

Australian Standard[®]

**Information technology—Biometric
performance testing and reporting**

Part 1: Principles and framework



This Australian Standard® was prepared by Committee IT-032, Biometric and Identification. It was approved on behalf of the Council of Standards Australia on 9 March 2007. This Standard was published on 2 July 2007.

The following are represented on Committee IT-032:

- AUSTROADS
 - Attorney General's Department
 - Australian Business Limited
 - Australian Electrical and Electronic Manufacturers Association
 - Centrelink
 - The Biometric Institute
-

This Standard was issued in draft form for comment as DR 06645.

Standards Australia wishes to acknowledge the participation of the expert individuals that contributed to the development of this Standard through their representation on the Committee and through public comment period.

Keeping Standards up-to-date

Australian Standards® are living documents that reflect progress in science, technology and systems. To maintain their currency, all Standards are periodically reviewed, and new editions are published. Between editions, amendments may be issued.

Standards may also be withdrawn. It is important that readers assure themselves they are using a current Standard, which should include any amendments that may have been published since the Standard was published.

Detailed information about Australian Standards, drafts, amendments and new projects can be found by visiting **www.standards.org.au**

Standards Australia welcomes suggestions for improvements, and encourages readers to notify us immediately of any apparent inaccuracies or ambiguities. Contact us via email at **mail@standards.org.au**, or write to Standards Australia, GPO Box 476, Sydney, NSW 2001.

Australian Standard[®]

Information technology—Biometric performance testing and reporting

Part 1: Principles and framework

First published as AS ISO/IEC 19795.1—2007.

COPYRIGHT

© Standards Australia

All rights are reserved. No part of this work may be reproduced or copied in any form or by any means, electronic or mechanical, including photocopying, without the written permission of the publisher.

Published by Standards Australia GPO Box 476, Sydney, NSW 2001, Australia

ISBN 0 7337 8245 0

PREFACE

This Standard was prepared by the Standards Australia Committee IT-032, Biometric and Identification.

The objective of this Standard is to make available to the Australian biometrics community the core ISO SC37 standards published over the last 24 months, specifically testing principles and framework.

This Standard is identical with, and has been reproduced from ISO/IEC 19751-1:2006, *Information technology—Biometric performance testing and reporting—Part 1: Principles and framework*.

As this Standard is reproduced from an international standard, the following applies:

- (a) Its number appears on the cover and title page while the international standard number appears only on the cover.
- (b) In the source text ‘this part of ISO/IEC 19795’ should read ‘this Australian Standard’.
- (c) A full point substitutes for a comma when referring to a decimal marker.

Reference to International Standard

Australian Standard

ISO/IEC

AS ISO/IEC

17025 General requirements for the
competence of testing and
calibration laboratories

17025 General requirements for the
competence of testing and
calibration laboratories

Only international references that have been adopted as Australian or Australian/New Zealand Standards have been listed.

The term ‘informative’ has been used in this Standard to define the application of the annex to which it applies. An ‘informative’ annex is only for information and guidance.

CONTENTS

	<i>Page</i>
1	Scope 1
2	Conformance 1
3	Normative references 1
4	Terms and definitions..... 2
4.1	Biometric data 2
4.2	User interaction with a biometric system..... 3
4.3	Personnel involved in the evaluation 3
4.4	Types of evaluation 4
4.5	Biometric applications 5
4.6	Performance measures 5
4.7	Data presentation curves 7
4.8	Statistical terms 7
5	General biometric system 8
5.1	Conceptual diagram of general biometric system 8
5.2	Conceptual components of a general biometric system..... 8
5.3	Functions of general biometric system..... 10
5.4	Enrolment, verification & identification transactions 12
5.5	Performance measure 12
6	Planning the evaluation..... 14
6.1	General..... 14
6.2	Use of other parts of ISO/IEC 19795 14
6.3	Determine information about the system..... 14
6.4	Controlling factors that influence performance 15
6.5	Test subject selection 16
6.6	Test size..... 17
6.7	Multiple tests 18
7	Data collection 19
7.1	Avoidance of data collection errors..... 19
7.2	Data and details collected..... 20
7.3	Enrolments 20
7.4	Genuine transactions 21
7.5	Identification transactions of users enrolled in the system..... 22
7.6	Impostor transactions 23
7.7	Identification transactions of users not enrolled in the system 25
8	Analyses 26
8.1	General..... 26
8.2	Fundamental performance metrics 26
8.3	Verification system performance metrics 28
8.4	(Open-set) Identification system performance metrics 30
8.5	Closed-set identification 31
8.6	Detection error trade-off / Receiver operating characteristic curves..... 31
8.7	Uncertainty of estimates 32
9	Record keeping 32
10	Reporting performance results 33
10.1	Fundamental metrics 33

