INTERNATIONAL STANDARD

**ISO/IEC 29794-1**

Second edition
2016-01-15

Corrected version
2016-09-15

# Information technology — Biometric sample quality —

## Part 1: Framework

*Technologies de l'information — Qualité d'échantillon biométrique — Partie 1: Cadre*

Reference number
ISO/IEC 29794-1:2016(E)

**COPYRIGHT PROTECTED DOCUMENT**



ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

# Contents

Page

# Foreword

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

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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information.

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

This second edition cancels and replaces the first edition (ISO/IEC 29794-1:2009), which has been technically revised to revise Clause 8 and Table 2, which describes the structure of quality record.

ISO/IEC 29794 consists of the following parts, under the general title *Information technology — Biometric sample quality*:

— *Part 1: Framework*

— *Part 4: Finger image data*

— *Part 5: Facial image data* [Technical Report]

— *Part 6: Iris image data*

ISO/IEC 29794 series is prepared to accommodate new, additional parts that address other modalities specified by ISO/IEC 19794, with part numbers and titles aligning appropriately. However, as Part 1 is intended for use by all modalities, a modality does not necessarily need a modality-specific part in order to make use of quality scores.

It is anticipated that a future version of each part of the ISO/IEC 19794 series will reference this part of ISO/IEC 29794 normatively, and their respective data fields will be updated as required.

This corrected version of ISO/IEC 29794:2016 incorporates the following corrections.

1. "as given in Formula (C.1)" has been deleted from C.2 a).

2. Table 2, row: 5-byte Quality Block, column: Governing Section + Description + Notes:

QAID values of 0 to 32767

is changed to

QAID values of 1 to 32767

3. A.2, table, row: 5, column: Block 1 Byte 4+5 (QAID)

0

is changed to

10

# Introduction

Quality metrics are useful for several applications in the field of biometrics. While ISO/IEC 19784-1 specifies a structure and gives guidelines for quality score categorization, ISO/IEC 29794 defines and specifies methodologies for objective, quantitative quality score expression, interpretation, and interchange. This International Standard is intended to add value to a broad spectrum of applications in a manner that encourages competition, innovation, interoperability and performance improvements, and avoids bias towards particular applications, modalities, or techniques.

This International Standard presents several biometric sample quality scoring tools, the use of which is generally optional but can be determined as mandatory by particular Application Profiles or specific implementations.

A number of applications can benefit from the use of biometric sample quality data; an example is the use of real-time quality feedback upon enrolment to improve the operational efficiency and performance of a biometric system. The association of quality data with biometric samples is an important component of quality metric standardization. Quality fields as specified in 7.1 and 7.2 will be incorporated into data interchange formats. If a CBEFF header is present, then CBEFF_BDB_quality may additionally be used to express quality data. Useful analyses can be performed using quality data along with other data in order to improve the performance of a biometric system. For example, correlating quality data to other system metrics can be used to diagnose problems and highlight potential areas of performance improvement.

This edition introduces encoding of a vector of quality metrics.

# Information technology — Biometric sample quality —

# Part 1: Framework

## 1 Scope

This part of ISO/IEC 29794, for any or all biometric sample types as necessary, establishes the following:

— terms and definitions that are useful in the specification and use of quality metrics;

— purpose and interpretation of biometric quality scores;

— encoding of quality data fields in biometric data interchange formats;

— methods for developing biometric sample datasets for the purpose of quality score normalisation;

— format for exchange of quality algorithm results;

— methods for aggregation of quality scores.

The following are outside the scope of this part of ISO/IEC 29794:

— specification of minimum requirements for sample, module, or system quality scores;

— performance assessment of quality algorithms;

— standardization of quality algorithms.

## 2 Conformance

A biometric sample quality record shall conform to this part of ISO/IEC 29794 if its structure and data values conform to the formatting requirements of Clause 7. Conformance to normative requirements of 7.1 and 7.2 fulfils Level 1 and Level 2 conformance as specified in ISO/IEC 19794-1:2011, Annex A. Conformance to normative requirements of 7.3 is Level 3 conformance as specified in ISO/IEC 19794-1:2011, Annex A.

## 3 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 19794–1:2011, *Information technology — Biometric data interchange formats — Part 1: Framework*

ISO/IEC 19785–1, *Information technology — Common Biometric Exchange Formats Framework — Part 1: Data element specification*

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

## 4 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 2382-37, ISO/IEC 19794-1 and the following apply.

**4.1**
**acquisition fidelity**
*fidelity* (4.6) of a sample attributed to the acquisition process

**4.2**
**character**
contributor to *quality* (4.11) of a sample attributable to inherent properties of the *source* (4.17)

**4.3**
**environment**
physical surroundings and conditions where biometric capture occurs, including operational factors such as *operator* (4.9) skill and enrolee cooperation level

**4.4**
**extraction fidelity**
component of the *fidelity* (4.6) of a sample attributed to the biometric feature extraction process

**4.5**
**extrinsic**
when used to describe a *quality score* (4.12), requiring reference to an external *source* (4.17), such as a standard, register, or technical specifications for full *interpretation* (4.8) and normalisation

**4.6**
**fidelity**
expression of how accurately a biometric sample represents its *source* (4.17) biometric characteristic

Note 1 to entry: The fidelity of a sample comprises components attributable to one or more of the processing steps: acquisition, extraction, signal processing.

**4.7**
**intrinsic**
when used to describe a *quality score* (4.12), conveying fully *interpreted* (4.8), normalised data without the requirement for additional *extrinsic* (4.5)information for *quality score normalisation* (4.13)

**4.8**
**interpretation**
process of analysing a *quality score* (4.12) along with other data in order to give that score contextual, relative meaning

**4.9**
**operator**
individual who processes a capture subject in a biometric system, performing or supervising capture and recapture

**4.10**
**performance**
assessment of false match rate, false non-match rate, failure to enrol rate and failure to acquire rate of a biometric system

**4.11**
**quality**
degree to which a biometric sample fulfils specified requirements for a targeted application

Note 1 to entry: Specified quality requirements may address aspects of quality such as focus, resolution, etc. Implicit quality requirements address the likelihood of achieving a correct comparison result.

