

---

---

**Information technology — Biometric data  
interchange formats —**

**Part 5:  
Face image data**

*Technologies de l'information — Formats d'échange de données  
biométriques —*

*Partie 5: Données d'image de la face*



Reference number  
ISO/IEC 19794-5:2011(E)

**AVA Documents**  
Document Sharing Hub  
© ISO/IEC 2011



**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2011

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

|                                                                                   |    |
|-----------------------------------------------------------------------------------|----|
| Foreword .....                                                                    | v  |
| Introduction.....                                                                 | vi |
| 1 Scope .....                                                                     | 1  |
| 2 Conformance .....                                                               | 1  |
| 3 Normative references.....                                                       | 2  |
| 4 Terms and definitions .....                                                     | 2  |
| 5 The Face Image Data Record Format .....                                         | 4  |
| 5.1 Overview.....                                                                 | 4  |
| 5.2 Data Conventions .....                                                        | 7  |
| 5.3 The General Header.....                                                       | 8  |
| 5.4 The Representation Header.....                                                | 9  |
| 5.5 The Facial Information Block .....                                            | 11 |
| 5.6 The Landmark Point Block .....                                                | 16 |
| 5.7 The Image Information Block .....                                             | 22 |
| 5.8 The Representation Data block.....                                            | 25 |
| 5.9 The Image Data Block .....                                                    | 25 |
| 5.10 The 3D Information Block.....                                                | 25 |
| 5.11 The 3D Data Block.....                                                       | 32 |
| 6 The Basic Face Image Type .....                                                 | 35 |
| 6.1 Inheritance requirements for the Basic Face Image Type .....                  | 35 |
| 6.2 Image data encoding requirements for the Basic Face Image Type.....           | 35 |
| 6.3 Image data compression requirements for the Basic Face Image Type .....       | 35 |
| 6.4 Format requirements for the Basic Face Image Type .....                       | 35 |
| 7 The Frontal Face Image Type .....                                               | 35 |
| 7.1 Inheritance requirements for the Frontal Face Image Type.....                 | 35 |
| 7.2 Scene requirements for the Frontal Image Type.....                            | 36 |
| 7.3 Photographic Requirements for the Frontal Image Type .....                    | 37 |
| 7.4 Digital requirements for the Frontal Image Type .....                         | 38 |
| 7.5 Format requirements for the Frontal Image Type .....                          | 39 |
| 8 The Full Frontal Image Type.....                                                | 39 |
| 8.1 Inheritance requirements for the Full Frontal Face Image Type .....           | 39 |
| 8.2 Scene requirements for the Full Frontal Face Image Type .....                 | 39 |
| 8.3 Photographic requirements for the Full Frontal Face Image Type .....          | 40 |
| 8.4 Digital requirements for the Full Frontal Face Image Type.....                | 42 |
| 8.5 Format requirements for the Full Frontal Image Type.....                      | 42 |
| 9 The Token Face Image Type .....                                                 | 42 |
| 9.1 Inheritance requirements for Token Face Image Type.....                       | 42 |
| 9.2 Digital requirements for the Token Face Image Type.....                       | 42 |
| 9.3 Format requirements for the Token Face Image Type .....                       | 44 |
| 10 The Post-processed Frontal Face Image Type.....                                | 44 |
| 10.1 Introduction.....                                                            | 44 |
| 10.2 Inheritance requirements for the Post-processed Frontal Face Image Type..... | 44 |
| 10.3 Format requirements for the Post-processed Frontal Face Image Type.....      | 44 |
| 11 The Basic 3D Image Type .....                                                  | 45 |
| 11.1 Inheritance Requirements for the Basic 3D Image Type.....                    | 45 |
| 11.2 The Basic 3D Image Type using the 3D Point Map representation.....           | 45 |

|              |                                                                                         |     |
|--------------|-----------------------------------------------------------------------------------------|-----|
| 11.3         | The Basic 3D Image Type using the 3D Vertex representation .....                        | 45  |
| 12           | The Full Frontal 3D Image Type .....                                                    | 45  |
| 12.1         | Inheritance requirements .....                                                          | 45  |
| 12.2         | Coordinate System Type.....                                                             | 46  |
| 12.3         | Pose of the 3D representation.....                                                      | 46  |
| 12.4         | Calibration Texture Projection Accuracy .....                                           | 46  |
| 12.5         | Requirements on Full Frontal 3D Image Types using the Range Image Representation .....  | 46  |
| 12.6         | Requirements on Full Frontal 3D Image Types using the 3D Point Map Representation.....  | 47  |
| 12.7         | Requirements on Full Frontal 3D Image Types using the 3D Vertex Representation .....    | 47  |
| 13           | The Token Frontal 3D Image Type .....                                                   | 47  |
| 13.1         | Inheritance requirements .....                                                          | 47  |
| 13.2         | Requirements on Token Frontal 3D Image Types using the Range Image Representation ..... | 48  |
| 13.3         | Requirements on Token Frontal 3D Image Types using the 3D Point Map Representation..... | 48  |
| 13.4         | Requirements on Token Frontal 3D Image Types using the Vertex Representation .....      | 48  |
| 14           | Registered Format Type Identifier .....                                                 | 48  |
| Annex A      | (normative) Conformance test methodology.....                                           | 49  |
| Annex B      | (informative) Best practices for Face Images .....                                      | 50  |
| Annex C      | (informative) Conditions for Taking Photographs.....                                    | 70  |
| Annex D      | (informative) Experimental studies .....                                                | 100 |
| Annex E      | (informative) The Frankfurt Horizon.....                                                | 110 |
| Bibliography | .....                                                                                   | 111 |

## Foreword

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

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

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

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

ISO/IEC 19794-5 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

This second edition constitutes a technical revision of the first edition (ISO/IEC 19794-5:2005), which is provisionally retained. It also incorporates the Amendments ISO/IEC 19794-5:2005/Amd.1:2007 and ISO/IEC 19794-5:2005/Amd.2:2009, and the Technical Corrigenda ISO/IEC 19794-5:2005/Cor.1:2008 and ISO/IEC 19794-5:2005/Cor.2:2008. This edition reflects the harmonization across the second generation of ISO/IEC 19794. Clause 5 contains descriptions of the harmonized general and representation headers; and Clauses 5 to 13 have been technically revised. Annexes C, D, and E have been added.

