

Project no.: IST-2005-027351
 Project acronym: MTIT
 Project title: Minutiae template interoperability testing
 Instrument: Specific targeted research project
 Thematic Priority: Information Society Technologies
 Start date of project: November 1, 2005
 Duration: 14 months

Deliverable D 3.2:

Specification of a test and certification protocol

Version 1.1

Due date of deliverable: February 28, 2006
 Actual submission date: July 30, 2006
 Lead contractor for this deliverable: NPL Management Ltd

Project co-funded by the European Commission within the Sixth Framework Programme (2002–2006)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

1	Purpose of the document	1
2	Objectives of the testing effort.....	1
2.1	Objectives	1
2.2	Phases of testing	2
2.3	Investigations.....	2
2.4	Certification	2
3	Interfaces of the subsystems under test.....	3
3.1	Finger minutiae data interchange formats	3
3.2	Calling interface	6
3.3	Finger image data interchange format	8
4	Finger image databases.....	8
4.1	Acquisition.....	8
4.2	Corpus size vs. confidence levels	9
4.3	Requirements on metadata.....	10
4.4	Format of the collected finger image databases	10
5	Test preparation	11
5.1	Overview	11
5.2	Conversion of finger image databases.....	11
5.3	Assessment of finger image quality.....	12
5.4	Partitioning of finger image databases	12
6	Conformance testing	13
6.1	Conformance testing of minutiae extraction subsystems	13
6.2	Conformance testing of a minutiae comparison subsystem	14
7	Interoperability testing	15
7.1	Interoperability tests to be carried out	15
7.2	Overview of single interoperability test	16
7.3	Generation of minutiae template databases	17
7.4	Minutiae comparisons.....	17
7.5	Computation of single-supplier and cross-supplier performance metrics	18
7.6	Determination of maximum classes of interoperable subsystems.....	18
8	Investigations	18
	References	20

1 Purpose of the document

This report test plan specifies, the MTIT project test plan for investigation, measurement and certification of interoperability of finger minutiae extraction and comparison subsystems conforming to [ISO/IEC 19794-2].

The test plan is based on:

1. the criteria of interoperability specified in [MTIT D3.1]
2. test standards and drafts [ISO/IEC 19795-1, ISO/IEC 19795-2] and particularly [ISO/IEC 19795-4] for interoperability testing of data interchange formats.

The required test bed functionality [MTIT D4.1] is extracted from this test plan.

The report is organised as follows:

- Section 2 states the objectives of this testing effort.
- Section 3 discusses what interfaces the subsystems under test shall provide and which data interchange formats shall be used.
- Section 4 introduces the test databases.
- Section 5 deals with preparatory steps before test execution.
- Section 6 specifies how minutiae extraction and comparison subsystems are tested for conformance to [ISO/IEC 19794-2].
- Section 7 specifies how the interoperability of finger minutiae extraction and comparison subsystems is tested.
- Section 8 describes the analyses to be performed in the investigation phase of testing. These investigations may cause modifications to the precise test specification for the certification phase of testing.

2 Objectives of the testing effort

2.1 Objectives

The objectives of the described testing effort are

- to assess the conformance of minutiae extraction subsystems and minutiae comparison subsystems to [ISO/IEC 19794-2],

- to assess the cross-supplier data-format interoperability of minutiae extraction subsystems and minutiae comparison subsystems interchanging minutiae records conforming to [ISO/IEC 19794-2] (minutiae-minutiae comparison scenario),
- to assess the cross-supplier data-format interoperability of minutiae extraction subsystems with comparison subsystems that compare minutiae records conforming to [ISO/IEC 19794-2] with proprietary finger templates (minutiae-image comparison scenario).

2.2 Phases of testing

There will be two phases of testing. The first phase will be an investigation phase, based on the outline interoperability criteria (in [MTIT D3.1]). This phase will assess what level of performance is achieved in interoperating systems and will investigate the feasibility of the approaches to interoperability criteria and the approaches to interoperability assessment.

After the initial test campaign, any interoperability issues that may have arisen shall be analysed, and the suppliers of the subsystems under test shall have the opportunity to mend the subsystems under test. After mending the subsystems under test, a second test campaign shall be conducted.

The second phase will be a certification phase, based on an instantiation of the interoperability criteria. This phase will use a revised performance interoperability assessment and can be expected to validate the approach to interoperability assessment.

The tests will be replicated for two of the format types defined in [ISO/IEC 19794-2] and comparing data from one finger and from two fingers at a time.

2.3 Investigations

- Compare the performance obtainable in the minutiae-image comparison scenario against the performance obtainable in the minutiae-minutiae comparison scenario (MINEX scenario 1)
- Compare MTIT results against MINEX/ILO results
- Relate minutiae extraction discrepancies to performance discrepancies
- Investigate statistical models for interoperable performance

2.4 Certification

Use data set such that there is 95% confidence of interoperable performance; e.g. better than an FNMR of 1% at an FMR of 0.1%. An appropriate performance level must be set following the investigation phase of the test, when the level of difficulty of the collected test data sets is clearer.