**4.12**
**quality score**
quantitative expression of *quality* (4.11)

**4.13**
**quality score normalisation**
rescaling of *quality scores* (4.12) to improve consistency in scale and *interpretation* (4.8)

**4.14**
**quality score normalisation dataset**
**QSND**
dataset of biometric samples annotated with *quality scores* (4.12) for use in *quality score normalisation* (4.13)

Note 1 to entry: Target quality scores may be assigned on the basis of *performance* (4.10) outcomes using the sample in question, or may be based on quality factors recorded in acquisition of the dataset.

**4.15**
**quality score percentile rank**
**QSPR**
percentile rank of the *quality score* (4.12) of a biometric sample, derived from its own utility score and those of other samples in an identified control dataset

Note 1 to entry: See *QSND* (4.14).

**4.16**
**raw quality score**
*quality score* (4.12) that has not been *interpreted* (4.8), either by the creator or recipient of the score, and alone may not intrinsically provide contextual information

**4.17**
**source**
physical body part or function represented by a biometric sample

**4.18**
**utility**
observed *performance* (4.10) of a biometric sample or set of samples in one or more biometric systems

Note 1 to entry: The *character* (4.2) of the sample *source* (4.17) and the *fidelity* (4.6) of the processed samples contribute to, or similarly detract from, the utility of the sample.

Note 2 to entry: Utility may combine performance measures such as false match rate, false non-match rate, failure to enrol rate, and failure to acquire rate.

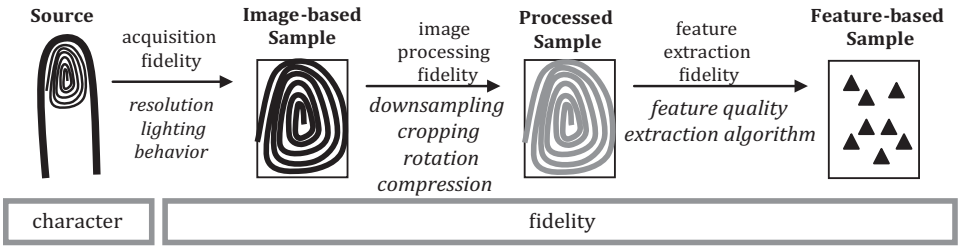
## 5 Abbreviated terms

| | |
|---|---|
| BDB | biometric data block |
| BDIR | biometric data interchange record |
| BIR | biometric information record |
| CBEFF | common biometric exchange formats framework (ISO/IEC 19785) |
| FERET | facial recognition technology database |
| FNMR | false non-match rate |
| QAID | quality algorithm identifier |
| QSND | quality score normalisation dataset |
| QSPR | quality score percentile rank |
| QVID | quality algorithm vendor identifier |
| XML | eXtensible Markup Language |

# 6 Biometric sample quality criteria

## 6.1 Reference model

In biometrics, the term "quality" is used to describe several different aspects of a biometric sample that contribute to the overall performance of a biometric system. For the purposes of standardization, this part of ISO/IEC 29794 defines terms, definitions, and a reference model for distinguishing between these different aspects of quality, illustrated in Figure 1. Figure 2 illustrates the relationship between character, fidelity, quality, utility, and system performance.




Quality = Function [character, fidelity components]
Utility reflects the impact of the quality of a single sample on system performance

**Figure 1 — Quality reference model illustration**

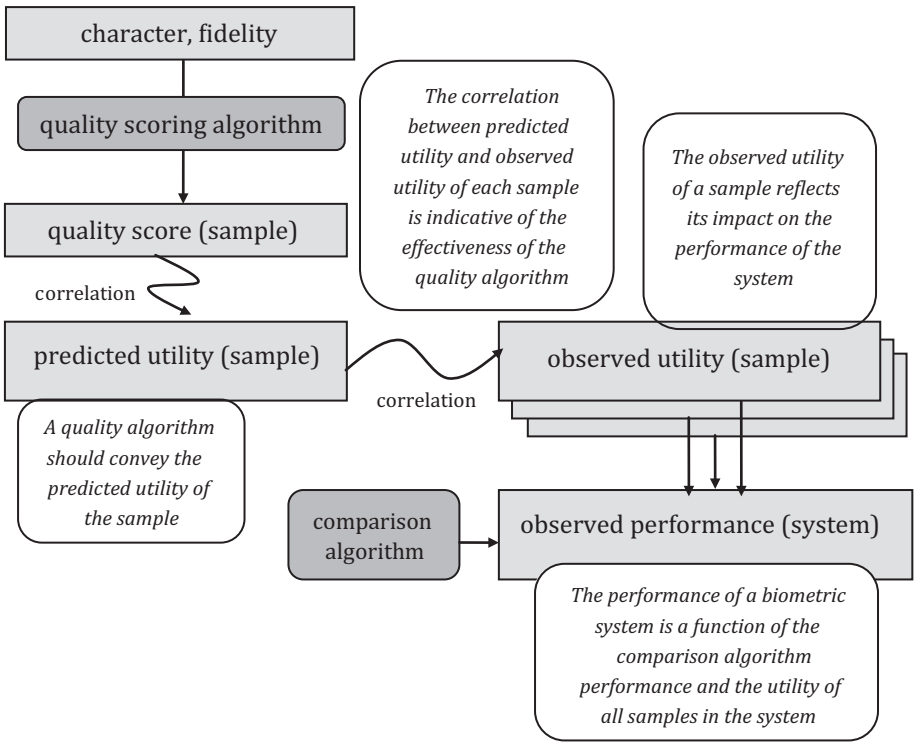



**Figure 2 — Relationship between quality and system performance**

## 6.2 Quality components: character, fidelity, utility

The term "quality" as it is currently used in the field of biometrics has several connotations, depending on context. Three prevalent uses are to subjectively reflect the following.