10.2	Verification system metrics	33
10.3	Identification system metrics.....	33
10.4	Closed-set identification system metrics	34
10.5	Reporting test details	34
10.6	Graphical presentation of results.....	35
Annex A	(informative) Differences between evaluation types	38
Annex B	(informative) Test size and uncertainty	39
B.1	Confidence intervals and test size assuming independent identically distributed comparisons	39
B.1.1	Rule of 3	39
B.1.2	Rule of 30	39
B.1.3	Number of comparisons to support a claimed error rate	39
B.2	Variance of performance measures as a function of test size	40
B.3	Estimates for variance of performance measures.....	41
B.3.1	General	41
B.3.2	Variance of observed false non-match rate	41
B.3.3	Variance of observed false match rate	43
B.4	Estimating confidence intervals	44
B.4.1	General	44
B.4.2	Bootstrap estimates of the variance and confidence intervals.....	44
B.4.3	Subset sampling	45
Annex C	(informative) Factors influencing performance	46
C.1	General	46
C.2	List of factors.....	46
C.2.1	Population demographics	46
C.2.2	Application.....	47
C.2.3	User physiology	47
C.2.4	User behaviour	48
C.2.5	User appearance	48
C.2.6	Environmental influences	49
C.2.7	Sensor and hardware.....	49
C.2.8	User interface	50
C.3	Examples for reporting.....	50
C.3.1	Finger position	50
C.3.2	Illumination	50
C.3.3	Glasses.....	50
C.3.4	Dirt on platen	51
C.3.5	Weather	51
Annex D	(informative) Pre-selection	52
D.1	Pre-selection algorithm performance	52
Annex E	(informative) Identification performance as a function of database size	53
Annex F	(informative) Algorithms for generating ROC, DET and CMC curves	54
F.1	Algorithm for ROC and DET	54
F.2	Algorithm for generating CMC.....	54
	Bibliography	55

INTRODUCTION

This part of ISO/IEC 19795 is concerned solely with the scientific “technical performance testing” of biometric systems and devices. Technical performance testing seeks to determine error and throughput rates, with the goal of understanding and predicting the real-world error and throughput performance of biometric systems. The error rates include both false positive and false negative decisions, as well as failure-to-enrol and failure-to-acquire rates across the test population. Throughput rates refer to the number of users processed per unit time based both on computational speed and human-machine interaction. These measures are generally applicable to all biometric systems and devices. Technical performance tests that are device-specific — for example, fingerprint scanner image quality — are not considered in this part of ISO/IEC 19795.

It is acknowledged that technical performance testing is only one form of biometric testing. Other types of testing not considered in this part of ISO/IEC 19795 include

- reliability, availability and maintainability;
- security, including vulnerability;
- conformance;
- safety;
- human factors, including user acceptance;
- cost/benefit;
- privacy regulation compliance.

Methods and philosophies for these other types of test are currently being considered internationally by a broad range of groups.

The purpose of this part of ISO/IEC 19795 is to present the requirements and best scientific practices for conducting technical performance testing. This is necessary because even a short review of the technical literature on biometric device testing over the last two decades or more reveals a wide variety of conflicting and contradictory testing protocols [1-11]. Even single organizations have produced multiple tests, each using a different test method. Test protocols have varied not only because test goals and available data are different from one test to the next, but also because no standard has existed for protocol creation.

Biometric technical performance testing can be of three types: technology, scenario or operational evaluation. Each type of test requires a different protocol and produces different types of results. Even for tests of a single type, the wide variety of biometric devices, sensors, vendor instructions, data acquisition methods, target applications and populations makes it impossible to present precise uniform testing protocols. Other parts of ISO/IEC 19795 will provide specific advice and requirements for the development and use of such different test protocols. This part of ISO/IEC 19795 addresses specific philosophies and principles that can be applied over a broad range of test conditions.

This part of ISO/IEC 19795 has been developed from the UK Biometrics Working Group’s Best Practices in Testing and Reporting Performance of Biometric Devices [12] which itself drew from two primary source documents developed by the US National Institute of Standards and Technology (NIST) [13, 14], a variety of evaluation reports [7-10], and comments from the Biometrics Consortium Working Group on Interoperability, Performance and Assurance.

AUSTRALIAN STANDARD

Information technology — Biometric performance testing and reporting —

Part 1: Principles and framework

1 Scope

This part of ISO/IEC 19795

- establishes general principles for testing the performance of biometric systems in terms of error rates and throughput rates for purposes including prediction of performance, comparison of performance, and verifying compliance with specified performance requirements;
- specifies performance metrics for biometric systems;
- specifies requirements on test methods, recording of data and reporting of results; and
- provides a framework for developing and describing test protocols, to help avoid bias due to inappropriate data collection or analytic procedures, to help achieve the best estimate of field performance for the expended effort, and to improve understanding of the limits of applicability of the test results.