3 Interfaces of the subsystems under test

3.1 Finger minutiae data interchange formats

3.1.1 Format types, options, and parameters defined in [ISO/IEC 19794-2]

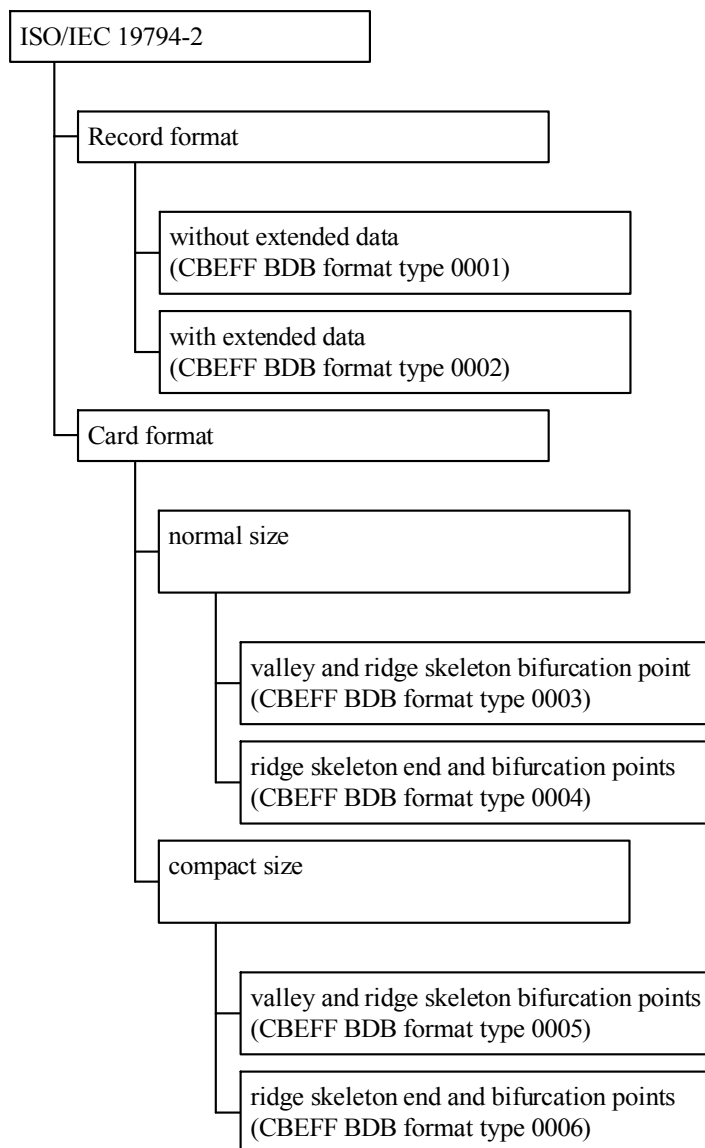


Figure 1 Format types according to [ISO/IEC 19794-2]

[ISO/IEC 19794-2] defines six different minutiae data format types, each of which is labelled with an individual CBEFF BDB format type identifier (see Figure 1): two record formats, one without and the other with extended data in a standardised form, two normal-size card formats, and two compact-size card formats. In the card formats, ridge endings can be represented either by valley skeleton bifurcation points or by ridge skeleton end points. In the record formats, ridge endings are always represented by valley skeleton bifurcation points. In both the record formats and the normal-size card formats, the x and y coordinates of each minutia are expressed in 14 bits and minutia angles are expressed in 8 bits. In the compact-

size card formats, the x and y coordinates of each minutia are expressed in 8 bits only and minutia angles are expressed in 6 bits. In the record formats, the resolutions of the x and y coordinates are adjustable, but must be at least 99 pixels per centimetre (250 pixels per inch). In the normal-size card formats, the resolutions of the x and y coordinates are fixed to 1000 pixels per centimetre, in the compact-size card formats to 100 pixels per centimetre. All formats provide further format options and parameters.

In a minutiae record in record format, a capture device type identifier may or may not be reported in the record header (anyway, the capture device type identifier is vendor-specified and, hence, useless in a multi-vendor system). A minutiae record in a record format may contain up to 16 views of each of 10 fingers. In a minutiae record in record format, for each individual minutia a minutia quality score may or may not be reported using a supplier-specific algorithm.

In any of the formats, except for the record format without extended data, a minutiae record may contain optional extended data, i.e. one or more of the following sorts of data:

- ridge count data, extracted either by
 - a non-specific ridge count extraction method,
 - the four-neighbour ridge count extraction method, or
 - the eight-neighbour ridge count extraction method,
- core data,
- delta data,
- zonal quality data,
- vendor-defined extended data.

In the card formats, the minutiae can either be not particularly ordered or be ordered by

- ascending x-y coordinates,
- descending x-y coordinates,
- ascending y-x coordinates,
- descending y-x coordinates,
- ascending angle,
- descending angle,
- ascending polar coordinates,
- descending polar coordinates, or
- x or y coordinate extension.

3.1.2 Chosen format types, options, and parameter values

The interoperability tests need to be restricted to the formats that are supported by more than one supplier.