a) ***Character*** of a sample. An expression of quality based on the inherent properties of the source from which the biometric sample is derived. For example, a scarred fingerprint has poor character and blepharoptosis (droopy eyelid) causes poor iris character.

b) ***Fidelity*** of a sample to the source from which it is derived. An expression of quality based on fidelity reflects the degree of its similarity to its source. Sample fidelity is comprised of fidelity components contributed by different processes.

c) ***Utility*** of a sample within a biometric system. An expression of quality based on utility reflects the predicted positive or negative contribution of an individual sample to the overall performance of a biometric system. Utility-based quality is dependent on both the character and fidelity of a sample. Utility–based quality is intended to be more predictive of system performance, e.g. in terms of false match rate, false non-match rate, failure to enrol rate, and failure to acquire rate, than measures of quality based on character or fidelity alone. See Table 1.

The term "quality" should not be solely attributable to the acquisition settings of the sample, such as image resolution, dimensions in pixels, grey scale/colour bit depth, or number of features. Though such factors can affect sample utility and could contribute to the overall quality score.

Note that the character and utility of an acquired sample depend on the features to be considered by the comparator. For instance, the same finger image may be of low character and utility with respect to minutiae recognition (because of too few minutiae), but of high character and utility with respect to spectral pattern recognition.

**Table 1 — Illustration of relationship between fidelity, utility, and character**

| | | Fidelity | |
|---|---|---|---|
| | | Low | High |
| Character | Low | Low fidelity and low character results in low utility. Recapture might improve utility. However, if possible, use of other biometric characteristics is recommended. | High fidelity and low character results in low utility. Recapture will not improve utility. Use of other biometric characteristics is recommended. |
| | High | Samples with high character and low fidelity typically will not demonstrate high utility. Utility can be improved upon recapture or image enhancement techniques. | Samples with high character and high fidelity indicate capture of useful sample. High utility is expected. |

## 6.3 Usefulness of quality data

### 6.3.1 Real-time quality assessment

Real-time quality data can be used by an operator, automated system, or a user to help improve the average quality of biometric samples submitted upon enrolment. This feedback might indicate the character, fidelity, utility, and improvability of a sample. In this way, operational efficiency and overall system performance can be improved by assisting an operator, or augmenting an automated quality control system, in decisions to accept the sample, reject the sample, reattempt a capture, or declare a failure to acquire or failure to enrol. Quality data can be retained for later use in, for example, determining whether an enrolment sample should be replaced when the next sample is captured.

### 6.3.2 Use in different applications

A particular biometric sample might be used for several different applications and therefore, its associated quality data should be applicable to all of these. This would include both one-to-one and one-to-many comparisons involving the use of comparison algorithms from different vendors that would interpret sample features differently and yield different comparison scores. The challenge in establishing a universal quality standard is in defining a metric that is sufficiently adaptable for use by all applications for which a particular sample might be used given that metrics of utility vary greatly between applications. Therefore, it should be recognized that it is a technical challenge to define a singular metric that accurately conveys the utility of a biometric sample for all applications for which it may be used, and this should be taken into consideration in defining quality standards. Thus, a quality metric (ideally predicting performance for a comparator or class of comparators) will likely be designed to capture only some of the failure modes and sensitivities of only a limited number of biometric systems. It may be useful to apply more than one quality metric in order to improve predictability of various failure modes.

It is useful for recipients of quality score data to be able to differentiate between scores generated by different quality algorithms and capture equipment. This data may be used to enable recipient software to be configured so that different thresholds or quality classifications can be applied to scores generated by different algorithms. In addition, by differentiating between scores from different algorithms, a recipient may isolate results from different algorithms and use the data to optimize thresholds accordingly.

### 6.3.3 Use as a survey statistic

Quality scores may be used to monitor operational quality. For example, aggregated quality scores could be compared with pre-set limits or monitored against an operational requirement. If, for example, quality scores are generated from biometric samples collected at many sites, or over different time periods, then they may be used to identify anomalous operation. For example, if face image quality is computed at the license issuance desks at a Department of Motor Vehicles, then a ranked list of aggregated quality scores could be used to identify desks that exhibit a lower than average quality, or to monitor trends over weeks or months.

### 6.3.4 Accumulation of relevant statistics

Reliable quality scores may be used to survey users and transactions to accumulate statistics giving conditional probabilities of the kind “given a quality X sample on finger A, what is the likelihood of a quality Y sample from finger A (or finger B)”. This will inform the system and/or operators over whether a higher quality sample is likely if another capture is attempted.

### 6.3.5 Reference dataset improvement

The association of quality data with a sample that is to be entered into a reference dataset is important for the maintenance and improvement of the reference dataset quality. The tracking of sample quality can lead to detection of potential deterioration of operator training or it may indicate deterioration in the performance of the sample capture equipment. Tracking of the sample quality data should be an important part of biometric systems operating procedures. The quality data may also be used to improve the quality of the reference file, and hence the performance of the biometric system. Improvement can be made by the replacement or possible augmentation of the stored information so as to make use of the best available quality data. Typically, the replacement decisions are linked to the comparator performance of the system processing the data.

### 6.3.6 Quality-based conditional processing

Biometric samples can be processed differently based on quality metrics. In particular, poor quality data can be processed using algorithms or thresholds different from those used for high-quality data.

### 6.3.7 Interchange of quality data by disparate systems

Standardized exchange of quality data between disparate systems is useful in retaining the modular interchangeability of local or remote system hardware and software components, and the integrity of quality data in the event of such an interchange.

For example, by using standardized exchange of quality data, consumers of quality data from a component require minimal modification if that component is replaced.

## 7 Data interchange format field definition

### 7.1 Binary encoding

A quality record shall consist of a length field followed by zero, one, or multiple 5-byte Quality Blocks, as shown in Figure 3.