This part of ISO/IEC 19795 is applicable to empirical performance testing of biometric systems and algorithms through analysis of the matching scores and decisions output by the system, without detailed knowledge of the system's algorithms or of the underlying distribution of biometric characteristics in the population of interest.

Not within the scope of this part of ISO/IEC 19795 is the measurement of error and throughput rates for people deliberately trying to circumvent correct recognition by the biometric system (i.e. active impostors).

2 Conformance

To conform to this part of ISO/IEC 19795, a biometric performance test shall be planned, executed and reported in accordance with the mandatory requirements contained herein.

3 Normative references

The following referenced documents are indispensable for the application 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 17025, *General requirements for the competence of testing and calibration laboratories*

4 Terms and definitions

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

4.1 Biometric data

4.1.1

sample

user's biometric measures as output by the data capture subsystem

EXAMPLE Fingerprint image, face image and iris image are samples.

NOTE In more complex systems, the sample may consist of multiple presented characteristics (e.g., 10-print fingerprint record, face image captured from different angles, left and right iris image pair).

4.1.2

features

digital representation of the information extracted from a sample (by the signal processing subsystem) that will be used to construct or compare against enrolment templates

EXAMPLE Minutiae coordinates and principal component coefficients are features.

4.1.3

template

model

user's stored reference measure based on features extracted from enrolment samples

NOTE The reference measure is often a template comprising the biometric features for an ideal sample presented by the user. More generally, the stored reference will be a model representing the potential range of biometric features for that user. In this part of ISO/IEC 19795, we normally use "template" to include "model".

4.1.4

matching score

similarity score

measure of the similarity between features derived from a sample and a stored template, or a measure of how well these features fit a user's reference model

NOTE 1 A match or non-match decision may be made according to whether this score exceeds a decision threshold.

NOTE 2 As features derived from a presented sample become closer to the stored template, similarity scores will increase.

4.1.5

verification decision

determination of the probable validity of a user's claim to identity in the system

4.1.6

candidate list

set of potential enrolled identifiers for a subject produced by an identification attempt (or by a pre-selection algorithm)

4.1.7

identification decision

determination of a candidate list for the user's probable identity in the system

4.2 User interaction with a biometric system

4.2.1

presentation

submission of a single biometric sample on the part of a user

4.2.2

attempt

submission of one (or a sequence of) biometric samples to the system

NOTE An attempt results in an enrolment template, a matching score (or scores), or possibly a failure-to-acquire.

4.2.3

transaction

sequence of attempts on the part of a user for the purposes of an enrolment, verification or identification

NOTE There are three types of transaction: enrolment sequence, resulting in an enrolment or a failure-to-enrol; a verification sequence resulting in a verification decision; or identification sequence, resulting in an identification decision.

4.2.4

genuine attempt

single good-faith attempt by a user to match their own stored template

4.2.5

zero-effort impostor attempt

attempt in which an individual submits his/her own biometric characteristics as if he/she were attempting successful verification against his/her own template, but the comparison is made against the template of another user

4.2.6

active impostor attempt

attempt in which an individual tries to match the stored template of a different individual by presenting a simulated or reproduced biometric sample, or by intentionally modifying his/her own biometric characteristics

NOTE Error rates for active impostor attempts will vary from those for zero-effort impostor attempts. Defining the methods and skill used in active impostor attempts is outside the scope of this part of ISO/IEC 19795.

4.2.7

presentation effects

broad category of variables affecting the way in which the users' inherent biometric characteristics are displayed to the sensor

EXAMPLE In facial recognition, this could include pose angle and illumination; in fingerprinting, finger rotation and skin moisture. In many cases, the distinction between changes in the fundamental biometric characteristic and the presentation effects may not be clear (e.g. facial expression in facial recognition or pitch change in speaker verification systems).

4.2.8

channel effects

changes imposed on the presented signal in the transduction and transmission process due to the sampling, noise and frequency response characteristics of the sensor and transmission channel

4.3 Personnel involved in the evaluation

4.3.1

user

person presenting a biometric sample to the system

4.3.2**test subject**

user whose biometric data is intended to be enrolled or compared as part of the evaluation

4.3.3**crew**

set of test subjects gathered for an evaluation

4.3.4**target population**

set of users of the application for which performance is being evaluated

4.3.5**administrator**

person performing the testing or enrolment

4.3.6**operator**

individual with function in the actual system

EXAMPLE

Staff conducting enrolments or overseeing verification or identification transactions.

4.3.7**observer**

test staff member recording test data or monitoring the crew

4.3.8**experimenter**

person responsible for defining, designing and analysing the test

4.3.9**test organization**

functional entity under whose auspices the test is conducted

4.4 Types of evaluation**4.4.1****technology evaluation**

offline evaluation of one or more algorithms for the same biometric modality using a pre-existing or specially-collected corpus of samples

4.4.2**scenario evaluation**

evaluation in which the end-to-end system performance is determined in a prototype or simulated application

4.4.3**operational evaluation**

evaluation in which the performance of a complete biometric system is determined in a specific application environment with a specific target population

4.4.4**online**

pertaining to execution of enrolment and matching at the time of image or signal submission

NOTE

Online testing has the advantage that the biometric sample can be immediately discarded, saving the need for storage and for the system to operate in a manner different from usual. However, it is recommended that images or signals are collected if possible.