Interoperability is especially interesting with respect to the card formats because these formats are intended for use in environments urgently requiring cross-supplier data format interoperability (like electronic ID cards).

It can be expected that the interoperability problems will be similar for similar formats and will probably not depend on the choice between ridge skeleton end points and valley skeleton bifurcations (as long as all implementations under test support the same choice) or on the ordering of the minutiae. The coding of a minutia in 3 or 5 bytes (compact or normal size), however, may have an influence on the interoperability.

Therefore, the following format profiles are chosen for data interchange between minutiae extraction and comparison subsystems:

1. 1-finger record format profile

- record format without extended data
- capture device type identifier not reported
- number of finger views: 1
- minutiae quality reported using supplier-specific algorithm
- resolution of minutiae coordinates: the same as the resolution of the underlying finger image
- not embedded in a CBEFF patron format

2. 1-finger card format profile

- compact-size card format using valley and ridge skeleton bifurcation points as minutiae
- minutiae not particularly ordered
- no extended data
- not embedded in a CBEFF patron format
- 1 finger

The main differences between the chosen format profiles are listed in Table 1. Since the fingerprints of a person may vary slightly from print to print and fingerprints must be accepted anyway within a certain range of tolerance, it cannot be clearly predicted whether the lower resolutions and the smaller minutiae window in the compact-size card format would

Table 1 Differences between the chosen format profiles

	1-finger record format profile	1-finger card format profile
Resolution of x coordinates, in pixels per cm	196.85 (500 dpi)	100
Resolution of y coordinates, in pixels per cm	196.85 (500 dpi)	100
Maximum range of x coordinates, in cm	0–83.16	0–2.55
Maximum range of y coordinates, in cm	0–83.16	0–2.55
Resolution of minutia angles, in elements per 90°	64	16
Minutia quality	reported using supplier-specific algorithm	not reported

lead to a performance loss compared to the performance achievable using the minutiae record format or not.

3.2 Calling interface

3.2.1 Constraints by the biometric performance testing toolkit

The commercial off-the-shelf biometric performance testing product DaonAnalytics [Daon05] will be applied for performing the interoperability tests. Minutiae extraction and minutiae comparison will be performed in loose succession:

- First, the minutiae extraction subsystems under test will be invoked to generate all needed minutiae templates from a finger image database. The generated minutiae templates will be stored in minutiae template databases.
- Afterwards, the minutiae comparison subsystems under test will be invoked with minutiae templates from the minutiae template databases as input.

There are several interface options for integrating the supplier-specific biometric subsystems into the DaonAnalytics toolkit:

- External executable interface (based on the submission requirements for the fingerprint verification competition 2000 [FVC00]): The suppliers of the subsystems under test have to provide Win32 console applications. This interface is simple, but adds a significant delay overhead by creating a new process for every biometric comparison.
- BioAPI 1.1 [BioAPI01]: The suppliers of the subsystems under test have to provide Biometric Service Provider (BSP) components, which are loaded and called through the BioAPI framework.
- Daon snap-in interface (designed to overcome some of the limitations of the external executable and BioAPI interfaces [Daon05]): The suppliers of the subsystems under test have to provide a Win32 dynamic link library (DLL).
- Proprietary vendor SDK integration: The suppliers of the subsystems under test provide their proprietary biometric software development kits (SDKs), and Daon, the vendor of the test toolkit, has to make efforts to integrate these proprietary SDKs into DaonAnalytics. Because it takes time to become familiar with each independent vendor SDK and to develop code to call the appropriate SDK functions from DaonAnalytics, it is more efficient and cost-effective to use one of the other interface options.

3.2.2 Chosen calling interface

Because a large number of subsystem invocations are required, an interface option that yields an efficient implementation needs to be chosen. The interface option chosen for minutiae

template interoperability testing in the MTIT project is an extended version of the Daon snap-in interface:

1. For minutiae extraction subsystems to be tested for interoperability with other subsystems using the 1-finger record format, suppliers are required to submit a Win32 DLL including a template generation function that takes a single finger image as input, extracts minutiae from it, and generates a minutiae template in the 1-finger record format.
2. For minutiae extraction subsystems to be tested for interoperability with other subsystems using the 1-finger card format, suppliers are required to submit a Win32 DLL including a template generation function that takes a single finger image as input, extracts minutiae from it, and generates a minutiae template in the 1-finger card format.
3. For proprietary extraction subsystems to be tested with other subsystems using the 1-finger record format or 1-finger card format, suppliers are required to submit a Win32 DLL including a finger template generation function that takes a single finger image as input, extracts information from it, and generates a proprietary finger template.
4. For minutiae comparison subsystems to be tested for interoperability with other subsystems using the 1-finger record format, suppliers are required to submit a Win32 DLL including a comparison function that takes one or more pairs of minutiae templates in 1-finger record format as input, compares each pair of minutiae templates, and returns a single similarity score.
5. For minutiae comparison subsystems to be tested for interoperability with other subsystems using the 1-finger card format, suppliers are required to submit a Win32 DLL including a comparison function that takes one or more pairs of minutiae templates in 1-finger card format as input, compares each pair of minutiae templates, and returns a single similarity score.
6. For comparison subsystems comparing minutiae templates in 1-finger record format with proprietary finger templates, suppliers are required to submit a Win32 DLL including a comparison function that takes one or more pairs of templates as input, compares each pair of templates, and returns a single similarity score. Each pair of templates consists of a minutiae template in 1-finger record format and a proprietary finger template.
7. For comparison subsystems comparing minutiae templates in 1-finger card format with proprietary finger templates, suppliers are required to submit a Win32 DLL including a comparison function that takes one or more pairs of templates as input, compares each pair of templates, and returns a single similarity score. Each pair of templates consists of a minutiae template in 1-finger card format and a proprietary finger template.

