)	2 (always)	2 (always)	2 (usually) ^b	at least 1
Multialgorithmic	1 (always)	2 (always)	1 (always)	1 (always)	at least 1
Multi-instance	1 (always)	1 (always)	2 (always)	1 (usually) ^c	at least 1
Multi-sensorial	1 (always)	1 (usually) ^a	1 (always, and same instance)	2 (always)	at least 1
Multipresentation	1	1	1	1	at least 2

^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 An exception is a multi-characteristic-type system with a single sensor used to capture two different characteristic types. For example, a high resolution image used to extract face and iris or face and skin texture.

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

- Multi-characteristic-type biometric systems — these systems take input from single or multiple sensors that capture two or more different types of biometric characteristic. For example, a single system combining face and iris information for biometric recognition would be considered a “multi-characteristic-type” 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 combined in any mathematically complex way. For example, a system with fingerprint and voice recognition would be considered “multi-characteristic-type” even if the “OR” rule was being applied, allowing data subjects to be verified using either of the characteristic types.
- Multialgorithmic biometric systems — these 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 characteristic type. Maximum benefit (theoretically) would be derived from algorithms that are based on distinctly different and independent principles such as either features they extract from the biometric sample (e.g. finger minutiae versus finger pattern) or approaches to comparison (e.g. different algorithms comparing minutiae).
- Multi-instance biometric systems — these 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 multi-instance rather than multi-characteristic-type. However, systems capturing, for example, sequential frames of facial or iris images are considered to be multipresentation rather than multi-instance.

- Multi-sensorial biometric systems — these 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.
- Multipresentation — the biometric system uses multiple samples of one instance of a biometric characteristic.

For a specific application in an operational environment, there are numerous system design considerations and trade-offs that would need to 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.^[40]

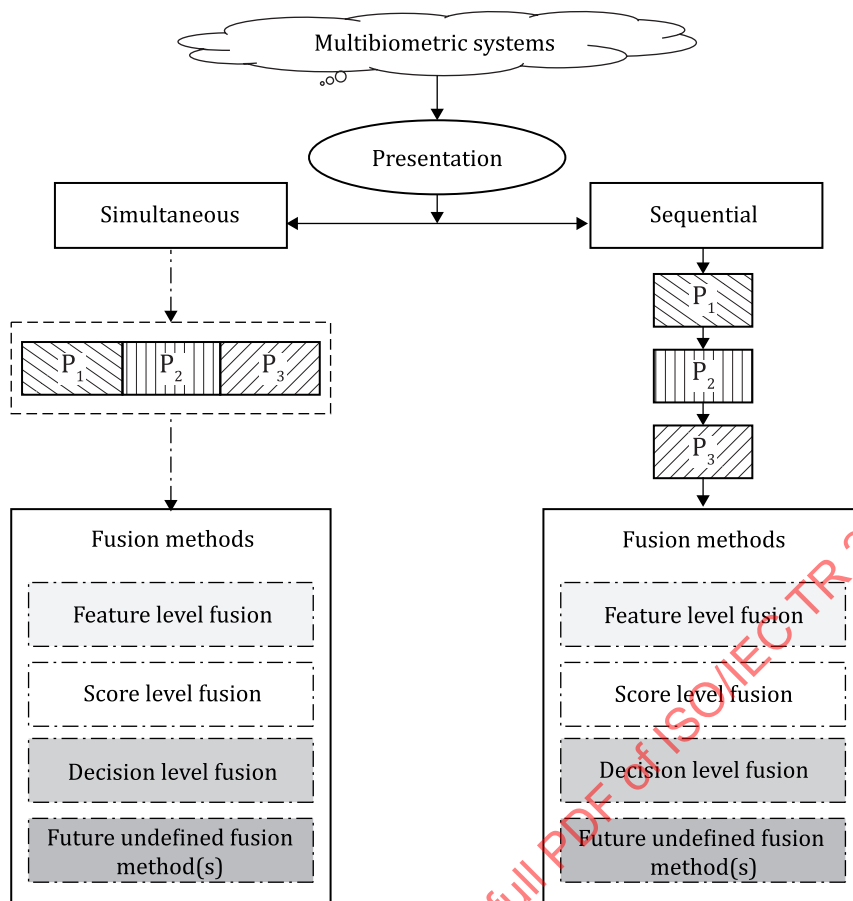
Especially for a large-scale human identification system, there are additional system design considerations, such as operation and maintenance, reliability, system acquisition cost, life cycle cost, and planned system response to identified susceptible means of attack, all of which will affect the overall deployability of the system.^[40]

4.2 Simultaneous and sequential presentation

4.2.1 Overview

Dependent upon the system design, there are two methods of presenting a biometric characteristic for capture by the system:

- 1) simultaneous; and
- 2) sequential.



NOTE The presentation (simultaneous or sequential) method generally induces different fusion processes. The purpose of including this information is to illustrate considerations that can potentially influence multibiometric system design.

Figure 1 — Classification of multibiometric systems by simultaneity of presenting biometric characteristic

4.2.2 Simultaneous presentation

Simultaneous presentation (with successful capture) provides biometric sample(s) from multiple characteristic types in a single event (e.g. a face and iris taken from the same camera). System designs that utilize simultaneous acquisition would tend towards high throughput applications at the expense of possibly adding complexity (to synchronize sample collection) or difficulty of use (dual sensor interaction, data subject multitasking).