4.4.5**offline**

pertaining to execution of enrolment and matching separately from image or signal submission

NOTE 1 Collecting a corpus of images or signals for offline enrolment and calculation of matching scores allows greater control over which attempts and template images are to be used in any transaction.

NOTE 2 Technology testing will always involve data storage for later, offline processing. However, with scenario and operational testing, online transactions might be simpler for the tester – the system is operating in its usual manner and, although recommended, storage of images or signals is not absolutely necessary.

4.5 Biometric applications**4.5.1****verification**

application in which the user makes a positive claim to an identity, features derived from the submitted sample biometric measure are compared to the enrolled template for the claimed identity, and an accept or reject decision regarding the identity claim is returned

NOTE The claimed identity might be in the form of a name, personal identification number (PIN), swipe card, or other unique identifier provided to the system.

4.5.2**identification**

application in which a search of the enrolled database is performed, and a candidate list of 0, 1 or more identifiers is returned

4.5.3**closed-set identification**

identification for which all potential users are enrolled in the system

4.5.4**open-set identification**

identification for which some potential users are not enrolled in the system

4.6 Performance measures**4.6.1****failure-to-enrol rate****FTE**

proportion of the population for whom the system fails to complete the enrolment process

NOTE The observed failure-to-enrol rate is measured on test crew enrolments. The predicted/expected failure-to-enrol rate will apply to the entire target population.

4.6.2**failure-to-acquire rate****FTA**

proportion of verification or identification attempts for which the system fails to capture or locate an image or signal of sufficient quality

NOTE The observed failure-to-acquire rate is distinct from the predicted/expected failure-to-acquire rate (the former may be used to estimate the latter).

4.6.3**false non-match rate****FNMR**

proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same user supplying the sample

NOTE The measured/observed false non-match rate is distinct from the predicted/expected false non-match rate (the former may be used to estimate the latter).

4.6.4**false match rate****FMR**

proportion of zero-effort impostor attempt samples falsely declared to match the compared non-self template

NOTE The measured/observed false match rate is distinct from the predicted/expected false match rate (the former may be used to estimate the latter).

4.6.5**false reject rate****FRR**

proportion of verification transactions with truthful claims of identity that are incorrectly denied

4.6.6**false accept rate****FAR**

proportion of verification transactions with wrongful claims of identity that are incorrectly confirmed

4.6.7**(true-positive) identification rate**

identification rate

proportion of identification transactions by users enrolled in the system in which the user's correct identifier is among those returned

NOTE 1 This identification rate is dependent on (a) the size of the enrolment database, and (b) a decision threshold for matching scores and/or the number of matching identifiers returned.

4.6.8**false-negative identification-error rate****FNIR**

proportion of identification transactions by users enrolled in the system in which the user's correct identifier is not among those returned

NOTE False-negative identification-error rate = $1 - \text{true-positive identification rate}$.

4.6.9**false-positive identification-error rate****FPIR**

proportion of identification transactions by users not enrolled in the system, where an identifier is returned

NOTE 1 The false-positive identification-error rate is dependent on (a) the size of the enrolment database, and (b) a decision threshold for matching scores and/or the number of matching identifiers returned.

NOTE 2 With closed-set identification false-positive identification is not possible, as all users are enrolled.

4.6.10**pre-selection algorithm**

algorithm to reduce the number of templates that need to be matched in an identification search of the enrolment database

4.6.11**pre-selection error**

⟨pre-selection algorithm⟩ error that occurs when the corresponding enrolment template is not in the pre-selected subset of candidates when a sample from the same biometric characteristic on the same user is given

NOTE In binning pre-selection, pre-selection errors happen when the enrolment template and a subsequent sample from the same biometric characteristic on the same user are placed in different partitions.

4.6.12**penetration rate**

⟨pre-selection algorithm⟩ measure of the average number of pre-selected templates as a fraction of the total number of templates

4.6.13**identification rank**

smallest value k for which a user's correct identifier is in the top k identifiers returned by an identification system

NOTE Identification rank is dependent on the size of the enrolment database, and should be quoted "rank k out of n ".

4.7 Data presentation curves**4.7.1****detection error trade-off curve**

DET curve

modified ROC curve which plots error rates on both axes (false positives on the x-axis and false negatives on the y-axis)

NOTE An example set of DET curves is shown in 10.6.2, Figure 3.