ISO/IEC 19794 consists of the following parts, under the general title *Information technology — Biometric data interchange formats*:

- *Part 1: Framework*
- *Part 2: Finger minutiae data*
- *Part 3: Finger pattern spectral data*
- *Part 4: Finger image data*
- *Part 5: Face image data*
- *Part 6: Iris image data*
- *Part 7: Signature/sign time series data*
- *Part 8: Finger pattern skeletal data*
- *Part 9: Vascular image data*
- *Part 10: Hand geometry silhouette data*
- *Part 11: Signature/sign processed dynamic data*
- *Part 13: Voice data*
- *Part 14: DNA data*

## Introduction

Face images, also commonly referred to as displayed portraits, have been used for many decades to verify the identity of persons. In recent years, digital face images are used in many applications including human examination as well as computer automated face recognition. Although photographic formats have been standardized in some cases such as for passports and driver licenses, there is a need to define a standard data format of digital face images to allow interoperability among vendors.

This part of ISO/IEC 19794 is intended to provide a face image format for face recognition applications requiring exchange of face image data. The typical applications are

- 1) human examination of facial images with sufficient resolution to allow a human examiner to ascertain small features such as moles and scars that might be used to verify identity,
- 2) human verification of identity by comparison of persons against facial images,
- 3) computer automated face biometric identification (one-to-many searching), and
- 4) computer automated face biometric verification (one-to-one comparison).

To enable many applications on a variety of devices, including devices that have limited resources available for data storage, and to improve face recognition accuracy, this part of ISO/IEC 19794 specifies not only a data format, but also scene constraints (lighting, pose, expression, etc.), photographic properties (positioning, camera focus, etc.) and digital image attributes (image resolution, image size, etc.).

Several face image types are introduced to define categories that satisfy requirements of some applications:

- **Basic:** This is the fundamental Face Image Type that specifies a record format including header and representation data. All Face Image Types adhere to the properties of this type. No mandatory scene, photographic and digital requirements are specified for this image type.
- **Frontal:** A Basic Face Image Type that adheres to additional requirements appropriate for frontal face recognition and/or human examination. Two types of Frontal Face Image Types are defined in this part of ISO/IEC 19794, Full Frontal and Token Frontal (or simply Token).
- **Full Frontal:** A Face Image Type that specifies frontal images with sufficient resolution for human examination as well as reliable computer face recognition. This type of Face Image Type includes the full head with all hair in most cases, as well as neck and shoulders. This image type is suitable for permanent storage of the face information, and it is applicable to portraits for passport, driver license, and “mugshot” images.
- **Token Frontal:** A Face Image Type that specifies frontal images with a specific geometric size and eye positioning based on the width and height of the image. This image type is suitable for minimizing the storage requirements for computer face recognition tasks such as verification while still offering vendor independence and human verification (versus human examination which requires more detail) capabilities.
- **Post-processed Frontal:** Applying digital post-processing to a captured image can modify this image in a way that it is more suitable for automatic face recognition. The Post-processed Frontal Face Image Type is thought of as the interchange format for these kinds of facial images.
- **Basic 3D:** The Basic 3D Image Type is the base Image Type of all 3D Face Image Types. All 3D Face Image Types obey normative requirements of this image type.
- **Full Frontal 3D:** The Full Frontal 3D Image Type combines a Full Frontal 2D image with additional 3D information.
- **Token Frontal 3D:** The Token Frontal 3D Image Type combines a Token Frontal 2D image with additional 3D information.

Table 1 shows the relationships between Face Image Types using the notion of inheritance. For example, Frontal inherits properties from Basic, which means that all normative clauses that apply to Basic also apply to Frontal.

**Table 1 — Inheritance of Face Image Types**

| Face Image Type        | Inherits from | Normative clauses | Informative annexes |
|------------------------|---------------|-------------------|---------------------|
| Basic                  | None          | 1, 2, 3, 4, 5, 6  | B.1                 |
| Frontal                | Basic         | 7                 | B.2                 |
| Full Frontal           | Frontal       | 8                 | B.3                 |
| Token Frontal          | Frontal       | 9                 | B.4                 |
| Post-processed Frontal | Frontal       | 10                |                     |

Figure 1 gives a general overview of the scene, photographic, digitization, and format requirements for the face image types specified in this part of ISO/IEC 19794.

| Requirements                                                                                                                                                                                                   |                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                               |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Scene                                                                                                                                                                                                          | Photographic                                                                                                                                                                                                      | Digital                                                                                                                                                                                                                                                                                                                                | Format                                                                                                                                                                                                                                        |
| <br>Lighting<br><br><br>Image and Subject | <br>Positioning<br><br><br>Camera Attributes | <br>Digital Camera<br><br><br>Analogue to Digital<br><br><br>Image Scanning | <br>Digital Specifications<br><br><br>Record Format and Organization |
| <i>Clauses:</i><br>Basic Face      None<br><br>Frontal Face      7.2<br>Full Frontal Face      8.2                                                                                                             | <i>Clauses:</i><br>Basic Face      None<br><br>Frontal Face      7.3<br>Full Frontal Face      8.3                                                                                                                | <i>Clauses:</i><br>Basic Face      None<br><br>Frontal Face      7.4<br>Full Frontal Face      8.4<br>Token Face      9.2                                                                                                                                                                                                              | <i>Clauses:</i><br>Basic Face      5<br>6.2<br>6.3<br>6.4<br>Frontal Face      7.5<br>Full Frontal Face      8.5<br>Token Face      9.3<br>Post-processed Frontal Face      10.3                                                              |

**Figure 1 — The types of imaging requirements specified in this part of ISO/IEC 19794. The Basic Face Image Type has no scene, photographic, or digital requirements**

## ISO/IEC 19794-5:2011(E)

This is a revision of ISO/IEC 19794-5:2005. The structure of the data format is not compatible with the previous version.