4.2.3 Sequential presentation

Sequential capture acquires biometric sample(s) from one or multiple characteristic types in separate events. Sequential capture may be utilized in the three concepts discussed in the literature.^[65] The first is multi-instance, which is the use of two or more instances within one characteristic type 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 multi-characteristic-type, which is the use of multiple different biometric characteristic types captured from one or more sensors for a subject, i.e. Hand + Face in sequence. The third concept is multi-sensorial, which is the use of two or more distinct sensors for capturing the same biometric feature(s) (e.g. traits) for a subject, but not at the same time. To avoid confusion with multi-characteristic-type, which can also capture biometric instance(s) from two or more distinct sensors, multi-sensorial can be clarified as “uni-characteristic-type multi-sensorial”. 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 can be correlated at several different levels^[53] as illustrated in the following examples.

- Correlation between characteristic types: this refers to biometric samples that are physically related such as the speech and lip movement of a data subject.
- 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 can be used by one algorithm and part of the sample by another) is subjected to different feature extraction and comparison algorithms (e.g. a minutiae-based comparator and a texture-based comparator).
- Correlation between feature values: a subset of feature values constituting the feature vectors of different characteristic types can be correlated. For example, the area of a data subject's palm (hand geometry) can 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 comparison scores (or the ACCEPT/REJECT decision) pertaining to the comparators 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 biometric performance.^[53]

For two classifiers of reasonable accuracy involved in a fusion scheme, score outputs from inputs that come from the same subject can, but need not, be correlated. In the case of decision level fusion, it is appropriate to consider the correlation of classifier errors as described by Goebel, Yan, and Cheetham.^[20] The correlation ρ_{n_c} is given by [Formula \(1\)](#):

$$\rho_{n_c} = \frac{nN_c^f}{N - N_c^t - N_c^f + nN_c^f} \quad (1)$$

where

n is the number of classifiers under test;

N is the total number of multibiometric information channels;

C is the threshold

N_c^f is the number of multibiometric information channels where all classifiers have an incorrect output at threshold C ;

N_c^t is the number of multibiometric information channels where all classifiers have a correct output at a threshold C .

Assessing score level correlation is inherently more difficult as it will depend on the normalization used (see [5.3.3](#)), but [Formula \(1\)](#), together with a relevant threshold on score can give a very rough first idea even in those cases.

5 Levels of combination

5.1 Overview

As a basis for the definition of levels of combination in multibiometric systems, this document first introduces the single-biometric process and its building blocks, using the example of an authentication system for simplification without PAD mechanisms. [Figure 2](#) shows the block diagram of a single-biometric process.

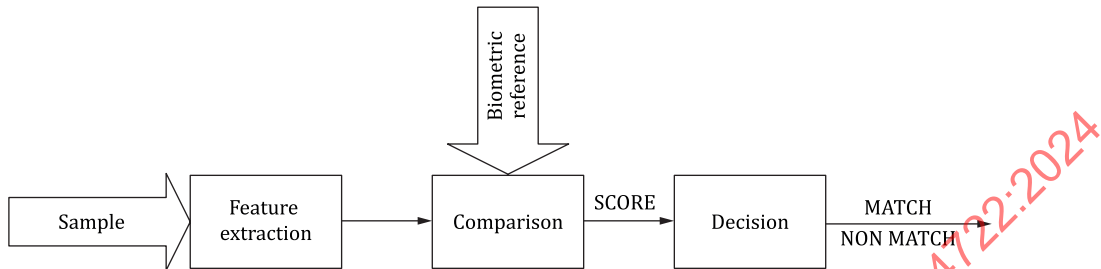
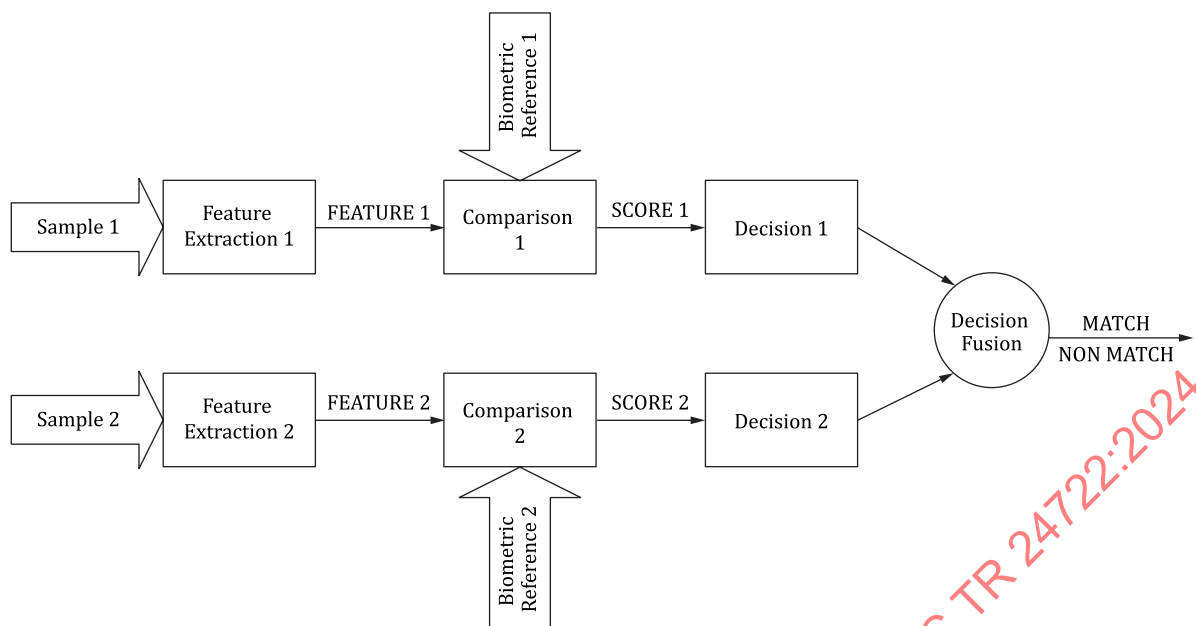