4.7.2**receiver operating characteristic curve**

ROC curve

plot of the rate of false positives (i.e. impostor attempts accepted) on the x-axis against the corresponding rate of true positives (i.e. genuine attempts accepted) on the y-axis plotted parametrically as a function of the decision threshold

NOTE An example set of ROC curves is shown in 10.6.3, Figure 4.

4.7.3**cumulative match characteristic curve**

CMC curve

graphical presentation of results of an identification task test, plotting rank values on the x-axis and the probability of correct identification at or below that rank on the y-axis

NOTE An example set of CMC curves is shown in 10.6.4, Figure 5.

4.8 Statistical terms**4.8.1****variance**

V

measure of the spread of a statistical distribution

NOTE 1 If $E(X)$ represents the distribution mean of a random variable X , then $V(X) = E((X - \mu)^2)$, where $\mu = E[X]$.

NOTE 2 The variance, if known, shows how close an estimated result is likely to be to its true value.

4.8.2**confidence interval**

a lower estimate L and an upper estimate U for a parameter x such that the probability of the true value of x being between L and U is the stated value (e.g. 95 %)

EXAMPLE If $[L, U]$ is a (95 %) confidence interval for parameter x , then probability $(x \in [L, U]) = 95 \%$.

NOTE The smaller the test size, the wider the confidence interval.

5 General biometric system

5.1 Conceptual diagram of general biometric system

Given the variety of applications and technologies, it might seem difficult to draw any generalizations about biometric systems. All such systems, however, have many elements in common. Biometric samples are acquired from a subject by a sensor. The sensor output is sent to a processor which extracts the distinctive but repeatable measures of the sample (the features), discarding all other components. The resulting features can be stored in the database as a template, or compared to a specific template, many templates or all templates already in a database to determine if there is a match. A decision regarding the identity claim is made based upon the similarity between the sample features and those of the template or templates compared.

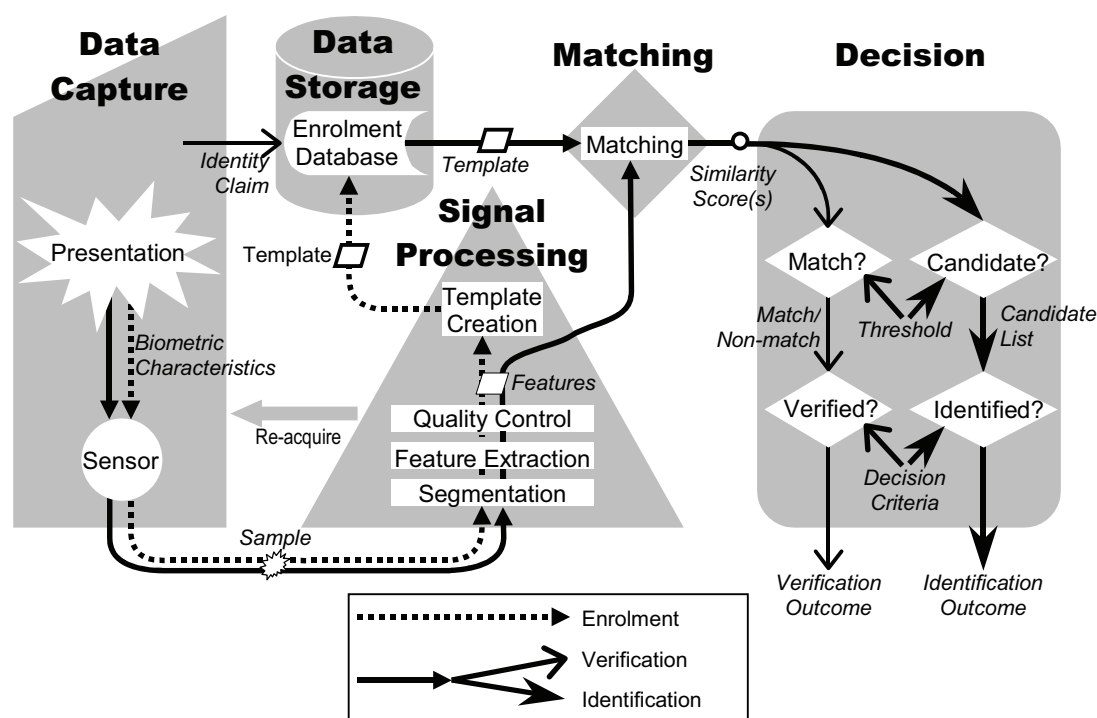


Figure 1 — Components of general biometric system

Figure 1 illustrates the information flow within a general biometric system consisting of *data capture*, *signal processing*, *storage*, *matching*, and *decision subsystems*. This diagram illustrates both enrolment, and the operation of verification and identification systems. The following subclauses describe each of these subsystems in more detail. It should be noted that, in any real biometric system, these conceptual components may not exist or may not directly correspond to the physical components, for example quality control could also take place before segmentation or before feature extraction.