NOTE This part of ISO/IEC 19794 relies on other ISO International Standards.

# Information technology — Biometric data interchange formats —

## Part 5: Face image data

### 1 Scope

This part of ISO/IEC 19794

- specifies a record format for storing, recording, and transmitting the information from one or more facial images or a short video stream of facial images,
- specifies scene constraints of the facial images,
- specifies photographic properties of the facial images,
- specifies digital image attributes of the facial images,
- provides best practices for the photography of faces.

### 2 Conformance

A biometric data record conforms to this part of ISO/IEC 19794 if it satisfies all of the normative requirements related to:

A) Its data structure, data values and the relationships between its data elements, as specified in Clauses 5, 6, 7, 8 for the Full Frontal Face Image Type, Clauses 5, 6, 7, 9 for the Token Frontal Image Type, and Clauses 5, 6, 7, 10 for the Post-processed Frontal Image Type of this part of ISO/IEC 19794, respectively.

B) The relationship between its data values and the input biometric data from which the biometric data record was generated, as specified in Clauses 5, 6, 7, 8 for the Full Frontal Face Image Type, Clauses 5, 6, 7, 9 for the Token Frontal Image Type, and Clauses 5, 6, 7, 10 for the Post-processed Frontal Image Type of this part of ISO/IEC 19794, respectively.

A system that produces biometric data records is conformant to this part of ISO/IEC 19794 if all biometric data records that it outputs conform to this part of ISO/IEC 19794 (as defined above) as claimed in the Implementation Conformance Statement associated with that system. A system does not need to be capable of producing biometric data records that cover all possible aspects of this part of ISO/IEC 19794, but only those that are claimed to be supported by the system in the Implementation Conformance Statement.

A system that uses biometric data records is conformant to this part of ISO/IEC 19794 if it can read, and use for the purpose intended by that system, all biometric data records that conform to this part of ISO/IEC 19794 (as defined above) as claimed in the Implementation Conformance Statement associated with that system. A system does not need to be capable of using biometric data records that cover all possible aspects of this part of ISO/IEC 19794, but only those that are claimed to be supported by the system in an Implementation Conformance Statement.

### 3 Normative references

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

ISO/IEC 10918-1, *Information technology — Digital compression and coding of continuous-tone still images: Requirements and guidelines*

ISO/IEC 14496-2:2004, *Information technology — Coding of audio-visual objects — Part 2: Visual*

ISO/IEC 15444-1, *Information technology — JPEG 2000 image coding system: Core coding system*

ISO/IEC 15948, *Information technology — Computer graphics and image processing — Portable Network Graphics (PNG): Functional specification*

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

ISO/IEC 29794-1, *Information technology — Biometric sample quality — Part 1: Framework*

### 4 Terms and definitions

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

#### 4.1

##### **2D image**

two-dimensional face representation that encodes the luminance and/or colour texture of a capture subject in a given lighting environment

#### 4.2

##### **3D image**

three-dimensional face representation that encodes a surface in a 3D space

#### 4.3

##### **3D point map**

3D point cloud representing a capture subject, where each surface point is encoded with a triplet, representing the x, y and z value of the point in 3D respectively

#### 4.4

##### **3D vertex representation**

representation using 3D vertices and triangles between these points for coding of a 3D surface

#### 4.5

##### **anthropometric landmark**

landmark point on the face used for identification and classification of humans

#### 4.6

##### **anthropometric landmark code**

two-part code that defines an anthropometric landmark uniquely

#### 4.7

##### **Cartesian coordinate system**

3D orthogonal coordinate system

#### 4.8

##### **chin**

central forward portion of the lower jaw

**4.9****colour image**

continuous-tone image that has more than one channel, each of which is coded with one or multiple bits

**4.10****colour space**

way of representing colours of pixels in an image

EXAMPLE RGB, YUV and greyscale colour spaces are typically used in this part of ISO/IEC 19794.

**4.11****crown**

top of the head, or (if obscured by hair or headwear) where the top of the head/skull would be if it could be seen

**4.12****cylindrical coordinate system**

three-dimensional polar coordinate system describing a point by the three components radius, azimuth and height

**4.13****dots per inch (DPI)**

measurement of scanner and printer spatial sampling rate

**4.14****facial image**

electronic image-based representation of the portrait of a person

**4.15****face image type**

category of facial images that satisfy specific requirements

**4.16****feature point**

reference point in a face image as used by face recognition algorithms

NOTE Commonly referred to as a landmark, an example being the position of the eyes.

**4.17****fish eye**

type of distortion where central objects of the image erroneously appear closer than those at the edge

**4.18****greyscale image**

continuous-tone image encoded with one luminance channel

NOTE For example, if the luminance channel is coded with 8 bits, the greyscale image is also referred to as a monochrome or black and white image.

**4.19****human examination**

process of careful human comparison of a face image with a person or another face image to ascertain the identity of the respective person by a detailed examination of facial characteristics and structures

**4.20****human verification**

validation of the identity of a face image by means of comparison with a person or other face image

NOTE Also known as one-to-one (1:1) comparison.