Figure 2 — Single biometric process (generic)

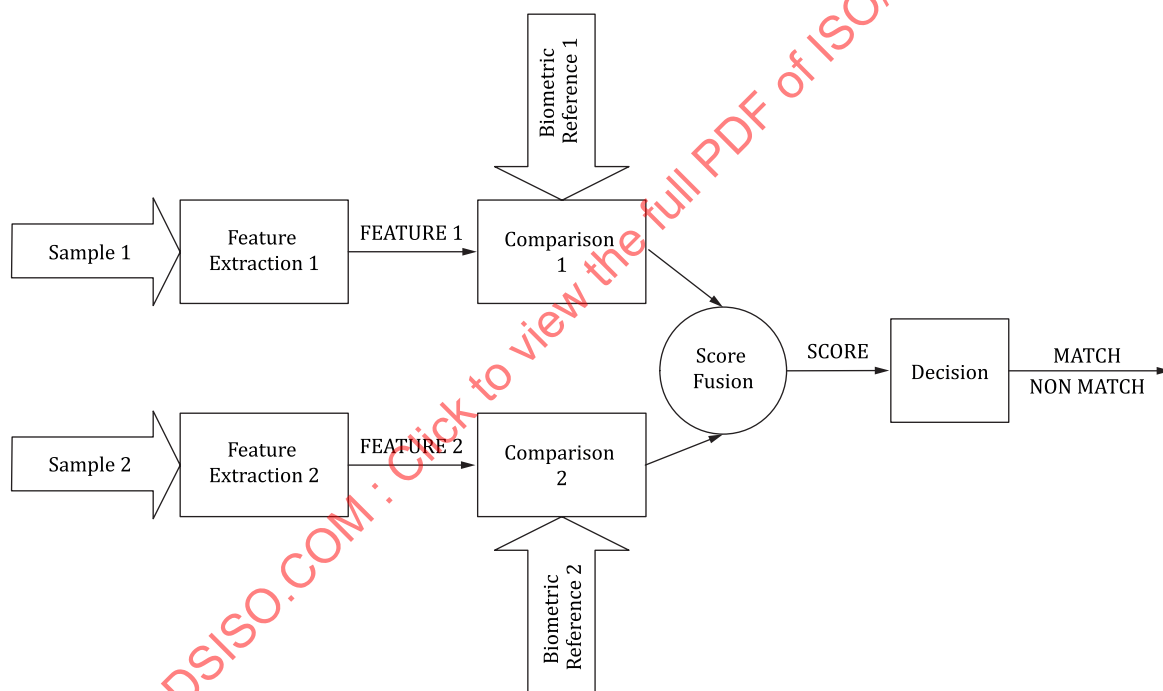
A biometric sample captured by a biometric sensor (e.g. a fingerprint image) is fed into the feature extraction module. Using signal processing methods, the feature extraction module converts a sample into features (e.g. fingerprint minutiae), which form a representation apt for comparison. Usually, multiple features are collected into a feature vector. The comparison module takes the feature vector as input and compares it to a biometric reference. The result is a comparison 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 biometric information channels, there are several levels at which fusion can take place. These include consolidating information at the (i) decision level, (ii) comparison score level, (iii) feature level, and (iv) sample level. Fusion at levels (i) and (ii) occurs after the comparison module is invoked, while levels (iii) and (iv) occur before the comparator. Although integration is possible at these different levels, fusion at the feature set level, the comparison score level and the decision level are the most commonly used. [Figure 3](#) illustrates the different levels of fusion for the case of a multimodal system.^{[7], [41]}