5.2 Conceptual components of a general biometric system

5.2.1 Data capture subsystem

The *data capture subsystem* collects an image or signal of a subject's *biometric characteristics* that they have *presented* to the *biometric sensor*, and outputs this image/signal as a *biometric sample*.

5.2.2 Transmission subsystem (Not portrayed in diagram)

The *transmission subsystem* (not always present or visibly present in a biometric system) will transmit *samples, features, and/or templates* between different subsystems. *Samples, features or templates* may be transmitted using standard biometric data interchange formats. The biometric *sample* may be compressed and/or encrypted before transmission, and expanded and/or decrypted before use. A biometric *sample* may be altered in transmission due to noise in the transmission channel as well as losses in the compression/expansion process. It is advisable that cryptographic techniques be used to protect the authenticity, integrity, and confidentiality of stored and transmitted biometric data.

5.2.3 Signal processing subsystem

The *signal processing subsystem* extracts the distinguishing *features* from a biometric *sample*. This may involve locating the signal of the subject's *biometric characteristics* within the received *sample* (a process known as *segmentation*), *feature extraction*, and *quality control* to ensure that the extracted features are likely to be distinguishing and repeatable. Should *quality control* reject the received *sample/s*, control may return to the *data capture subsystem* to collect a further *sample/s*.

In the case of enrolment, the *signal processing subsystem* creates a *template* from the extracted biometric *features*. Often the enrolment process requires *features* from several presentations of the individual's *biometric characteristics*. Sometimes the *template* comprises just the *features*.

5.2.4 Data storage subsystem

Templates are stored within an *enrolment database* held in the *data storage subsystem*. Each *template* is associated with details of the enrolled subject. It should be noted that prior to being stored in the *enrolment database*, *templates* may be re-formatted into a biometric data interchange format. *Templates* may be stored within a biometric capture device, on a portable medium such as a smart card, locally such as on a personal computer or local server, or in a central database.

5.2.5 Matching subsystem

In the *matching subsystem*, the *features* are compared against one or more *templates* and *similarity scores* are passed to the *decision subsystem*. The *similarity scores* indicate the degree of fit between the *features* and *template/s* compared. In some cases, the *features* may take the same form as the stored *template*. For verification, a single specific claim of subject enrolment would lead to a single *similarity score*. For identification, many or all *templates* may be compared with the *features*, and output a *similarity score* for each comparison.

5.2.6 Decision subsystem

The *decision subsystem* uses the *similarity scores* generated from one or more attempts to provide the *decision outcome* for a verification or identification transaction.

In the case of verification, the *features* are considered to match a compared *template* when the *similarity score* exceeds a specified *threshold*. A claim about the subject's enrolment can then be verified on the basis of the *decision policy*, which may allow or require multiple attempts.

In the case of identification, the enrollee identifier or *template* is a potential *candidate* for the subject when the *similarity score* exceeds a specified *threshold*, and/or when the *similarity score* is among the highest k values generated for a specified value k. The *decision policy* may allow or require multiple attempts before making an identification decision.

NOTE Conceptually, it is possible to treat multi-biometric systems in the same manner as uni-biometric systems, by treating the combined biometric *samples/templates/scores* as if they were a single *sample/template/score* and allowing the *decision subsystem* to operate score fusion or decision fusion as and if appropriate.

5.2.7 Administration subsystem (Not portrayed in diagram)

The *administration subsystem* governs the overall policy, implementation and usage of the biometric system, in accordance with the relevant legal, jurisdictional and societal constraints and requirements. Illustrative examples include:

- providing feedback to the subject during and/or after data capture;
- requesting additional information from the subject;
- storage and format of the biometric *templates* and/or biometric interchange data;
- provide final arbitration on output from decision and/or scores;
- set *threshold* values;
- set biometric system acquisition settings;
- control the operational environment and non-biometric data storage;
- provide appropriate safeguards for end-user privacy;
- interact with the application that utilizes the biometric system.

5.2.8 Interface (Not portrayed in diagram)

The biometric system may or may not interface to an external application or system via an Application Programming Interface, Hardware Interface or a Protocol Interface.

5.3 Functions of general biometric system

5.3.1 Enrolment

In enrolment, a transaction by a subject is processed by the system in order to generate and store an enrolment template for that individual.

Enrolment typically involves:

- sample acquisition,
- segmentation and feature extraction,
- quality checks, (which may reject the sample/features as being unsuitable for creating a template, and require acquisition of further samples),
- template creation (which may require features from multiple samples), possible conversion into a biometric data interchange format and storage,
- test verification or identification attempts to ensure that the resulting enrolment is usable,
- and should the initial enrolment be deemed unsatisfactory, repeat enrolment attempts may be allowed (dependent on the enrolment policy).