#### 4.21

##### **JPEG**

image compression standard specified as ISO/IEC 10918-1

NOTE The JPEG baseline standard was published as ISO/IEC 10918-1 and ITU-T Rec. T.81.

#### 4.22

##### **JPEG2000**

image compression standard specified as ISO/IEC 15444-1

NOTE The JPEG2000 baseline standard was published as ISO/IEC 15444-1 and ITU-T Rec. T.800.

#### 4.23

##### **PNG format**

lossless image compression standard specified in ISO/IEC 15948

#### 4.24

##### **portrait**

photograph of a person which includes the full head, with all hair in most cases, as well as neck and top of shoulders

#### 4.25

##### **range image**

numerical matrix that encodes a surface point in 3D space, where the position encodes the first two coordinates and the value at that position encodes the third coordinate

#### 4.26

##### **red-eye**

red glow from a subject's eye caused by light from flash reflecting from blood vessels behind the retina

#### 4.27

##### **texture**

two-dimensional representation of the luminance and/or colour of a capture subject in a given lighting environment

#### 4.28

##### **texture projection matrix**

3x4 matrix to transform a 3D surface coordinate from a metric Cartesian coordinate system to a 2D texture image coordinate, where the transformation makes use of the 3D homogeneous coordinates of the 3D point as well as the 2D homogeneous coordinates of the 2D point

NOTE See reference [2] for details.

## 5 The Face Image Data Record Format

### 5.1 Overview

The ISO/IEC 19794-5 BDIR format specified in this part of ISO/IEC 19794 is a format to store face representations within a biometric data record. Each BDIR shall pertain to a single subject and shall contain at least one or more 2D image and zero or more geometric representations (range images, 3D point maps, 3D vertex representations) of a human face. Depending on the face image type, a 3D representation of a face may be included in addition to the 2D representation. The record structure is depicted in Figure 2, and Figure 3.

Adherence to this format requires compliance to the standards referred to above. In particular, 2D image data will be encoded using JPEG, JPEG2000 or PNG.

When referring to elements of the record format, “field” denotes the singular element such as Face Image Type and Image Data Type, “block” denotes the group of fields such as Facial Information block or Image Information block, and “record” denotes the data that consists of the General Header and one or more Representations.

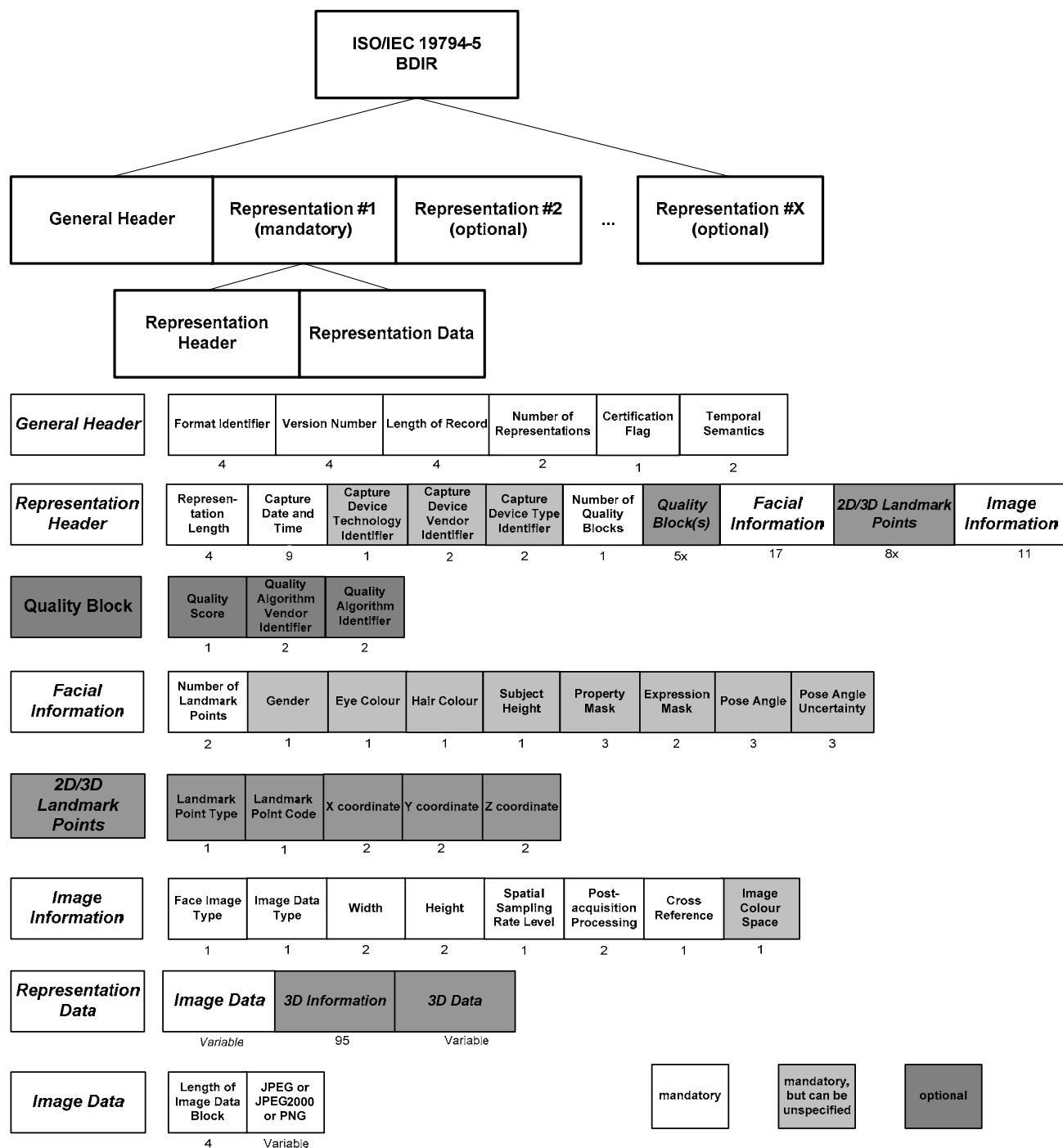


Figure 2 — The ISO/IEC 19794-5 Biometric Data Interchange Record. The length value of each field in bytes is shown below the field. The white boxes indicate fields or blocks that shall be specified, and dark grey boxes indicate optional fields.