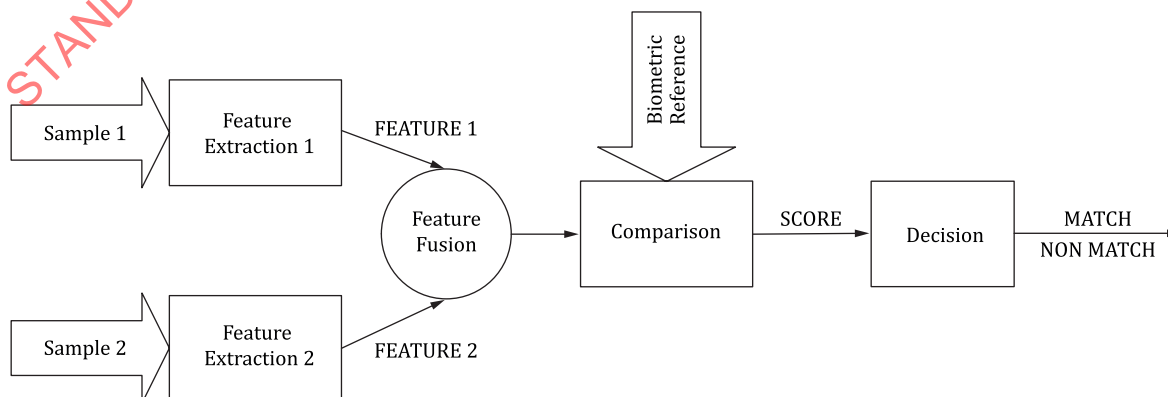
- Decision level: each individual biometric process outputs its own Boolean result. The fusion process fuses them together by a combination algorithm such as AND and OR, possibly taking further parameters such as sample quality scores as input.
- Score level: each individual biometric process typically outputs a single comparison 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.
- 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.
- 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

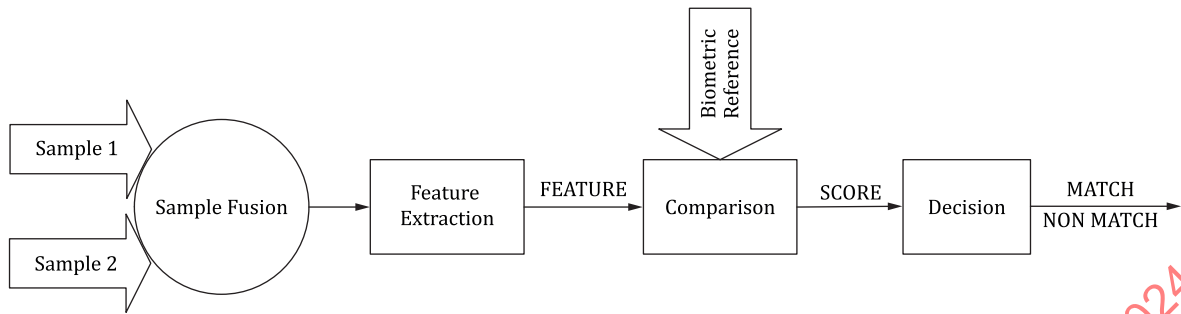


b) Score-level fusion



NOTE Sample 1 and Sample 2 can be the same sample.

c) Feature-level fusion



d) Sample-level fusion

Figure 3 — Different levels of fusion for the case of a multimodal system

For simultaneous or sequential biometric sample acquisition, features are extracted and are compared against the template. How the comparison scores are determined is system-dependent and outside the scope of this document. The comparison 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 can 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 comparison decision has been made for each biometric information channel. 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 information channels, it is convenient to assign logical values to comparison 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 2](#), assuming a pair of decision-level outputs.

Table 2 — AND & OR fusion of decisions for a case of two biometric characteristic types

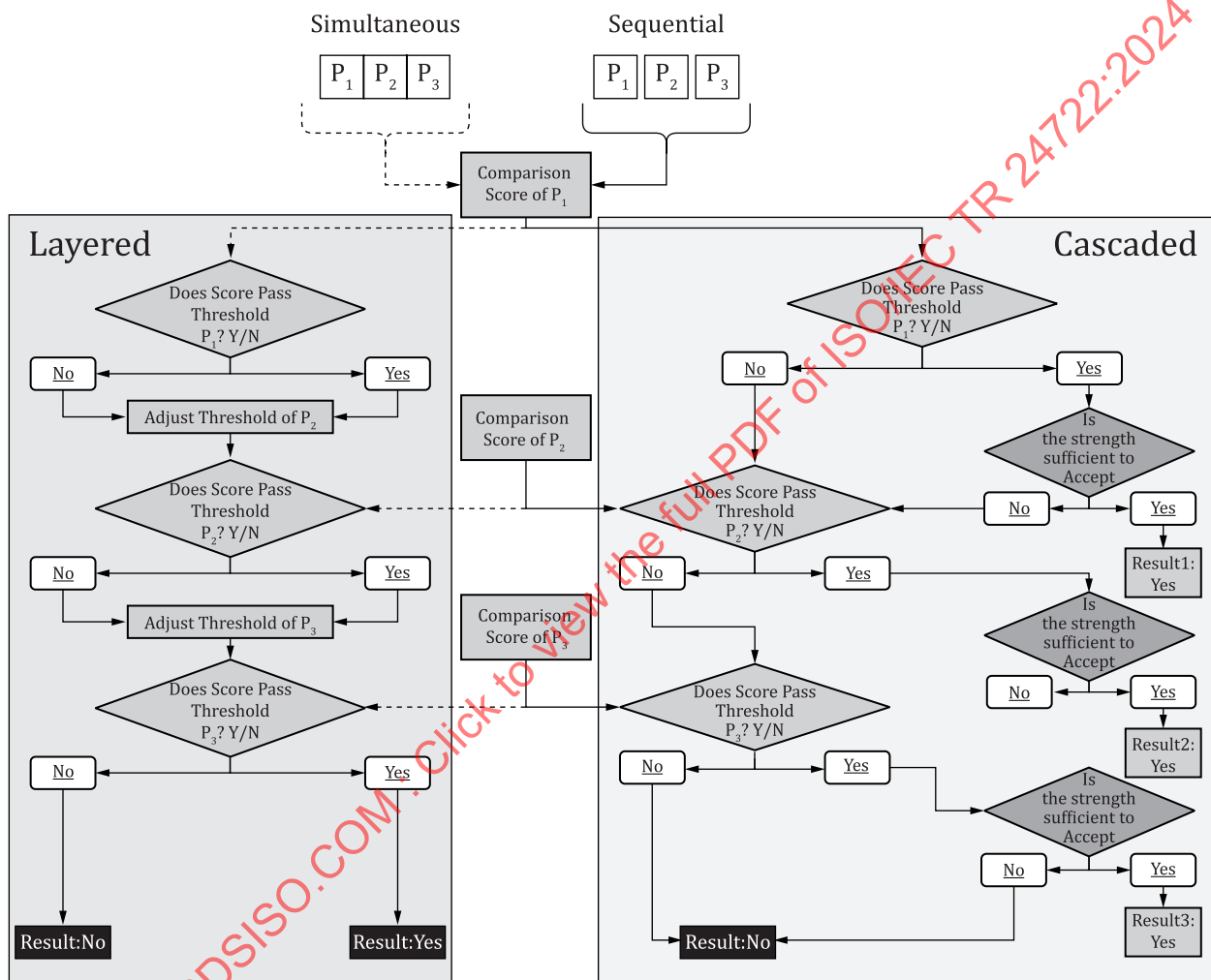
Decision Biometric information channel 1	Decision Biometric information channel 2	AND-fused decision	OR-fused decision
X	X	X	X
X	•	X	•
•	X	X	•
•	•	•	•
Key X Non-match • Match			