Where multiple pairs of minutiae templates are to be compared in a single call, vendors are expected to apply their own fusion techniques, within the comparison function, to derive a single similarity score.

The existing DaonAnalytics snap-in interface will be extended to meet the above interface requirements, such as providing the ability to pass multiple separate pairs of template to the comparison subsystem, rather than using a single consolidated record format to do this.

The functions in the submitted DLLs must use the C calling convention. There will be one set of functions defined for generating finger templates, including both standard and proprietary formats. There will be a second set of functions defined for comparing one or more pairs of these finger templates. Since the same function interface is re-used, despite the format or quantity of the finger data being passed to that function, suppliers will need to provide a separate DLL for each of the above seven subsystem scenarios. Further technical details of the interface and DLL naming conventions are specified in [MTIT D4.1].

3.3 Finger image data interchange format

3.3.1 Formats to chose from

A pure image data format, a format option of [ANSI/NIST-ITL 1-2000], or a format option of [ISO/IEC 19794-4] could be chosen.

3.3.2 Chosen format

Uncompressed greyscale finger images have been chosen as input to the minutiae extraction subsystems. This removes the need for each individual supplier subsystem or other subsystems within the testbed, such as quality scoring algorithms, from having to parse a complex image record format. Conversion from the collected image databases format (detailed in Section 4.4) to an uncompressed greyscale finger image will be performed once by the biometric performance testing product.

4 Finger image databases

4.1 Acquisition

Within the MTIT project, two finger image databases are collected for offline processing: One finger image database is collected using a LS Check sensor from Smiths Heimann Biometrics¹ with a large capture platen for the simultaneous capture of four flat fingerprints or two thumbprints; the other finger image database is collected using a Verifier 300 2.0 single-fingerprint sensor from Cross Match Technologies.

Even though for unbiased technology tests, the suppliers of subsystems under test are supposed not to have any opportunity to tune their subsystems to the sample corpus [ISO/IEC

¹ merged with Cross Match Technologies as of August 1, 2005

19795-1], in order to achieve a sufficient corpus size, each of the vendors who provide the initial set of subsystems under test contributes also a finger image database containing 2000 pairs of fingerprint sets (4000 records). Because finger image data will be contributed by several vendors, the test is expected not to be biased in favour of any of the contributing vendors.

To sum up, six finger image databases are used for testing within the MTIT project: For more details of the finger image databases see [MTIT D2.1]. Each of the finger image databases consists of a corpus of finger images and associated metadata describing these data.

LS2 Check finger image database (dataset A):

500 subjects: 10 fingerprints x 2 visits (fingerprints collected on Smiths Heimann, segmented into individual prints)

Verifier 300 2.0 finger image database:

500 subjects: 2 fingerprints (left and right index fingers) with mates in Dataset A

Sagem, Cogent, Motorola, and NEC finger image databases

each 2000 subjects, 10 fingerprints segmented into individual prints, with mates.

The datasets will be divided uniformly into two, the first part to be used in the investigative phase of testing, the second part sequestered for the certification phase of testing.

To identify interoperability issues that may be due to differing conditions under which the finger images in different databases are captured, minutiae templates extracted from finger images from different databases may also be compared with each other.

4.2 Corpus size vs. confidence levels

The corpus must be large enough to allow statistically significant tests. The lower the magnitude of the error rate to be verified, the higher the desired statistic confidence level, and the closer the observed error rate is to the error rate to be verified, the larger a corpus is required.

An observed error rate \bar{p} supports the working hypothesis that the corresponding honest error rate is below a chosen threshold p if and only if the corresponding null hypothesis, that the corresponding honest error rate is equal or above the threshold, is to be rejected. Provided the observed comparison attempts are mutually independent, it can be decided by a one-sided one-sample z test whether or not the null hypothesis is to be rejected: If

$$z = \frac{p - \bar{p}}{\sqrt{\frac{p(1-p)}{n}}}, \text{ with } n \text{ being the total number of observations, } np > 10, \text{ and } n(1-p) > 10,$$

is greater than the tabulated value of $z_{0.05} = 1.6449$, then the null hypothesis is to be rejected at the 95% confidence level, and the working hypothesis is to be accepted. Otherwise, there is not enough evidence to reject the null hypothesis, and the working hypothesis cannot be accepted.