5.3.2 Verification

In verification, a transaction by a subject is processed by the system in order to verify a positive specific claim about the subject's enrolment (e.g. "I am enrolled as subject X"). Verification will either accept or reject the

claim. The verification decision outcome is considered to be erroneous if either a false claim is accepted (false accept) or a true claim is rejected (false reject). Note that some biometric systems will allow a single end-user to enrol more than one instance of a biometric characteristic (for example, an iris system may allow end-users to enrol both iris images, while a fingerprint system may have end-users enrol two or more fingers as backup, in case one finger gets damaged).

Verification typically involves:

- sample acquisition,
- segmentation and feature extraction,
- quality checks, (which may reject the sample/features as being unsuitable for comparison, and require acquisition of further samples),
- comparison of the sample features against the template for the claimed identity producing a similarity score,
- judgement on whether the sample features match the template based on whether the similarity score exceeds a threshold, and
- a verification decision based on the match result of one or more attempts as dictated by the decision policy.

EXAMPLE In a verification system allowing up to three attempts to be matched to an enrolled template, a false rejection will result with any combination of failures-to-acquire and false non-matches over three attempts. A false acceptance will result if a sample is acquired and falsely matched to the enrolled template for the claimed identity on any of three attempts.

5.3.3 Identification

In identification, a transaction by a subject is processed by the system in order to find an identifier of the subject's enrolment. Identification provides a candidate list of identifiers that may be empty or contain only one identifier. Identification is considered correct when the subject is enrolled, and an identifier for their enrolment is in the candidate list. The identification is considered to be erroneous if either an enrolled subject's identifier is not in the resulting candidate list (false-negative identification error), or if a transaction by a non-enrolled subject produces a non-empty candidate list (false-positive identification error).

Identification typically involves:

- sample acquisition,
- segmentation and feature extraction,
- quality checks, (which may reject the sample/features as being unsuitable for comparison, and require acquisition of further samples),
- comparison against some or all templates in the enrolment database, producing a similarity score for each comparison,
- judgement on whether each matched template is a potential candidate identifier for the user, based on whether the similarity score exceeds a threshold and/or is among the highest k scores returned, producing a candidate list,
- an identification decision based on the candidate lists from one or more attempts, as dictated by the decision policy.

NOTE 1 In fully automated systems, the assigned identifier might be that corresponding to the template giving the highest similarity score (providing that this exceeds the specified threshold). When there is a human operator, the system may show a candidate list of the top *r* matches, for a final decision by the operator. Determining the ultimate performance metrics for systems using examination of possible matches by human operators is beyond the scope of this document.

NOTE 2 If all subjects using an identification system are known to be enrolled in the system, there can never be false-positive identification errors. In this degenerate case, known as closed-set identification, performance evaluation interest is normally focussed on how the rate of correct identification is related to the size of the candidate list returned.

5.4 Enrolment, verification & identification transactions

Each of the above biometric functions will depend on a user transaction. The transaction will consist of one or more attempts as allowed or required by the corresponding decision policy. For example, the decision policy may allow three attempts to verify; then the transaction may consist of one attempt, two attempts if the first attempt is rejected, or three attempts if the first two attempts are rejected.

Each attempt may consist of one or more presentations dependent on sensor operation, policy on sample quality, and any settings limiting the number of presentations or time permitted per attempt. For example an enrolment attempt may require the biometric sample to be submitted more than one time. Biometric verification systems often process a sequence of samples in a single attempt, for example: (a) collecting samples over some fixed period to find the best matching sample; (b) collecting samples until either a match is obtained or the system times out; or (c) collecting samples until one of sufficient quality is obtained, or the system times out.