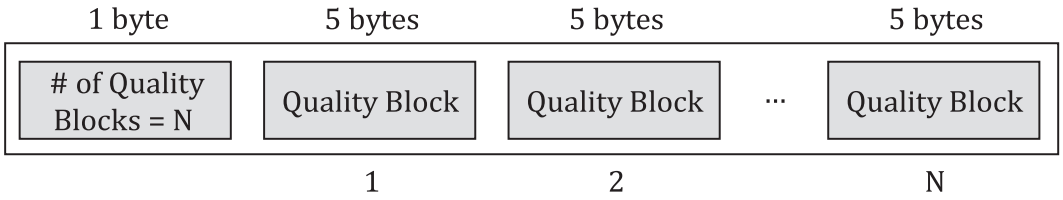



**Figure 3 — Location of number of Quality Block fields**

Table 2 defines the structure of a quality record. Each quality score of the standard metrics (specified in modality specific parts of ISO/IEC 29794), if computed, shall be encoded in a 5-byte Quality Block as specified in ISO/IEC 19794-1:2011.

The first byte of a 5-byte Quality Block shall contain the quality score.

The second and third bytes shall contain the CBEFF biometric organization identifier for organisation whose algorithm was used to compute the quality score. CBEFF biometric organization identifier for the ISO/IEC JTC 1 SC 37 *Biometrics*, as registered by IBIA, is 257 or $0101_{HEX}$ and shall be used if and only if an SC 37 approved reference implementation is used to compute the quality scores of the standard quality metrics defined in the modality specific parts of ISO/IEC 29794.

The fourth and fifth bytes shall contain the numeric identifier of the quality metric, which was computed for the representation image. If vendor ID is 257 (or $0101_{Hex}$) which is the SC 37 IBIA registered QVID, the QAID identifiers for the standard quality metrics shall be defined in modality specific parts of ISO/IEC 29794.

Quality score of 255 ($FF_{Hex}$) shall indicate that an attempt to calculate a quality score failed.

Figure 4 shows the following various possible encoding of a quality block:

a) fully standard quality component, where the quality component computation is defined in the modality specific part of ISO/IEC 29794 and its implementation is provided by SC 37;

b) semi-standard, where quality component computation is defined in the modality specific part of ISO/IEC 29794 but its implementation is provided by a vendor;

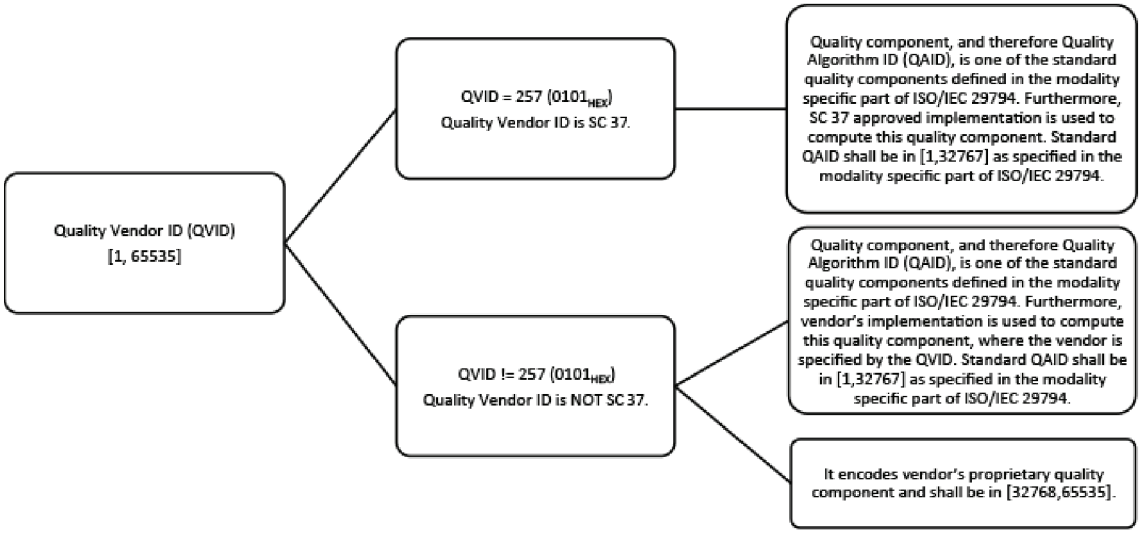
c) fully proprietary, where quality component definition, computation and implementation are provided by a vendor.

**Table 2 — Quality data record structure**

| | Byte # | Name | Length | Valid values | Governing section + Description + Notes |
|---|---|---|---|---|---|
| Record Length | 0 | Number of Quality Blocks | 1 byte | 0 to 255 | This field is followed by the number of 5-byte Quality Blocks reflected by its value.<br>A value of zero (0) means that no attempt was made to assign a quality score. In this case, no Quality Blocks are present. |
| 5-byte Quality Block (0 or more) | 1 | Quality score | 1 byte | 0 to 100, 255 | Quality score of the metric identified by the Quality Algorithm Identifier (QAID) in bytes 4 and 5 of this Quality Block.<br>If quality score is equal to 255 ($FF_{Hex}$), an attempt to calculate a quality score has failed. |
| | 2 to 3 | Quality Algorithm Vendor Identifier | 2 bytes | 0 to 65535<br>257 ($0101_{HEX}$) for standard quality. | This field shall contain the identifier of the vendor whose algorithm was used to compute quality. Quality algorithm vendor identifier shall be registered with IBIA or other approved registration authority as a CBEFF biometric organization in accordance with CBEFF vendor ID registry procedures in ISO/IEC 19785-1. A value of all zeros shall indicate that the value for this field is unreported.<br>SC 37 vendor ID (257 or $0101_{HEX}$) shall be used if and only if an SC 37 approved reference implementation is used to compute the quality score. |
| | 4 to 5 | Quality Algorithm Identifier (QAID) | 2 bytes | 1 to 65535 | The quality algorithm identifier shall be encoded in two bytes. A value of all zeros is not permitted.<br>If encoding standard quality metrics using the computation methods defined in the modality specific parts, quality algorithm identifiers defined in the modality specific parts shall be used.<br>For encoding of quality components not defined in the specific modality parts the quality algorithm identifier shall be assigned by the vendor or an approved registration authority.<br>QAID values of 1 to 32767 are reserved for quality components defined in the specific modality parts.<br>If vendor ID is 257 (or $0101_{Hex}$) which is the SC 37 IBIA registered QVID, the QAID identifiers for the standard quality metrics shall be defined in modality specific parts of ISO/IEC 29794.<br>For QVID other than 257 (i.e., SC 37), QAIDs equal or larger than 32768 specify the vendor's implementation of the quality metrics defined in the modality specific part of ISO/IEC 29794. QAIDs larger than 32768 can be used to encode vendor specific proprietary metrics. |