For instance, the ILO criterion of interoperability (an FRR of 1% or less at an FAR of 1%, i.e., an EER below 1% [ILO04]) could be verified at the 95% confidence level if in 2000 independent comparisons not more than 12 errors were observed. For verifying at the 95% confidence level that an error rate does not exceed 0.1%, only 12 errors would be admissible in 20000 independent comparisons.

In case of cross-comparisons, the assumption of mutually independent comparisons does not apply, and the statistical hypothesis test will be applied with care. The real impact for MTIT, where cross-comparisons are performed, will have to be investigated.

The required corpus size for MTIT, taken from the MTIT databases (8500 pairs of 10 fingerprints), and the predicted statistical confidence will be determined after the initial (investigative) phase of testing.

4.3 Requirements on metadata

For identifying the fingerprints, the following data elements are needed:

- person identifier,
- finger position,
- sensor type, and
- date and time of capturing.

For deciding whether the collected data is representative, the following information about the persons presenting the fingerprint is needed:

- date and place of birth,
- gender, and
- occupation.

To diagnose interoperability issues, we may also want to know the capture device (and ideally capture environment), as some issues may be observed more with some device or supplier sets than with others.

4.4 Format of the collected finger image databases

The collected data are stored in the format supported by the finger image capturing software from Smiths Heimann Biometrics, which is Type 4 defined in [ANSI/NIST-ITL 1-2000] for high-resolution greyscale fingerprint image records, augmented by several fields from Type-1 transaction information records and from Type-2 user-defined descriptive text records, and with WSQ compressed finger images. The image resolution is the minimum resolution for Type-4 images, i.e. 500 pixels per inch. The Type-4 record contains the following additional information beside the finger image:

- record length in bytes (LEN),
- image designation character (IDC),
- impression type (IMP),
- finger position (FGP),
- image scanning resolution (ISR),
- image dimensions (horizontal line length, HLL, and vertical line length, VLL), and
- greyscale compression algorithm (GCA).

The Type-1 and Type-2 fields with the following numbers are stored along with the Type-4 records [MTIT D2.1]:

- 1.008: originating agency identifier (ORI),
- 2.010: person identifier,
- 2.019: date of capturing,
- 2.020: time of capturing,
- 2.022: place of capturing,
- 2.035: date of birth,
- 2.036: date of birth range,
- 2.037: place of birth,
- 2.039: gender,
- 2.049: occupation, and
- 2.067: sensor type and version.

5 Test preparation

5.1 Overview

Before the subsystems under test can be tested, the following steps shall be carried out:

- to convert the collected finger image databases into the format required as input by the biometric performance testing toolkit,
- to assess, and record, the quality of each finger image, and
- to partition the finger image databases into reference finger image databases and verification finger image databases, into representative and difficult databases, and into sequestered and unsequestered databases.

5.2 Conversion of finger image databases

The biometric performance testing toolkit DaonAnalytics takes biometric data, each piece of which is typically stored in a single file and may have any format, as well as an XML description of that data as input.

The collected finger image databases have to be converted to the format expected by the biometric performance testing toolkit. The decompressed finger images with XML description are stored as separate databases. With the manageable amount of data being captured and

processed within the project, storing the finger image databases in two different formats is considered viable.

Once the initial conversion has been performed, a once-off process, the original collected finger image database of NIST records is not required in order to perform the interoperability testing. Where storage space is an issue, these original records may be archived on long-term storage media.

5.3 Assessment of finger image quality

A non-proprietary assessment of finger image quality will be assessed using the NIST Fingerprint Image Quality software (NFIQ). The NIST NFIQ algorithm is applied to each individual finger image, recording the quality score. The test tools should allow filtering of image, template, and match score datasets, based on the quality score obtained for the original image.

5.4 Partitioning of finger image databases

5.4.1 Partitions

There will be different partitions of the finger image databases used for the two testing phases: investigation and certification. The method for partitioning each database will try to ensure that the partitions are of equivalent difficulty for testing performance and interoperability: The NFIQ scores for the fingerprints in each partition will be calculated, and we shall try to ensure that both partitions have a similar distribution of NFIQ scores.

5.4.2 Subcorpus for in-depth conformance testing

The subcorpus used for in-depth conformance testing shall contain at least 10 pairs of finger images from the LS2 Check finger image database and at least 10 pairs of finger images from the Verifier 300 2.0 finger image database (altogether at least 40 finger images). Each pair shall consist of two finger images of the same finger of the same person. Among them, there shall not be finger images with low quality score, where even manual minutiae detection would be highly ambiguous.

5.4.3 Subcorpus of difficult data

Difficult data are finger images from which different minutiae extraction subsystems are likely to extract non-matching minutiae records.

As a first approximation, the subcorpus of difficult data shall contain finger images showing a large disparity in the minutiae records extracted.

6 Conformance testing

6.1 Conformance testing of minutiae extraction subsystems

6.1.1 Overview

In a multi-supplier system, conformance of the minutiae extraction and comparison subsystems to a common format is a prerequisite, though not a guarantee, for interoperability. Therefore, the conformance of the individual minutiae extraction and comparison subsystems to [ISO/IEC 19794-2] is tested prior to interoperability testing.