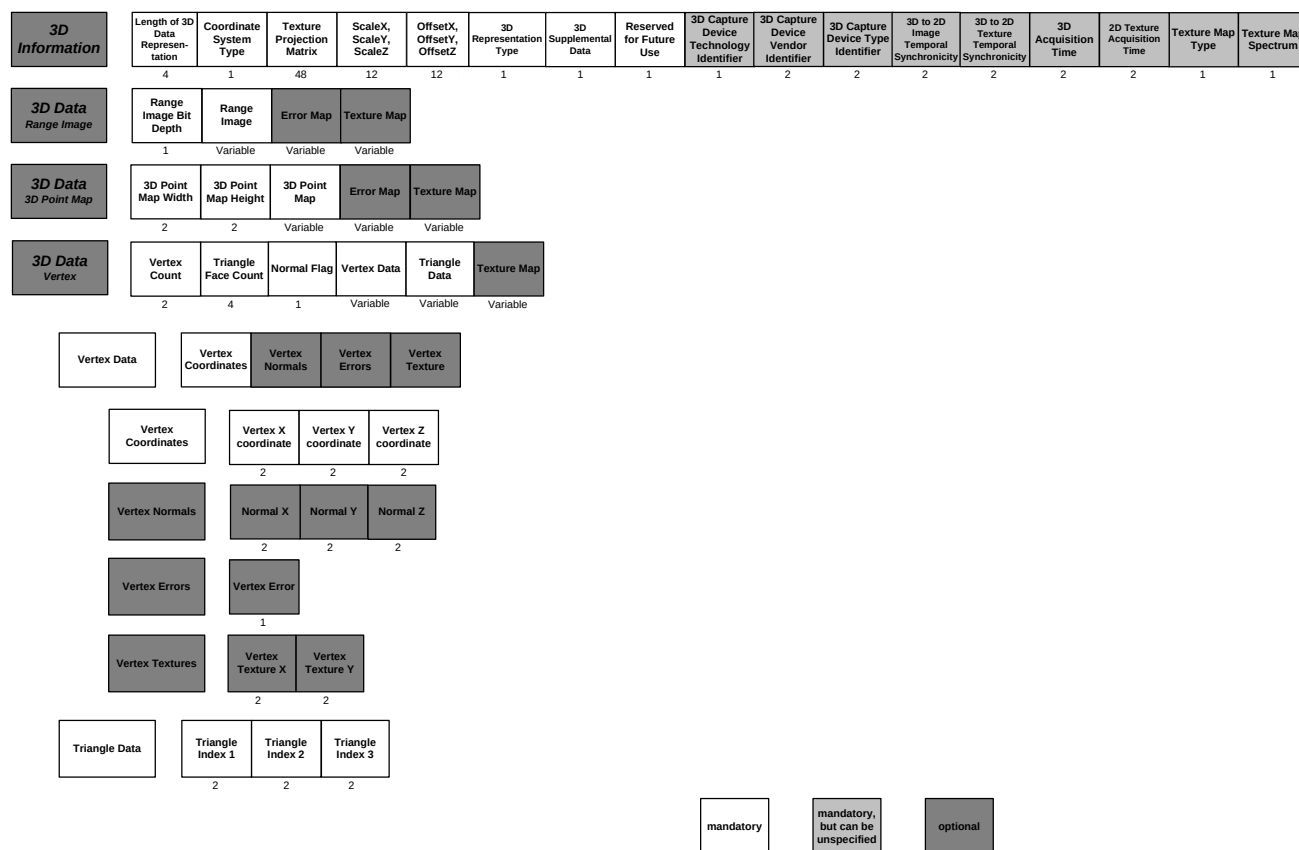


**Figure 3 — The 3D Information block and the three possible 3D Data blocks specified in this part of ISO/IEC 19794. The length value of each field in bytes is shown below the field. The white boxes indicate fields or blocks that shall be specified, light grey boxes that the fields are mandatory, but an unspecified value is acceptable, and dark grey boxes indicate optional fields.**

With the exception of the Format Identifier and the Version Number for the standard, which are null-terminated ASCII character strings, all data is represented in binary format.

There are no record separators or field tags; fields are parsed by byte count.

The organization of the record format is as follows:

- A fixed-length (17 byte) **General Header** containing information about the overall record, including the number of facial images represented and the overall record length in bytes.
- A Representation block for each facial representation. This data consists of a Representation Header and the Representation Data.
- The Representation Header consists of
  - o A fixed length (19 bytes) common elements defined in ISO/IEC 19794-1:2011
  - o Multiple (including none) fixed length (5 byte) **Quality** blocks describing the quality of the representation.
  - o A fixed length (17 byte) **Facial Information** block describing discernable characteristics of the subject such as gender.
  - o Multiple (including none) fixed length (8 byte) **Landmark Point** blocks describing Landmark Points in a facial image.
  - o A fixed length (11 byte) **Image Information** block describing digital properties of the image such as Face Image Type and dimensions such as width and height.