For biometric systems using many information channels, 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 features adjustable thresholds computed by using individual biometric scores to determine the pass/fail thresholds for other biometric data processes. A cascaded system features fixed thresholds as pass/fail thresholds of characteristic type-specific biometric samples to determine if additional biometric samples from other characteristic types are required to reach an overall system decision. Decision-level fusion for the two subgroups is shown in Figure 4.



NOTE 1 The left side of the figure presents the layered system and the right side presents the cascaded system.

NOTE 2 The processes, P_i , representing the fused biometric information channels are denoted as P_1 , P_2 and P_3 .

Figure 4 — Advanced decision-level fusion

5.2.2.2 Layered system

Independent of whether the presentation was simultaneous or sequential, the comparison 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 characteristic type P_1 , the output would adjust (raise or lower) the threshold needed to pass for characteristic type P_2 . If P_1 fails to meet the criteria/threshold for characteristic type P_1 , then the output would most likely increase the threshold required for characteristic type P_2 . Upon completion of processing P_1 and resetting the thresholds requirements for characteristic type P_2 , the comparison score of

P_2 enters the system. The process iterates as discussed above for P_2 and P_3 . Once the characteristic type 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 compared. Using [Figure 4](#) as the model for this discussion, comparison score P_1 enters the system and is compared 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 . It is not necessary for cascaded systems to 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 comparison scores indicating the proximity of the feature vector with the biometric reference vector. These scores can then be combined to improve the comparison performance.

From a theoretical point of view, biometric processes can be combined reliably to give a guaranteed improvement in comparison performance. Any number of suitably characterized biometric processes can have their comparison 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 correctly identify the method which will combine these comparison scores reliably and maximize the improvement in comparison performance.

The mechanism (for this sort of good combination of scores within a multibiometric system) will ideally follow at least two guidelines. Firstly, each biometric process is expected to 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 is expected to make available to the multibiometric combiner, its technical performance (such as score distributions) in the appropriate form (and with sufficient accuracy of characterization).

5.3.2 Rank-level fusion

Both verification (1:1) and identification (1:N) systems can support fusion at the comparison 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 comparison scores pertaining to these identities is available. Ho et al. [\[23\]](#) describe three methods for combining the ranks assigned by the different comparators. In the "highest rank" method, each possible match is assigned the highest (minimum) rank as computed by different comparators. 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 comparators 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.

5.3.3 Score normalization

Score normalization 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 can be based on modifying other statistical measures of the comparison score distributions.

The parameters used for normalization can be determined using a fixed training set or adaptively based on the current feature vector.

NOTE 1 The computed characteristic can represent only “estimates” of the underlying population characteristic.

Score normalization is closely related to score-level fusion since it affects how scores are combined and interpreted in terms of biometric performance. As discussed in Reference [32]:

- The comparison scores at the output of the individual comparators need not be homogeneous. For example, one comparator can output a distance (dissimilarity) measure while another can output a similarity measure;
- Further, the outputs of the individual comparators need not be on the same numerical scale (range).
- Finally, the comparison scores at the output of the comparators are permitted to 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 normalization into account.