Testing a minutiae extraction subsystem for conformance to the 1-finger record format profile or to the 1-finger card format profile (see Section 3.1.2), respectively, shall include the following steps:

- to automatically check the format and internal consistency of each minutiae record that is generated using the minutiae extraction subsystem from the finger images within the corpus and
- to manually check the contents of the minutiae records that are generated using the minutiae extraction subsystem from the finger images within the subcorpus for in-depth conformance testing.

The conformance test of a minutiae extraction subsystem is successful if and only if all minutiae templates that this subsystem generates from the finger image corpus conform to the 1-finger record format or the 1-finger card format, respectively.

6.1.2 Automated minutiae record conformance testing

The test tool for automated minutiae record conformance testing shall be able

- to take a minutiae record in the 1-finger record format profile or the 1-finger card format profile as input and
- to check the format and the internal consistency of the minutiae record.

The automated conformance test of a minutiae record is successful if and only if this minutiae record satisfies the relevant syntactic requirements specified in [ISO/IEC 19794-2].

6.1.3 In-depth minutiae record conformance testing

The test tool for in-depth minutiae record conformance testing shall be able

- to take a minutiae record in the 1-finger record format profile or the 1-finger card format profile as well as the finger image that underlies the minutiae record as input,

- to check the format and the internal consistency of the minutiae record,
- to graphically superimpose the scaled contents of the minutiae record onto the corresponding finger image to aid the test operator in contents checking, and
- to graphically superimpose the contents of the minutiae record under test onto the contents of a reference minutiae record generated from the same finger image (similar to the FBI's Universal Latent Workstation).

For lack of a universally accepted reference algorithm for generating minutiae records from finger images, minutiae records generated using other minutiae extraction subsystems will serve as reference records. Also for that reason, the test operator will carry out contents checking as a manual process. If

- a minutiae extraction algorithm were universally accepted as a reference algorithm or
- finger images annotated with undisputed reference minutiae were available,

checking the contents of minutiae records could also be automated.

The in-depth conformance test of a minutiae record is successful if and only if this minutiae record satisfies the relevant syntactic and semantic requirements specified in [ISO/IEC 19794-2].

6.2 Conformance testing of a minutiae comparison subsystem

6.2.1 Database of conforming minutiae templates

Before testing a minutiae comparison subsystem for conformance to the 1-finger record format profile or to the 1-finger card format profile, respectively, a database of minutiae templates conforming to the 1-finger record format or the 1-finger card format, respectively, shall be created.

The database of minutiae templates conforming to the 1-finger record format or the 1-finger card format, respectively, shall contain a minutiae template for each finger image of the sub-corpus of finger images set aside for in-depth conformance testing. Additionally, the database shall contain minutiae templates not corresponding to real fingerprints, e.g. a minutiae template containing 0 minutiae, in order to test special cases.

6.2.2 Minutiae-minutiae comparison testing

A minutiae-minutiae comparison subsystem is tested for

- conformance to the 1-finger record format profile or to the 1-finger card format profile, respectively, and

- satisfaction of the requirement to return a similarity score expressing the degree of resemblance of the compared minutiae records

by reading in and comparing all possible pairs of minutiae templates from the database of conforming minutiae templates, one or more pairs at a time.

The test of a minutiae-minutiae comparison subsystem is successful if and only if

- the subsystem under test accepts all minutiae templates from the database of minutiae templates conforming to the 1-finger record format or the 1-finger card format, respectively,
- comparisons of identical minutiae templates yield the highest similarity score, and
- comparisons of non-matching minutiae templates yield low similarity scores.

7 Interoperability testing

7.1 Interoperability tests to be carried out

The interoperability testing is carried out offline as a set of technology tests of for each combination of subsystems from different suppliers. Among the pairings (or triples) of systems, are the cases where all subsystems are from the same supplier. Such cases provide a benchmark for assessing performance degradations due to possible interoperability problems.

The interoperability test will be repeated over the different application scenarios being considered, and different databases being used.

Data format options

- Proprietary reference and verification records:
- Standard minutiae template, proprietary verification record
- Standard minutiae template, standard verification minutiae record

- Record format for minutiae records
- Compact card format for minutiae records

Number of fingerprints

- Single index finger
- 2 index fingers

Fingerprint image corpuses:

- MTIT-1 reference and verification images from Smiths Heimann sensor
- MTIT-2 reference images from Smiths Heimann sensor, verification images from Cross-match sensor.
- Sagem corpus
- Motorola corpus
- Cogent corpus

- NEC corpus

Filtering of data:

- All images
- Excluding low quality fingerprint images

7.2 Overview of single interoperability test

Figure 2 depicts the test architecture for the minutiae-minutiae comparison scenario where the algorithms for minutiae extraction of the recognition fingerprint and for minutiae comparison are provided by the same supplier. The test architecture for the minutiae-image comparison scenario is similar, though proprietary fingerprint records are generated from the verification image database. The general case, where the minutiae comparison system and the minutiae extraction subsystems for the reference template and verification minutiae record can be from three different suppliers is also similar (though is 3- rather than 2-dimensional).