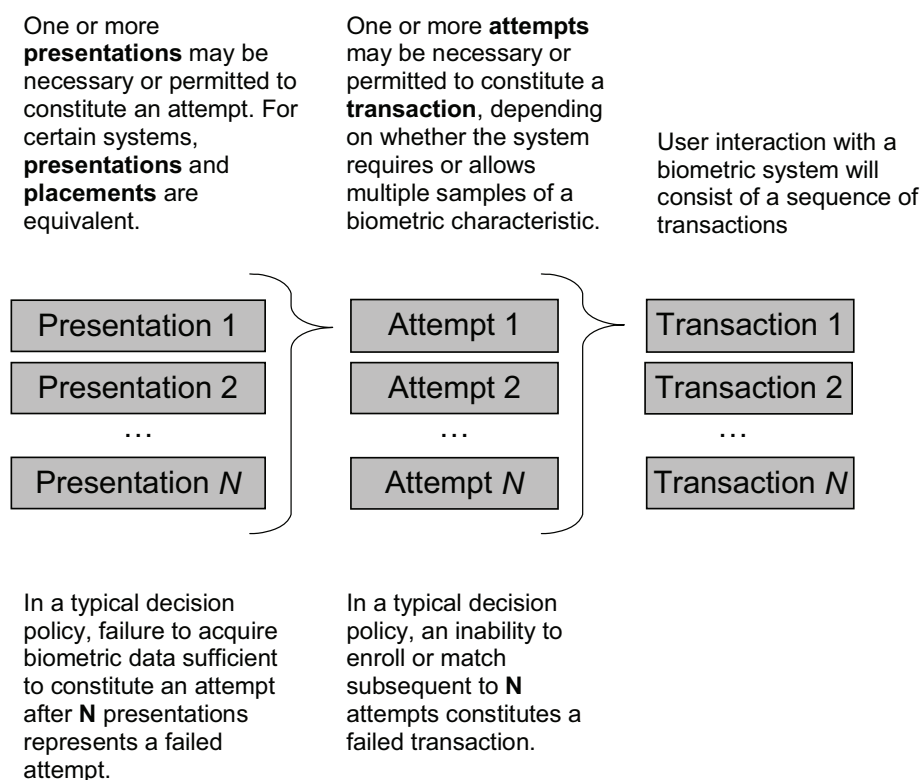


Figure 2 — Presentations, attempts and transactions

5.5 Performance measure

5.5.1 Error rates

Verification and identification decision errors are due to matching errors (i.e. false match and false non-match errors) or sample acquisition errors (i.e. failures-to-enrol and failures-to-acquire). How these fundamental errors combine to form decision errors will depend upon the number of comparisons required; whether there is a positive or negative claim of identity; and the decision policy, e.g. whether the system allows multiple attempts.

NOTE Though biometric performance has traditionally been stated in terms of the decision error rates, i.e. false accept rate and false reject rate, conflicting definitions are implicit in the literature. Literature on large-scale identification systems often refers to a “false rejection” occurring when a submitted sample is incorrectly matched to a template enrolled by another user. In access control literature, a “false acceptance” is said to have occurred when a submitted sample is incorrectly matched to a template enrolled by another user. False match rate and false non-match rate are not generally synonymous with false accept rate and false reject rate. False match or non-match rates are calculated over the number of comparisons, but false accept or reject rates are calculated over transactions and refer to the acceptance or rejection of the stated hypothesis, whether positive or negative. Furthermore, false accept or reject rates are inclusive of failures-to-acquire.

5.5.2 Throughput rates

Throughput rates show the number of users that can be processed per unit time based both on computational speed and human-machine interaction. These measures are generally applicable to all biometric systems and devices. Attaining adequate throughput rates is critical to the success of any biometric system. Throughput rates for verification systems, such as those for access control, are usually controlled by the speed of user interaction with the system in the process of submitting a good quality biometric sample. Throughput rates for identification systems, such as enrolment in a social service program, can be heavily impacted by the computer processing time required to compare the enrolment sample to the database of stored templates. So, depending upon the type of system, it may be appropriate to measure the interaction times of users with the system and also the processing rate of the computational hardware. Actual benchmark measurement of computer processing speed is covered in texts such as [12] and is considered outside the scope of this document. Measurement of the speed of the human-machine interaction will require a precise definition of the action which indicates the initiation of the interaction and the action that terminates the interaction. This definition should be determined prior to the start of the test and noted in the test report. The test report should also include a brief listing of the actions of the user included in the human-machine interaction.

5.5.3 Types of performance evaluation

Testing a biometric system will involve the collection of input images or signals, which are used for template generation at enrolment and for calculation of matching scores for verification or identification attempts. The images/signals collected can either be used immediately for an online enrolment, verification or identification attempt, or may be stored and used later for offline enrolment, verification or identification.

- a) In a technology evaluation, testing of all algorithms is carried out on a standardized corpus, ideally collected by a “universal” sensor (i.e. a sensor that collects samples equally suitable for all algorithms tested). Nonetheless, performance against this corpus will depend on both the environment and the population in which it is collected. Although example data may be distributed for developmental or tuning purposes prior to the test, the actual testing needs to be done on data that has not previously been seen by algorithm developers. Testing is carried out using offline processing of the data. Because the corpus is fixed, the results of technology tests are repeatable.
- b) In a scenario evaluation, testing is carried out on a complete system in an environment that models a real-world target application of interest. Each tested system will have its own acquisition sensor and so will receive slightly different data. Consequently, if multiple systems are being compared, care will be required that data collection across all tested systems is in the same environment with the same population. Depending on the data storage capabilities of each device, testing might be a combination of offline and online comparisons. Test results will be repeatable only to the extent that the modelled scenario can be carefully controlled.
- c) In an operational evaluation, depending on the data storage capabilities of the operational system, offline testing might not be possible. In general, operational test results will not be repeatable because of unknown and undocumented differences between operational environments. Furthermore, “ground truth” (i.e. who was actually presenting a “good faith” biometric measure) can be difficult to ascertain, particularly if an operational evaluation is performed under unsupervised conditions without an administrator, operator or observer present.

Annex A summarizes the different characteristics of the different types of evaluations.