- The Representation Data consists of
  - o **Image data** consisting of a JPEG, JPEG2000 or PNG encoded data block.
  - o For Face Image Types containing 3D information a **3D Information** block (95 byte) describing properties of this data.
  - o For Face Image Types containing 3D information the **3D Data** block describing the 3D shape of the face.

Multiple 2D / 3D representations of the same biometric data subject can be described in a single record. This is accomplished by including multiple representation blocks after the General Header block. Representation blocks containing 2D data can be stored together with Representation blocks also containing 3D data. The structure of this embedding is illustrated in Figure 4.

|                |                                        |                                         |                                                 |
|----------------|----------------------------------------|-----------------------------------------|-------------------------------------------------|
| General Header | Representation<br>#1<br>(Full Frontal) | Representation<br>#2<br>(Token Frontal) | Representation<br>#3<br>(Full Frontal<br>Range) |
|----------------|----------------------------------------|-----------------------------------------|-------------------------------------------------|

Figure 4 — Embedding multiple 2D / 3D representations in the same record

## 5.2 Data Conventions

### 5.2.1 Byte ordering

Within the record format and all well-defined data blocks therein, all multi-byte quantities are stored in Big-Endian format. That is, the more significant bytes of any multi-byte quantity are stored at lower addresses in memory than less significant bytes. For example, the value 1 025 (2 to the 10<sup>th</sup> power plus one) would be stored as first byte= 00000100 and second byte=00000001.

### 5.2.2 Numeric values

All numeric values are fixed-length unsigned integer quantities, unless otherwise specified.

### 5.2.3 Conversion to integer

The conversion of a numeric value to integer is given by rounding down if the fractional portion is less than 0,5 and rounding up if the fractional value is greater than or equal to 0,5.

### 5.2.4 Unspecified field value

In Figure 2 and Figure 3 the following fields are mandatory, but the value of the field can indicate that the field is unspecified: Capture Device Technology Identifier, Capture Device Vendor Identifier, Capture Device Type Identifier, Gender, Eye Colour, Hair Colour, Subject Height, Property, Expression, Pose Angle, Pose Angle Uncertainty, Image Colour Space, 3D Capture Device Technology Identifier, 3D Capture Device Vendor Identifier, 3D Capture Device Type Identifier, 3D to 2D Image Temporal Synchronicity, 3D to 2D Texture Temporal Synchronicity, 3D Acquisition Time, 2D Texture Acquisition Time, Texture Map Type, and Texture Map Spectrum.

### 5.2.5 Unknown field value

A field value labelled by the identifier “Unknown” shall be used to denote that the information encoded by the field cannot be determined by examination of the face image.

### 5.3 The General Header

#### 5.3.1 Structure

The General Header block consists of seven fields; Format Identifier, Version Number, Length of Record, Number of Representations, Capture Device Vendor Identifier, Capture Device Type Identifier and the Temporal Semantics field as shown in Table 2.

**Table 2 — The General Header**

| Field                     | Size    | Valid values                                             | Notes                                                                                                           |
|---------------------------|---------|----------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| Format Identifier         | 4 bytes | 46414300 <sub>HEX</sub> ('F' 'A' 'C' 0 <sub>HEX</sub> )  | Indicates face representation data                                                                              |
| Version Number            | 4 bytes | 30333000 <sub>HEX</sub> ('0' '3' '0' 00 <sub>HEX</sub> ) | "030" in ASCII                                                                                                  |
| Length of Record          | 4 bytes | $68 \leq \text{Length of Record} \leq 2^{32} - 1$        | Includes Facial Record Header and Facial Record Data. The minimum of 68 bytes includes the smallest JPEG image. |
| Number of Representations | 2 byte  | $1 \leq \text{Number} \leq 65535$                        | See 5.3.5                                                                                                       |
| Certification Flag        | 1 byte  | 00 <sub>HEX</sub><br>01 <sub>HEX</sub>                   | See 5.3.6                                                                                                       |
| Temporal Semantics        | 2 bytes | $0 \leq \text{Number} \leq 65535$                        | See 5.3.7 and Table 3                                                                                           |

#### 5.3.2 Format Identifier

The format identifier shall be recorded in four bytes. The format identifier shall consist of three characters "FAC" followed by a zero byte as a NULL string terminator.

#### 5.3.3 Version Number

The number for the version of ISO/IEC 19794-5 used for constructing the BDIR shall be placed in four bytes. This version number shall consist of three ASCII numerals followed by a zero byte as a NULL string terminator. The first and second character will represent the major version number and the third character will represent the minor revision number. The Version Number of ISO/IEC 19794-5:2010 shall be 30333000<sub>HEX</sub>; "030" – Version 3 revision 0.

"NOTE In 2009 ISO/IEC 19794-5:2005 AMD 2 was published, which defined that interchange records that are conformant to that amendment shall contain the version number "020".

#### 5.3.4 Length of Record

The length (in bytes) of the entire BDIR shall be recorded in four bytes. This count shall be the total length of the BDIR including the general record header and one or more representation records.

#### 5.3.5 Number of Representations

The total number of representation records contained in the BDIR shall be recorded in two bytes. A minimum of one representation is required.

#### 5.3.6 Certification Flag

The value shall be 00<sub>Hex</sub>.

NOTE No certification schemes are available for this part of ISO/IEC 19794.

#### 5.3.7 Temporal Semantics

This two byte (2 byte) field shall be assigned according to Table 3. This supports storage of multiple representations: from a single session (e.g. from a photo shoot); from distinct sessions (e.g. from cash

dispenser transactions); and from a temporal sequence (e.g. a video sequence of equally time-spaced representations).