Quality scores shall always be placed within the quality record of the biometric data interchange record (BDIR) as defined in ISO/IEC 19794 associated with the sample. CBEFF quality fields should not be used in place of ISO/IEC 19794 quality fields but rather as supplementary data. The prescribed use

of CBEFF quality fields may be supplied by each CBEFF patron format standard and is beyond the scope of this this part of ISO/IEC 29794. Note that multiple quality scores calculated by the same algorithm (same quality algorithm vendor identifier and same quality algorithm identifier) shall not be present in a single BDIR.




NOTE The quality block, which consists of 2-bytes QVID, 2-bytes QAID and 1-byte Quality Score is encoded in ISO/IEC 19794 data interchange record with the modality of the representation specified in the general header of the record.

**Figure 4 — Diagram of all possible encodings of Quality Vendor ID (QVID) and Quality Algorithm ID (QAID)**

## 7.2 XML encoding

This subclause defines the syntax for quality data elements in XML documents in terms of XML type definitions.

```
<xs:complexType name="RegistryIDType">
 <xs:sequence>
  <xs:element name="Organization" type="xs:unsignedShort"/>
  <xs:element name="Identifier" type="xs:unsignedShort"/>
 </xs:sequence>
</xs:complexType>

<xs:complexType name="QualityType">
 <xs:sequence>
  <xs:element name="Algorithm" type="RegistryIDType" />
  <xs:choice>
   <xs:element name="Score" type="QualityScoreType" />
   <xs:element name="QualityCalculationFailed" type="NoDataFlagType"/>
  </xs:choice>
 </xs:sequence>
</xs:complexType>

<xs:simpleType name="QualityScoreType">
 <xs:restriction base="xs:unsignedByte">
  <xs:maxInclusive value="100"/>
 </xs:restriction>
</xs:simpleType>

<xs:complexType name="QualityListType">
 <xs:sequence>
  <xs:element name="Quality" type="QualityType" minOccurs="1" maxOccurs="255" />
 </xs:sequence>
```

```
</xs:complexType>
```

## 7.3 Quality score

### 7.3.1 Purpose

Quality algorithms shall produce quality scores that predict performance metrics such as either false match or false non-match. In cases where the system utilizes components from multiple vendors, the quality scoring method should aim to reflect the aspects of performance important for each algorithm used. As noted in 6.3.2, it is challenging to find a single quality measure that is universal, not vendor-specific and yet adequately indicates performance, and it may be useful to apply more than one quality algorithm.

### 7.3.2 Data transformation considerations

Data transformation by an application system is likely to impact the data quality (i.e. down-sampling or further compression). The impact of such transformations on the data quality metrics may be recomputed by the application system in accordance with guidance provided by ISO/IEC 29794. Any time a biometric sample undergoes a transformation, the quality of the transformed sample should be reassessed and associated with the transformed sample. For example, throughout an identity management system, a biometric sample may be stored in multiple formats (e.g. high resolution fingerprint image stored centrally and a minutia-based representation stored on a smart card).

### 7.3.3 Failure modes

To be predictive of performance, it may benefit a quality algorithm designer to produce quality scores that are intended to model known failure modes/sensitivities of a biometric comparator and image or signal processing algorithms. Further, to achieve some measure of generality, the quality score should be based on the set of sensitivities that are common to a class of system (e.g. minutia comparators).

### 7.3.4 Resolution

A quality apparatus shall provide for a mapping to at least four discrete values, which, when utilized towards a variety of applications, still maintains the ability to discriminate between distinct levels of performance, such as “excellent”, “adequate”, “marginal”, and “unacceptable”.

### 7.3.5 Summarisation

Annex B suggests procedures for the appropriate aggregation of utility-based quality scores over a collection of samples, e.g. enterprise-wide summarisation. The result is a summary value which supports monitoring of quality. Quality summarisation should be performed across similar usage, e.g. quality summarisation over all enrolment samples of an enterprise, or quality summarisation over all verification samples of an enterprise. In operations where users frequently interact with a biometric system (e.g. time and attendance applications), quality scores may be aggregated on a per user basis. This will reveal the existence of individuals that consistently yield low quality samples.

## 7.4 Quality algorithm identification

### 7.4.1 Overview

The Quality Algorithm Identifier (QAID) is an identifier of the quality algorithm used to calculate the quality score of the sample. As long as there are no common criteria for quality assessment, it is indispensable to enable the recipient of a BIR to differentiate between quality scores generated by different quality algorithms and adjust for any differences in processing or analysis as necessary. The quality algorithm vendor identifier (QVID) shall indicate the vendor of the identified quality algorithm. Different versions of a quality algorithm that yield different results shall be assigned different quality algorithm identifiers to allow unique identification. The combination of QVID and QAID is considered

to be a solution that may be implemented quickly but only partially achieves the goals of quality score standardization.