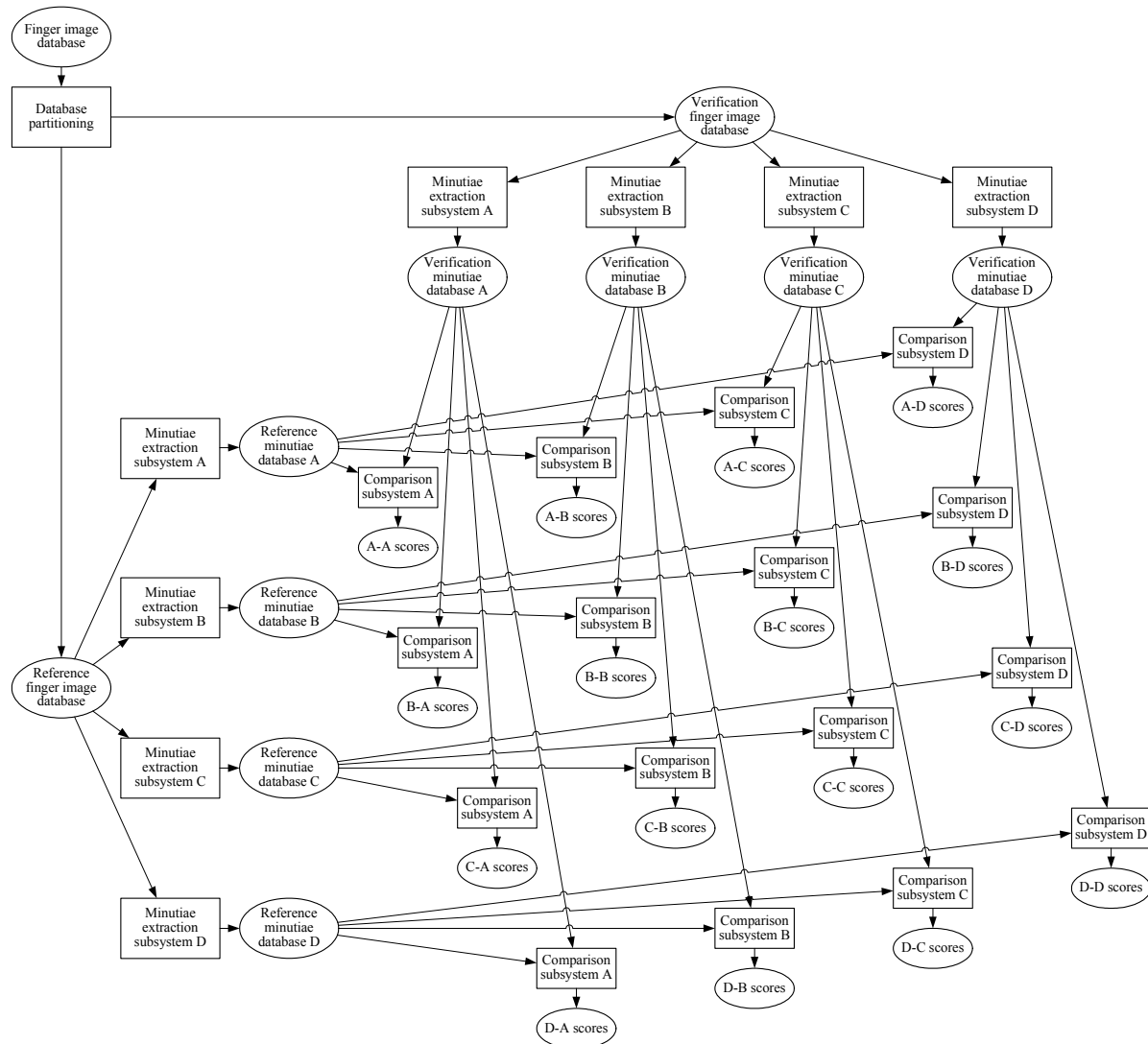


Figure 2 Test architecture for minutiae-minutiae comparison scenario

To test each combination of minutiae extraction and minutiae comparison involves the steps:

- Creation of minutiae template databases from the finger image databases, using each of the minutiae extraction subsystems under test. This results in a reference minutiae database, and a verification minutiae database for each minutiae extraction algorithm
- Comparison of minutiae templates against each other, using each minutiae comparison subsystems under test and generating similarity scores. This results in a set of genuine and impostor transaction scores for each grouping of minutiae-extraction and minutiae comparison subsystems being considered in the scenario.

After the subsystems under test have been tested, the following test analysis steps shall be carried out:

- Statistical analysis of the generated similarity scores to derive an interoperability matrix showing single-supplier and multi-supplier error rates as the cells of the interoperability matrix, with the rows and columns of the matrix indexing the suppliers.
- A statistical hypotheses test that the observed error rates within the interoperability matrix are below the specified target levels required for interoperability, and
- Group the subsystems under test into maximum classes of interoperable subsystems.

For the chosen performance metrics and grouping criteria, see [MTIT D3.1].