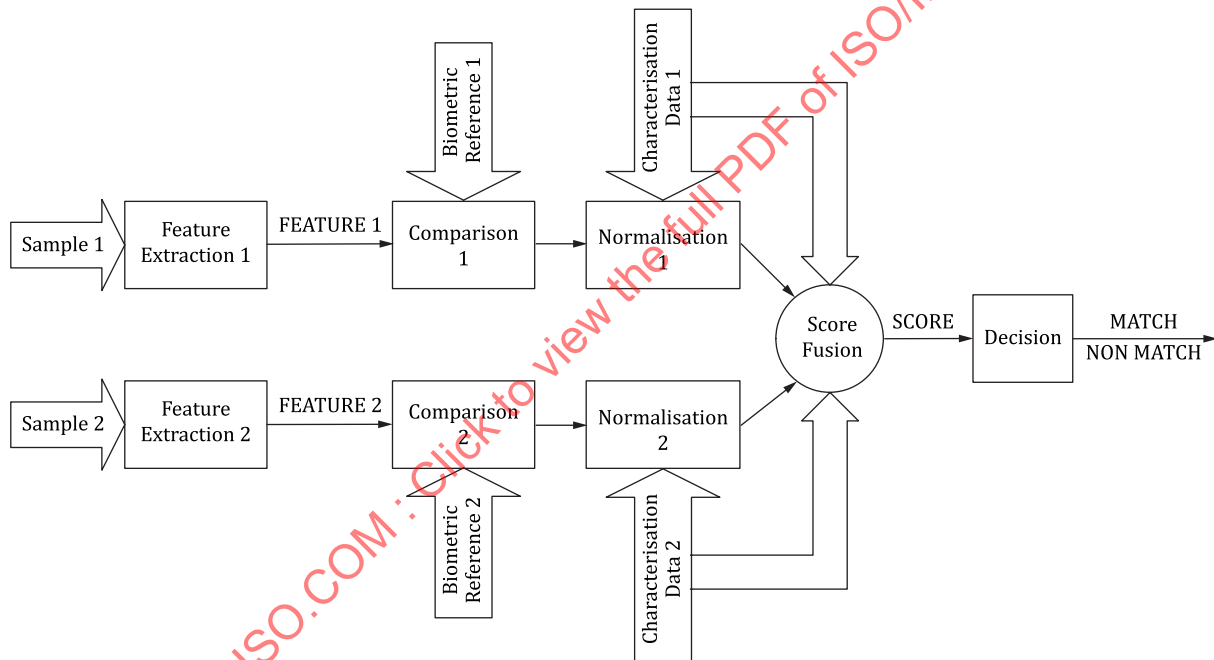


Figure 5 — A framework for score-level fusion

[Table 4](#) lists, under the framework of [Figure 5](#), several commonly used score normalization methods.

NOTE 2 Some fusion methods use probability density functions (PDFs) directly and do not require normalization methods.

[Table 3](#) defines the symbols used in [Table 4](#). In some cases, PDFs are used to convert raw/native scores directly into Probability of False Accept, and thus to a decision, without needing to have native scores brought to a common reference range using normalization.

Table 3 — Symbols used for score normalization formulae

Statistical measures	Characterization data		
	Mated distribution	Non-mated distribution	Both mated and non-mated distributions
Minimum score	$S_{M,min}$	$S_{NM,min}$	$S_{B,min}$
Maximum score	$S_{M,max}$	$S_{NM,max}$	$S_{B,max}$
Mean score	$S_{M,mean}$	$S_{NM,mean}$	$S_{B,mean}$
Median score	$S_{M,med}$	$S_{NM,mean}$	$S_{B,mean}$
Score standard deviation	$S_{M,SD}$	$S_{NM,SD}$	$S_{B,SD}$
Constant	C	C	C
Probability density function	PDF_M	PDF_{NM}	N.A.
Centre of PDF crossover	S_{centre}		
Width of PDF crossover	S_{width}		
Key S similarity score M mated score NM non-mated score B both $S_{M,min}$ minimum of mated similarity score $S_{M,max}$ maximum of mated similarity score $S_{M,mean}$ mean of mated similarity score $S_{M,med}$ median of mated similarity score $S_{M,SD}$ standard deviation of mated similarity score $S_{NM,min}$ minimum of non-mated similarity score $S_{NM,max}$ maximum of non-mated similarity score $S_{NM,mean}$ mean of non-mated similarity score $S_{NM,med}$ median of non-mated similarity score $S_{NM,SD}$ standard deviation of non-mated similarity score $S_{B,min}$ minimum of both mated and non-mated similarity score $S_{B,max}$ maximum of both mated and non-mated similarity score $S_{B,mean}$ mean of both mated and non-mated similarity score $S_{B,med}$ median of both mated and non-mated similarity score $S_{B,SD}$ standard deviation of both mated and non-mated similarity score PDF_M probability density function of mated distribution PDF_{NM} probability density function of non-mated distribution			