### 7.4.2 Methodology

This method requires as normative that if no quality scoring is attempted, then the value of the Number of Quality Blocks field is 0 and there are no Quality Blocks present. If Number of Quality Blocks is between 1 and 255, to uniquely identify the algorithm used to generate the encoded quality score, each of the quality blocks shall contain a quality algorithm vendor identifier (QVID) and a quality algorithm identifier (QAID) according to Table 2. Note that this method does not preclude, but rather complements, further work to standardize a universal quality scoring method (i.e. a score that intrinsically includes some degree of normalisation) such as QSND. See Clause 8.

A feature of this "quality algorithm identification" method is that the recipient of the raw quality score data may need to do some local analysis and/or processing to fully interpret the meaning of the scores. In other words, the sender of the score is not attempting to interpret the quality score for a potentially unknown application or destination. But importantly, the recipient can obtain the information on how the quality score is established from the quality algorithm vendor and develop appropriate means to automatically distinguish between quality scores generated by different quality algorithms and interpret them appropriately.

Some parts of ISO/IEC 29794 specify standardized computation method for the quality metrics defined in that part of ISO/IEC 29794.

## 7.5 Standardized exchange of quality algorithm results

Quality algorithm vendors should be able to offer results of their quality algorithms in a standardized way to the biometric community. On the other hand, consumers of ISO/IEC 19794 data interchange records are able to retrieve and process this information effectively in order to assess the value of the output of this quality algorithm to their implementation. This approach has the following benefits.

— Both quality algorithm vendors, as well as consumers have the ability to gain value from technical improvements, which is a necessary prerequisite in the starting phase of wide spread quality score use.

— In some applications, updates may be retrieved automatically, if the necessary infrastructure is there.

— It will re-shift the evaluation effort related with QAID from the consumer and integrators back to the quality algorithm vendors (which do the evaluation anyway).

— Over time, standardized test sets will evolve,

  — as it is in the interest of the quality algorithm vendor to use (a) reporting test set(s), that is of use for many costumers, and

  — need for new test sets will vanish over time and use of new test for non-obvious reasons will be critically reviewed by the community.

— Evolution of test sets will facilitate the development of QSND.

For the exchange, the following items shall be provided:

a) quality algorithm vendor ID;

b) quality algorithm ID;

c) minimum and maximum theoretical output value of the algorithm;

d) unique name of test set used (e.g. in form of "FERET-Greyscale" in the case of face recognition);

e) list of samples (e.g. for FERET “Duplicate 1” in the case of face recognition), that have been processed.

Everyone will be able to publish new test sets (biometric samples + a naming scheme).

A self-describing language like XML will be used for the description of the data sets, as well as for the evaluation results. The evaluation results could be maintained in a central registry or on a vendor site (via a link in the central registry).

An example implementation using XML can be found in Annex B.

## 8 Normalisation

Normalisation of quality score data is the process by which quality score data is processed by its recipient in order to give the scores local context and meaning, such as to make quality scores from different algorithms have similar meaning.

A raw quality score is assigned to a biometric sample by a particular quality algorithm. In order to interpret the raw score, the recipient of a score shall have some contextual information. The following information may be provided.

a) Extrinsically, in the form of metadata or off-line data (e.g. standard) that instructs the recipient on interpretation of the score. For example, if a quality score is accompanied by identification of the algorithm used to generate the quality score of the associated sample (i.e. QAID), then recipient software can be configured to use vendor-supplied data (e.g. suggested thresholds) to best process the sample. The algorithm could alternatively be used to perform analysis in order to fully optimize the interpretation of the scores given the local application and data. By identifying the algorithm, scores created by different algorithms could be differentiated so that, for example, different thresholds could be applied to the sample depending on the source of the quality score.

b) Intrinsically, in the form of a normalised quality score. Normalisation of quality score data provides contextual information about the score. An example is a quality score representing the perceived likelihood that a sample, when compared, will result in a false non-match.

QAID enables vendor-specific scaling, such that the 0 to 100 scale correlates to some other scale reflecting the above. For example, the recipient of a file would be encouraged to analyse the correlation of quality scores to false match rate and false non-match rate of the samples processed by their comparator. The results could be used to, for example, specify an acceptance operating threshold. This method provides the recipient the information necessary to interpret the scores in a way that is relevant to their own environment and application, and permits the use of many different algorithms or versions of algorithms in a single system.

The purpose of quality score normalisation (QSN) corpus is to provide a consistent interpretation of quality scores through normalizing quality scores or Quality Score Percentile Rank (QSPR). QSPR enables universal expression and interpretation of a quantitative sample quality score, which is that quality algorithm “X” considers biometric sample “Y” to have a quality percentile rank “Z”. The translation of raw quality scores to percentile rank scores is achieved by running a standardized corpus of samples through a given quality algorithm and pairing all possible raw score outcomes to percentile rank scores.

# Annex A
(informative)

# Example of encoding a biometric quality record

## A.1 XML example

```
<QualityList>
 <Quality>
  <Algorithm>
   <Organization>212</Organization>
   <Identifier>10</Identifier>
  </Algorithm>
  <QualityCalculationFailed />
 </Quality>
 <Quality>
  <Algorithm>
   <Organization>300</Organization>
   <Identifier>1</Identifier>
  </Algorithm>
  <Score>5</Score>
 </Quality>
 <Quality>
  <Algorithm>
   <Organization>257</Organization>
   <Identifier>9</Identifier>
  </Algorithm>
  <Score>74</Score>
 </Quality>
 <Quality>
  <Algorithm>
   <Organization>257</Organization>
   <Identifier>8</Identifier>
  </Algorithm>
  <Score>89</Score>
 </Quality>
 <Quality>
  <Algorithm>
   <Organization>21</Organization>
   <Identifier>8</Identifier>
  </Algorithm>
  <Score>48</Score>
 </Quality>
</QualityList>
```