**Table 3 — Temporal sequence flags and values**

| Description                                                                                                                                                                                                                                                                                                                            | Value                                                    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|
| One representation is present                                                                                                                                                                                                                                                                                                          | 0000 <sub>HEX</sub>                                      |
| Two or more representations are present. The temporal relationship between them is unspecified.                                                                                                                                                                                                                                        | 0001 <sub>HEX</sub>                                      |
| Two or more representations are present. The representations are taken at irregular intervals during a single session (e.g. images taken by a photographer during a shoot.)                                                                                                                                                            | 0002 <sub>HEX</sub>                                      |
| Two or more representations are present. The representations are taken at irregular intervals spanning multiple sessions (e.g. a lifetime of passport photo submissions). Unless representation qualities are known, or some other appropriate ordering mechanism is available, the most recent representation should be stored first. | 0003 <sub>HEX</sub>                                      |
| The number of milliseconds between sequential representations (e.g. the frames extracted from a video sequence).                                                                                                                                                                                                                       | 0004 <sub>HEX</sub> ≤<br>Number <<br>FFFF <sub>HEX</sub> |
| The representations correspond to a temporal sequence but are taken at regular intervals exceeding FFFE <sub>HEX</sub> milliseconds.                                                                                                                                                                                                   | FFFF <sub>HEX</sub>                                      |

NOTE 1 The minimum interval is  $4 * 0,001$  seconds = 0,004 seconds.

NOTE 2 The maximum interval is  $65\,534 * 0,001$  seconds = 65,534 seconds.

NOTE 3 While a proper video encoding, e.g. ITU-T H264, would offer a more compact encoding of the same sequence, it would add a decoding overhead and is anyway not supported by this part of ISO/IEC 19794.

## 5.4 The Representation Header

### 5.4.1 Structure

The Representation Header is intended to describe discrete properties of the individual discernable from the image, one is included for each facial representation included in the record. The structure of this block is shown in Figure 2.

The Representation Header consists of the Representation Length, the Capture Date and Time, the Capture Device Technology Identifier, the Capture Device Vendor Identifier. These are followed by the Number of Quality Blocks field and the related number of Quality blocks. Finally the Representation Header contains the Facial Information block, the optional multiple Landmark Point blocks, and the Image Information block.

### 5.4.2 Representation Length

The (4 byte) Representation Length field denotes the length in bytes of the representation including the representation header fields.

The minimum value of the Representation Length is 51 bytes, consisting of a minimum 47 bytes for the Representation Header plus the size of the Representation Data, i.e. minimum 4 bytes for the Length of Image Data Block field assuming 0 bytes for the variable data.

### 5.4.3 Capture Date and Time

The capture date and time field shall indicate when the capture of this representation started in Coordinated Universal Time (UTC). The capture date and time field shall consist of 9 bytes. Its value shall be encoded in the form given in ISO/IEC 19794-1.

### 5.4.4 Capture Device Technology Identifier

Capture device technology Identifier shall be encoded in one byte. This field shall indicate the class of device technology used to acquire the captured biometric sample. A value of 00<sub>HEX</sub> indicates unknown or unspecified

technology A value of 0 indicates “unknown or unspecified”. See Table 4 for the enumerated list of possible values.

Many different types of capture devices work in the visible spectrum or in near infra-red (NIR). To indicate that the capture device operates in NIR the highest bit in the Capture Device Technology Identifier field shall be set to 1. So, if the representation is a Video frame from a digital video camera operating in the visible spectrum, the resulting value of the Capture Device Technology Identifier field shall be 06<sub>HEX</sub>, if it operates in NIR, the code shall be 86<sub>HEX</sub>.

**Table 4 — Capture Device Technology Identifier codes**

| Description                                         | Value                                  |
|-----------------------------------------------------|----------------------------------------|
| Unknown or Unspecified                              | 00 <sub>HEX</sub>                      |
| Static photograph from an unknown source            | 01 <sub>HEX</sub>                      |
| Static photograph from a digital still-image camera | 02 <sub>HEX</sub>                      |
| Static photograph from a scanner                    | 03 <sub>HEX</sub>                      |
| Video frame(s) from an unknown source               | 04 <sub>HEX</sub>                      |
| Video frame(s) from an analogue video camera        | 05 <sub>HEX</sub>                      |
| Video frame(s) from a digital video camera          | 06 <sub>HEX</sub>                      |
| Reserved by SC37 for future use                     | 07 <sub>HEX</sub> to 7F <sub>HEX</sub> |
| Vendor specific                                     | 80 <sub>HEX</sub> to FF <sub>HEX</sub> |

#### 5.4.5 Capture Device Vendor Identifier

The (2 byte) Capture Device Vendor Identifier shall identify the biometric organisation that owns the product that created the BDIR. The capture device algorithm vendor identifier shall be encoded in two bytes carrying a CBEFF biometric organization identifier (registered by IBIA or other approved registration authority). A value of all zeros shall indicate that the capture device vendor is unreported.

#### 5.4.6 Capture Device Type Identifier

The (2 byte) Capture Device Type Identifier shall identify the product type that created the BDIR. It shall be assigned by the registered product owner or other approved registration authority. A value of all zeros shall indicate that the capture device type is unreported. If the capture device vendor identifier is 0000<sub>HEX</sub>, then also the capture device type identifier shall be 0000<sub>HEX</sub>.

#### 5.4.7 Number of Quality Blocks

This field is followed by the number of 5 byte Quality blocks reflected by its value. A value of zero (0) means that no attempt was made to assign a quality score. In this case, no Quality blocks are present.

NOTE According to ISO/IEC 19794-1:2011, 12.3.1, multiple Quality Blocks are considered as a Quality Record.