7.3 Generation of minutiae template databases

Using each of the minutiae extraction subsystems under test, from each of the finger image databases a set of databases of minutiae templates in the 1-finger record format and a set of databases of minutiae templates in the 1-finger card format shall be generated.

There is a requirement is to be able to do post-extraction filtering based on characteristics of the original image data, if needed. Therefore an XML data element allowing identification of the minutiae extraction subsystem used to generate the minutiae or proprietary record shall be added to the XML description of each minutiae template.

7.4 Minutiae comparisons

For each combination of reference minutiae database A (generated by minutiae extraction subsystem A), verification minutiae database B (generated by minutiae extraction subsystem B), and minutiae comparison subsystem C, generate sets of genuine and impostor scores for verification attempts. (Genuine attempts are the comparisons between minutiae records generated from the mates in the reference and verification fingerprint databases. Impostor attempts are the comparisons made between non-mates, i.e., same finger positions but on different individuals).

There is a requirement to be able to do post-match filtering based on characteristics of the original image and template data, and partitioning of the matching score data.

7.5 Computation of single-supplier and cross-supplier performance metrics

The “figure of merit” considered in the first, investigative phase of testing for each single-supplier or multi-supplier combination will be the false non-match rate at a given false match rate. Interoperability matrices will be produced containing FNMR values at FMR = 1%, FMR = 0.1%, and FMR = 0.01%, for each scenario as appropriate, with/without low quality prints.

For the investigative phase of testing, a decision threshold will be set so that, for the particular combination being tested, the false match rate attains the prescribed value. (A further investigation is proposed where the threshold is based solely on the single-supplier performance data.)

A combination of subsystems under test will be considered interoperable if the corresponding observed FNMR value in the interoperability matrix supports the working hypothesis that the corresponding honest FNMR is below a chosen threshold p . How to decide whether or not an observed error rate supports this working hypothesis has been outlined in Section 4.2. The statistical hypothesis test is based on the assumption of mutually independent attempts. If this assumption does not apply, then this shall be reported, and the test shall still be performed, but the results shall be interpreted with more caution.

7.6 Determination of maximum classes of interoperable subsystems

A maximum class of interoperable subsystems is a subset of subsystems under test such that

- all its elements are interoperable with each other, and
- there is no implementation under test that is interoperable with all elements of this subset, but not included in it.

Maximum classes of interoperable subsystems can be found by systematically grouping the interoperable subsystems under test.

8 Investigations

Several analyses will be made which involve comparing results of the different interoperability tests.

- What are the variations in performance over the different scenario options?
- Analysis of loss of performance from using fingerprint minutiae data-interchange format.

Specification of a test and certification protocol

- Analysis of loss of performance from using multi-supplier fingerprint minutiae systems as opposed to single supplier systems.
- Analysis of loss of performance from using the compact card minutiae format as opposed to the standard record format.
- Can we model interoperable performance by attribution of errors to the fingerprint images, minutiae extraction algorithms, minutiae comparison algorithms and interoperability problems?
- Are interoperability problems mostly confined to low quality fingerprint images?
- Are interoperability measures robust over the different fingerprint image data sets?
- Are the most interoperable systems the same over all the data sets?
- Can levels of interoperability be defined for which components are similarly categorised for each data set?

References

- [ANSI/NIST-ITL 1-2000] American National Standard for Information Systems—Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information
- [BioAPI 01] BioAPI Specification. Version 1.1, March 2001, [http://bioapi.org/NewDownloads\(AddtoSite\)/BioAPI1.1.pdf](http://bioapi.org/NewDownloads(AddtoSite)/BioAPI1.1.pdf)
- [FVC 2000] Fingerprint verification competition 2000, <http://bias.csr.unibo.it/fvc2000/>
- [ILO 2004] ILO seafarers' identity documents biometric testing campaign report, Part I. ILO, 2004
- [ISO/IEC 19794-2] International Standard ISO/IEC 19794-2:2005, Information technology – Biometric data interchange formats – Part 2: Finger minutiae data
- [ISO/IEC 19794-4] International Standard ISO/IEC 19794-4:2005, Information technology – Biometric data interchange formats – Part 4: Finger image data
- [ISO/IEC 19795-1] International Standard ISO/IEC 19795-1:2006, Information technology – Biometric performance testing and reporting – Part 1: Principles and framework
- [ISO/IEC 19795-2] Final Committee Draft ISO/IEC 19795-2:2006, Information technology – Biometric performance testing and reporting – Part 2: Testing methodologies for technology and scenario evaluation
- [ISO/IEC 19795-4] 3rd Committee Draft ISO/IEC 19795-4:2006, Information technology – Biometric performance testing and reporting – Part 4: Performance and interoperability testing of data interchange formats
- [MTIT D2.1] Data capture requirements. Report – restricted to other programme participants (including the Commission Services), Deliverable D2.1 of the MTIT Project
- [MTIT D3.1] Criteria of interoperability. Public Report, Deliverable D3.1 of the MTIT Project
- [MTIT D4.1] Required test bed functionality. Report – restricted to other programme participants (including the Commission Services), Deliverable D4.1 of the MTIT Project