## A.2 Binary example

| | **Block 1** | | | **Block 2** | | | **Block 3** | | | **Block 4** | | | **Block 5** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Q blocks | Quality computation failed. Vendor proprietary Quality Algorithm ID is 10 and Quality vendor ID is 212. | | | Quality score is 5. Quality vendor ID is 300 and QAID = 1. | | | Quality score is 74. Standard quality component 9 has been computed using an SC 37 approved reference. Implementation. Quality component 9 is defined in the modality specific part. | | | Quality score is 89. Standard quality component 8 has been computed using an SC 37 approved reference. Implementation. Quality component 8 is defined in the modality specific part. | | | Quality score is 48. Standard quality component 8 has been computed using quality algorithm from vendor with QVID = 21. Quality component 8 is defined in the modality specific part. | | |
| | Byte 1 (score) | Byte 2+3 (QVID) | Byte 4+5 (QAID) | Byte 1 (score) | Byte 2+3 (QVID) | Byte 4+5 (QAID) | Byte 1 (score) | Byte 2+3 (QVID) | Byte 4+5 (QAID) | Byte 1 (score) | Byte 2+3 (QVID) | Byte 4+5 (QAID) | Byte 1 (score) | Byte 2+3 (QVID) | Byte 4+5 (QAID) |
| 5 | 255 | 212 | 10 | 5 | 300 | 1 | 74 | 257 | 9 | 89 | 257 | 8 | 48 | 21 | 8 |

# Annex B
(informative)

# Example of standardized exchange of quality algorithm results

## B.1 General

As described in 7.5, Quality algorithm vendors should be able to offer results of their quality algorithms in a standardized way to the biometric community. Particularly, exchanging quality scores generated from public datasets will be most useful in providing technical insight and allow the consumers of the quality scores to examine and understand how the quality scores relate to intrinsic information content of the images.

This Annex provides an example of exchanging such information in XML format.

## B.2 Example quality exchange document

The following shows an example of an XML coding for vendor "SampleVendor" with id = 123 publishing the results of algorithm "SampleAlgo_v10" with id = 456 on test sets "FERET-grayscale" and "FERET-color".

```
<?xml version="1.0" encoding="UTF-8"?>
<iso-vendor-quality-report
    xmlns:iso="http://www.iso.org/29794-1"
    xmlns:xsi="xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.iso.org/29794-1
                        http://www.iso.org/29794-1.xsd"
    quality-vendor-id="123"
    quality-algorithm-id="456"
    quality-algorithm-min-value="0.0"
    quality-algorithm-max-value="100.0">
    <iso:testset
        name="FERET-grayscale"
        location="http://www.nist.gov/humanid/feret/feret_master.html">
       <iso:sample
           name="00001fa010_930831" quality-value="73.64"/>
       <iso:sample
           name="00002fa010_930831" quality-value="48.91"/>
    </iso:testset>
    <iso:testset
        name="FERET-color"
        location="http://www.nist.gov/humanid/colorferet/home.html">
       <iso:sample
           name="00002_931230_fa" quality-value="51.26"/>
       <iso:sample
           name="00002_931230_fb" quality-value="82.17"/>
    </iso:testset>
</iso-vendor-quality-report>
```

## B.3 Informative schema for sample XML quality exchange document

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
    xmlns:xs="http://www.w3.org/2001/XMLSchema"
    targetNamespace="http://www.iso.org/29794-1"
    elementFormDefault="qualified">
  <xs:annotation>
      <xs:documentation xml:lang="en">
          ISO/IEC 29794-1 Vendor Quality Report
      </xs:documentation>
  </xs:annotation>
```