Table 4 — Examples of score normalization 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	$S' = (S - S_{NM,mean}) / S_{NM,SD}$	$S_{NM,mean}$ $S_{NM,SD}$	— Assumes normal distribution. — Symmetric about mean. — Assumes stability of both distributions across populations.
Median absolute deviation (MAD)	$S' = (S - S_{B,med}) / (C \cdot \text{median} S - S_{B,med})$	$S_{B,med}$ C	— Assumes stability of both distributions across populations.
Hyperbolic tangent (Tanh)	$S' = 0.5(\tanh(C(S - S_{M,mean}) / S_{M,SD}) + 1)$	$S_{M,mean}$ $S_{M,SD}$ C	— Mean and variance of transformed data distribution. — Assumes stability of both distributions across populations.
Adaptive (AD) ^a a) Two-quadratics (QQ) b) Logistic c) Quadric-line-quadric (QLQ)	$n_{AD} = \begin{cases} \frac{1}{c} n_{MM}^2, & n_{MM} \leq c \\ c + \sqrt{(1-c)(n_{MM} - c)}, & \text{otherwise} \end{cases}$ $n_{AD} = \frac{1}{1 + A \cdot e^{-B \cdot n_{MM}}}$ $n_{AD} = \begin{cases} \frac{1}{\left(c - \frac{w}{2}\right)} n_{MM}^2, & n_{MM} \leq \left(c - \frac{w}{2}\right) \\ n_{MM}, & \left(c - \frac{w}{2}\right) < n_{MM} \leq \left(c + \frac{w}{2}\right) \\ \left(c + \frac{w}{2}\right) + \sqrt{\left(1 - c - \frac{w}{2}\right) \left(n_{MM} - c - \frac{w}{2}\right)}, & \text{otherwise} \end{cases}$	c w Δ $A = \frac{1}{\Delta} - 1$ $B = \frac{\ln A}{c}$	— Assumes non-linearity. — 3 modelling methods. — Assumes stability of both distributions across populations. — n_{AD} = adaptive normalization score; n_{MM} = min-max normalized score; c = centre of overlap of mated and non-mated score distributions; w = width of the overlap; Δ = a small value (0.01 in Reference [59]).
Biometric gain against impostors (non-mated) (BGI)	$PDF^{NM}(S_i) / PDF^M(S_i)$	PDF^M PDF^{NM}	— Assumes stability of both distributions across populations.
BioAPI	$S' = FAR_{\text{threshold}} = \text{score}$	PDF^{NM}	— Assumes stability of non-mated distribution.
Borda count	$N - \text{Rank}(S)$ (where N is the number of alternatives)	Rank	— Applicable only to 1:N comparison.

NOTE This table lists two types of normalization 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 mated and non-mated 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 Reference [59].

5.3.4 Score fusion methods

When each of a set of biometric comparators outputs a comparison score, integration can be done at the comparison score level. This is also known as fusion at the measurement level or confidence level. The comparison score output by a comparator 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 comparators. Consequently, integration of information at the comparison score level is the most common approach in multimodal biometric systems. [Table 5](#) provides an outline of several score fusion methods and their associated needs for data that characterize the comparator performance.

NOTE This is valid only in the case where a rank or/and a comparison score is/are available for all references present in the set of possible matches given by each algorithm.

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. [\[32\]](#), [\[35\]](#) In the classification approach, a feature vector is constructed using the comparison scores output by the individual comparators; this feature vector is then classified into one of two classes: "Accept" (assuming a mated comparison trial) or "Reject" (assuming a non-mated comparison trial). 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. [\[6\]](#), [\[61\]](#), [\[62\]](#) Hence, the output scores of the different characteristic types 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 comparison scores are combined to generate a single scalar score, which is then used to make the final decision. [\[38\]](#) To ensure a meaningful combination of the scores from the different characteristic types, if necessary, the scores can be first transformed to a common domain prior to combining them. This is known as score normalization (as discussed in [5.3.3](#)). [\[27\]](#)

As part of a pattern classification problem, in the classification approach, the fusion module design aims at finding an optimal two-class classifier for mated and non-mated classes. The classifier uses the vector of comparison scores provided by the comparators 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 mated class and one for non-mated 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 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 [\[38\]](#) 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 comparison scores based on performance. If more information on the distribution of comparison scores is available, then Bayesian statistics can be used in combining the scores of different biometric comparators 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} characteristic type) under the mated and non-mated hypothesis, respectively. A simple Bayesian classifier (SBC) would make a MATCH/NO-MATCH decision based on the posterior densities $P(G|S1, S2, \dots, SN)$ and $P(I|S1, S2, \dots, SN)$. In the absence of sufficient training data (i.e. mated and non-mated comparison scores) it is not possible to reliably estimate the "joint density" involving multiple characteristic types. Thus, the posterior probability could be estimated by the product of individual densities, i.e. $P(G|S1, S2, \dots, SN) = \prod P_i(S_i|G)$ and $P(I|S1, S2, \dots, SN) = \prod P_i(S_i|I)$.