```
  <xs:element
      name="iso-vendor-quality-report"
      type="iso-vendor-quality-report-type">
  </xs:element>
  <xs:complexType
      name="iso-vendor-quality-report-type">
      <xs:sequence minOccurs="0" maxOccurs="unbounded">
         <xs:element
             name="iso:testset"
             type="iso:testset-type"/>
      </xs:sequence>
      <xs:attribute
          name="quality-vendor-id"
          type="xs:ID"
          use="required"/>
      <xs:attribute
          name="quality-algorithm-id"
          type="xs:ID"
          use="required"/>
      <xs:attribute
          name="quality-algorithm-min-value"
          type="xs:float"
          use="required"/>
      <xs:attribute
          name="quality-algorithm-min-value"
          type="xs:float"
          use="required"/>
  </xs:complexType>
  <xs:complexType
      name="testset-type">
      <xs:sequence minOccurs="0" maxOccurs="unbounded">
          <xs:element
              name="iso:sample"
              type="iso:sample-type"/>
      </xs:sequence>
      <xs:attribute
          name="name"
          type="xs:ID"
          use="required"/>
      <xs:attribute
          name="location"
          type="xs:anyURI"
          use="required"/>
  </xs:complexType>
  <xs:complexType
      name="sample-type">
      <xs:attribute
          name="name"
          type="xs:ID"
          use="required"/>
      <xs:attribute
          name="quality-value"
          type="xs:float"
          use="required"/>
  </xs:complexType>
</xs:schema>
```

# Annex C
(informative)

# Procedures for aggregation of utility-based quality scores

## C.1 Purpose

This Annex suggests procedures for the appropriate aggregation of utility-based quality scores over a collection of samples, e.g. enterprise-wide summarisation. The result is a summary value which supports monitoring of quality. Quality summarisation should be performed across similar usage, e.g. quality summarisation over all enrolment samples of an enterprise, or quality summarisation over all verification samples of an enterprise. In operations where users frequently interact with a biometric system (e.g. time and attendance applications), quality scores may be aggregated on a per user basis. This will reveal the existence of individuals that consistently yield low quality samples.

## C.2 Method

Suppose some enterprise collects biometric samples and measures the quality of each using a quality scoring algorithm. We assume quality scores are quantized into $L$ levels so that (without loss of generality) $q = 0, \ldots, L$, where $q = 0$ and $q = L$ indicate lowest and highest quality scores, respectively. If the number of biometric samples collected over some interval in an operational situation is $n$ and this is composed of $n_q$ biometric samples of quality, $q$, then we could compute the mean quality across all $n$ samples. However, arithmetic mean is not the preferred method of summarizing quality scores because all samples, regardless of their quality scores, are given the same weight. If instead the expected utility of a biometric sample $t$ of quality $q$ is $u_q = \mathrm{U}(q)$, then a better summary statement of quality is:

$$\bar{q} = \frac{\sum_{q=0}^{L} u_q n_q}{\sum_{q=0}^{L} n_q} \tag{C.1}$$

If the utility $u_q$ is actually an estimate of the false reject rate for samples of quality $q$ of a reference biometric sample verification system operating at some reasonable threshold, then $\bar{q}$ will be an estimate of the expected error rate. We proceed by introducing a procedure to compute utility $u_q$ for different levels of a quality scoring algorithm such that the summarized quality score is an estimate of the expected error rate.

Consider a biometric corpus contains *2N* pairs of biometric samples from *N* persons. The first sample represents an enrolment sample, and the second represents the authentication sample. The samples have integer qualities $q_j^{(1)}$ and $q_j^{(2)}$ *for j=1,…, N*. Applying *V* comparison algorithms to the samples, we get the following:

— N genuine similarity scores, $s_{jj}^{(v)}$;

— up to $N(N-1)$ impostor scores, $s_{jk}^{(v)}$ with $j \neq k$

where $v = 1,\ldots, V$ and $\mathrm{V} \geq 1$.

a) For all comparison algorithms $v$ and quality scores $q$ compute $\mathrm{FNMR}^v(\tau,i)$ of authentication samples of quality $i$ with enrolment samples of quality better than or equal to $i$ at operating threshold $\tau$ using

genuine scores of comparison algorithm *v*. Note that we assumed higher quality scores indicate better quality.

for ($v$ = 1,...,$V$)

for ($i$ = 1,...,$L$)

$$\mathrm{FNMR}^{v}\left(\tau,i\right)=\frac{\left|\{s_{jj}^{(v)}:s_{jj}\leq\tau,q_{j}^{(1)}\geq i,q_{j}^{(2)}=i\}\right|}{\left|\{s_{jj}^{(v)}:s_{jj}<\infty,q_{j}^{(1)}\geq i,q_{j}^{(2)}=i\}\right|} \tag{C.2}$$

end

which results in the following array:

$$\begin{pmatrix}\mathrm{FNMR}^{1}\left(\tau,1\right)&\mathrm{FNMR}^{2}\left(\tau,1\right)\ldots\mathrm{FNMR}^{v}\left(\tau,1\right)\\\mathrm{FNMR}^{1}\left(\tau,2\right)&\mathrm{FNMR}^{2}\left(\tau,2\right)\ldots\mathrm{FNMR}^{v}\left(\tau,2\right)\\\mathrm{FNMR}^{1}\left(\tau,L\right)&\mathrm{FNMR}^{2}\left(\tau,L\right)\ldots\mathrm{FNMR}^{v}\left(\tau,L\right)\end{pmatrix}$$

b) Compute weight $u_i$ as given in Formula (C.3):

$$u_{i}=\frac{\sum_{v=1}^{V}\mathrm{FNMR}^{v}(\tau,i)}{\sum_{q=0}^{L}\sum_{v=1}^{V}\mathrm{FNMR}^{v}(\tau,q)} \tag{C.3}$$

Thus, the aggregated quality across an enterprise is as given in Formula (C.4):

$$Q=\sum_{i=0}^{L}u_{i}p_{i} \tag{C.4}$$

where $u_i$ are estimated posterior probabilities above. As probabilities, these values will not be on a range familiar to users. Note that if all samples were of best quality (i.e. $q=L$), the result would be $Q = u_L$. Similarly, the worst case is when all samples in the enterprise are of $q = 0$, which results in $Q = u_0$. Thus, this formulation would result in quality summaries on the range $[u_0, u_L]$. Users should regard Formula (C.4) as a measure of expected overall FNMR. However, it is recommended to transform $[u_0, u_L]$ to the recommended range [0,100], which has 0 as the lowest quality and 100 as the best. This can be accomplished by either of the following methods:

1) by relating the quality summary number $Q$ (i.e. expected error rate) back to the native quality range by using the inverse of the utility function, as given in Formula (C.5):

$$\bar{Q}=U^{-1}\left(Q\right)=U^{-1}\left(\sum_{i=0}^{L}u_{i}p_{i}\right) \tag{C.5}$$

where $U^{-1}$ is a function approximation (e.g. piece-wise linear interpolation) of pairs $(i,u_i)$;

2) by mapping (e.g. linear mapping) $[u_0, u_L]$ to [0, 100]. Thus, quality summaries mapped to [0, 100] are given by Formula (C.6):

$$\bar{Q}=\frac{100u_{0}}{u_{0}-u_{L}}-\sum_{i=0}^{L}\frac{100u_{i}}{u_{0}-u_{L}}p_{i} \tag{C.6}$$

NOTE 1 Weights in Formula (C.3) are estimates of the observed false non-match rates computed at some fixed threshold. The result is that these weights are most accurate for that particular threshold and not as accurate for biometric systems operating at other thresholds. In verification applications, where operating threshold is fixed at τ, users of a quality scoring algorithm are to follow the outlined procedure to establish dedicated weights.

NOTE 2 Weights in Formula (C.3) are consensus estimates. That is they were estimated using the observed false non-match rates from a set of comparison algorithms. The result is that the weights are not exactly the weights that would be used for any one algorithm, or for a specified set of algorithms. Weights in Formula (C.3) are regarded as best practice estimates to be used unless other details about the application are known. In verification applications, where a specific set of one or more comparison algorithms are known and available, users of an algorithm are to follow the outlined procedure to establish dedicated weights.

# Bibliography

[1] ISO/IEC 19784–1, *Information technology — Biometric application programming interface — Part 1: BioAPI specification*

**ICS 35.040**
Price based on 19 pages