Table 5 — Examples of score fusion methods, ignoring correlation

Method	Score fusion formula	Characterization data required					
		None	PDFG	PDFI	EER	V_G, V_I	Personal
Simple sum	$\sum_{(i=1 \text{ to } N)} S_i'$	0					
Minimum score	$\min_{(i=1 \text{ to } N)} S_i'$	0					
Maximum score	$\max_{(i=1 \text{ to } N)} S_i'$	0					
Comparator weighting	$\sum_{(i=1 \text{ to } N)} W_i \cdot S_i'$				0		
Comparator weighting with PDF fusion for decision ^a	$\sum_{(i=1 \text{ to } N)} W_i' \cdot S_i'$		0	0			
Data subject weighting	$\sum_{(i=1 \text{ to } N)} W_i^* \cdot S_i'$						0
Weighted product	$\prod_{(i=1 \text{ to } N)} W_i \cdot S_i'$				0		
Sum of probabilities mated	$\sum_{(i=1 \text{ to } N)} P_{M S_i}$		0				
Sum of probabilities non-mated	$\sum_{(i=1 \text{ to } N)} P_{NM S_i}$			0			
Product of probabilities mated	$\prod_{(i=1 \text{ to } N)} P_{M S_i}$		0				
Product of probabilities non-mated	$\prod_{(i=1 \text{ to } N)} P_{NM S_i}$			0			
Biometric gain against impostors (non-mated) ^b	$\prod_{(i=1 \text{ to } N)} BGI_i$		0	0			
Likelihood ratio ^c	PDF^M / PDF^{NM}		0	0			
K-nearest neighbour						0	
Decision trees						0	
Support vector machines						0	
Discriminant analysis						0	
Neural network						0	
Key <i>i</i> <i>i</i> -th biometric score <i>N</i> number of fusion inputs <i>S</i> ' <i>i</i> -th normalized comparison score <i>W</i> _{<i>i</i>} <i>i</i> -th comparator weight factor <i>W</i> ' _{<i>i</i>} <i>i</i> -th comparator weight factor in case of PDF fusion <i>W</i> * _{<i>i</i>} <i>i</i> -th data subject weight factor <i>PDF</i> ^M probability density functions of scores from mated data subjects for each dimension <i>PDF</i> ^{NM} probability density functions of scores from non-mated results for each dimension <i>EER</i> equal error rate <i>V</i> _M <i>N</i> -dimensional mated score vector; <i>N</i> is the number of modality <i>V</i> _{NM} <i>N</i> -dimensional non-mated score vector; <i>N</i> is the number of modality <i>P</i> _{M <i>S</i>_{<i>i</i>}} value of <i>PDF</i> _M at score <i>S</i> _{<i>i</i>} <i>P</i> _{NM <i>S</i>_{<i>i</i>}} value of <i>PDF</i> _{NM} at score <i>S</i> _{<i>i</i>}							
^a Refer to Reference [61].							
^b Refer to References [57] and [58].							
^c Refer to Reference [48].							

5.4 Feature-level fusion

In feature-level combination, biometric information is fused after feature extraction but before comparison [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, in some circumstances, good feature-level combination will potentially allow dependencies to be more fully exploited than solely using score-level combinations. Feature normalization is normally used before combining the real valued features (especially in case of feature concatenation). However, in case of binary features fusion feature normalization is not used. This is expected to give better overall performance. However, fusion at this level is difficult to achieve in practice because of the following reasons:

- a) the feature vectors of multiple characteristic types can be incompatible (e.g. minutiae set of fingerprints and Eigen-coefficients of face);
- b) the relationship between the feature spaces of different biometric systems can be unknown;
- c) concatenating two feature vectors can result in a feature vector with very large dimensionality leading to the "curse of dimensionality";
- d) a significantly more complex comparator can be required to operate on the concatenated feature vector.^[52]

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 characteristic types showing significant improvements in performance. Kumar et al.^[41] integrate the palm-print and hand geometry features of an individual in order to enhance biometric performance. In their experiments, fusion at the comparison score level was observed to be superior to fusion at the feature level. However, Ross and Govindarajan^[52] combine the hand and face characteristic types of a data subject (multibiometrics) as well as the R, G, B channels of the face image of a data subject (multi-sensorial) at the feature level and demonstrate that a feature selection scheme can be necessary to improve comparison performance at this level. Thus, it is imperative to use an appropriate feature selection scheme 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 normalization and combination for multibiometric systems is the origin of parameters for such normalization and/or combination. In the case of score-level combination using statistical pattern recognition theory, the PDFs of mated and non-mated 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 characterization data. Thus, this issue is all-pervading and conditions the relevance of theoretical analysis of the optimal fusion rule.

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

6.2 Use of characterization data in normalization and fusion

Score-level fusion combines the similarity scores from one or more comparators. In the multi-characteristic-type and multialgorithmic case there will generally be two or more such comparison systems. In the multi-sensor, multi-instance, and multipresentation cases only one comparator will usually be in use, but in any case, multiple scores will be available to a fusion module. The distribution of comparison scores will depend on the comparison system and the statistics of these variables will not usually be on any common range. Thus, the normalization process of 5.3.3 is a necessary precursor of the fusion process.

The characterization data, discussed in this subclause, is needed to support normalization and fusion. At its simplest, this can 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 4) to effect normalization of two scores. More usefully, a full specification (approximated) of the distribution of the scores would be used, and such a description would be provided for both the mated and non-mated distributions.

Thus, a biometric system's characterisation data is just a representative summary of the statistics of its output scores. One powerful and simple characterization is the cumulative distribution function (cdf), which can be expressed as N pairs of $(S_i, cdf(S_i))$ or a functional fit of the data (see References [18] and [36]).

In score normalization-based processes, fusion is preceded by a transformation of each score to a common domain. The fusion information format defined in ISO/IEC 29159-1 is intended to flexibly support any of the popular transformations. By establishing a standardized means of data exchange this standard supports a modular approach to biometric systems integration in which both the comparison and fusion algorithms remain protected as pieces of intellectual property. Thus, ISO/IEC 29159-1 envisages an application in which two (or more) underlying acquisition and comparison technologies (hand geometry and fingerprint, for example) each generate a score which is fed to a fusion module which has been initialized with an appropriate instance of the fusion information format defined in ISO/IEC 29159-1.