#### 5.4.8 Quality Score

The (1 byte) Quality Score, as defined in ISO/IEC 29794-1, shall be a quantitative expression of the predicted verification performance of the biometric sample. Valid values for Quality Score are integers between 0 and 100, where higher values indicate better quality. A value of 255 is to handle a special case. An entry of 255 shall indicate a failed attempt to calculate a quality score. This value of Quality Score is harmonized with ISO/IEC 19784-1, where 255 is -1.

NOTE BioAPI, unlike ISO/IEC 19794 uses signed integers

#### 5.4.9 Quality Algorithm Vendor Identifier

To enable the recipient of the quality score to differentiate between quality scores generated by different algorithms, the provider of quality scores shall be uniquely identified by this two-byte field. This is registered with the IBIA or other approved registration authority.

#### 5.4.10 Quality Algorithm Identifier

The (2 byte) Quality Algorithm Identifier specifies an integer product code assigned by the vendor of the quality algorithm. It indicates which of the vendor's algorithms (and version) was used in the calculation of the quality score and should be within the range 1 to 65 535.

NOTE Multiple quality scores calculated by the same algorithm (same vendor Identifier and algorithm Identifier) shall not be present in a single representation.

Table 5 summarizes the quality field. All values are fixed-length unsigned integer quantities represented in Big-Endian format.

**Table 5 — Structure of Quality blocks**

| Description              |                                     | Length  | Valid values       | Note                                                                                                                                                                                                                    |
|--------------------------|-------------------------------------|---------|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Number of Quality Blocks |                                     | 1 byte  | [0,255]            | This field is followed by the number of 5-byte Quality Blocks reflected by its value.<br><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. |
| Quality Block            | Quality Score                       | 1 byte  | [0,100]<br><br>255 | 0: lowest<br><br>100: highest<br><br>255: failed attempt to assign a quality score                                                                                                                                      |
|                          | Quality algorithm vendor Identifier | 2 bytes | [1,65535]          | Quality Algorithm Vendor Identifier shall be registered with IBIA or other approved registration authority as a CBEFF biometric organization. Refer to CBEFF vendor Identifier registry procedures in ISO/IEC 19785-2.  |
|                          | Quality algorithm Identifier        | 2 bytes | [1,65535]          | Quality Algorithm Identifier may be optionally registered with IBIA or other approved registration authority as a CBEFF Product Code. Refer to CBEFF product registry procedures in ISO/IEC 19785-2.                    |

## 5.5 The Facial Information Block

### 5.5.1 Structure

The Facial Information Block consists of the Number of Landmark Points, the Gender, the Eye Colour, the Hair Colour, the Subject Height, the Property Mask, the Expression Mask, the Pose Angle, and the Pose Angle Uncertainty fields.

### 5.5.2 Number of Landmark Points

The (2 byte) Number of Landmark Points field shall be the number of Landmark Point blocks that follow the Facial Information block. The Landmark Point block is defined in clause 5.6.

### 5.5.3 Gender

The (1 byte) Gender field shall represent the gender of the subject according to Table 6.

**Table 6 — Gender codes**

| Description | Value             |
|-------------|-------------------|
| Unspecified | 00 <sub>HEX</sub> |
| Male        | 01 <sub>HEX</sub> |
| Female      | 02 <sub>HEX</sub> |
| Unknown     | FF <sub>HEX</sub> |

#### 5.5.4 Eye Colour

The (1 byte) Eye Colour field shall represent the colour of irises of the eyes according to Table 7. If the eyes are different colours, then the right eye colour is to be encoded.

**Table 7 — Eye Colour codes**

| Description                                                               | Value                                  |
|---------------------------------------------------------------------------|----------------------------------------|
| Unspecified                                                               | 00 <sub>HEX</sub>                      |
| Black                                                                     | 01 <sub>HEX</sub>                      |
| Blue                                                                      | 02 <sub>HEX</sub>                      |
| Brown                                                                     | 03 <sub>HEX</sub>                      |
| Gray                                                                      | 04 <sub>HEX</sub>                      |
| Green                                                                     | 05 <sub>HEX</sub>                      |
| Multi-Coloured                                                            | 06 <sub>HEX</sub>                      |
| Pink <sup>1</sup>                                                         | 07 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use                                          | 08 <sub>HEX</sub> to FE <sub>HEX</sub> |
| Other or Unknown (e.g. cannot be determined from image, monochrome image) | FF <sub>HEX</sub>                      |

#### 5.5.5 Hair Colour

The (1 byte) Hair Colour field shall represent the colour of the hair according to Table 8.

**Table 8 — Hair Colour codes**

| Description                      | Value                                  |
|----------------------------------|----------------------------------------|
| Unspecified                      | 00 <sub>HEX</sub>                      |
| Bald                             | 01 <sub>HEX</sub>                      |
| Black                            | 02 <sub>HEX</sub>                      |
| Blonde                           | 03 <sub>HEX</sub>                      |
| Brown                            | 04 <sub>HEX</sub>                      |
| Gray                             | 05 <sub>HEX</sub>                      |
| White                            | 06 <sub>HEX</sub>                      |
| Red                              | 07 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 08 <sub>HEX</sub> to FE <sub>HEX</sub> |
| Unknown or Other                 | FF <sub>HEX</sub>                      |

#### 5.5.6 Subject Height

The (1 byte) Subject Height field shall represent the height of the subject according to Table 9.

**Table 9 — Subject Height codes**

| Description          | Value                                  |
|----------------------|----------------------------------------|
| Unspecified          | 00 <sub>HEX</sub>                      |
| Subject Height in cm | 01 <sub>HEX</sub> to FF <sub>HEX</sub> |

#### 5.5.7 Property Mask

The (3 byte) Property Mask is a bit mask of 3 bytes and each bit of the mask position listed in Table 10 shall be set to 1 if the corresponding property is present, and set to 0 if absent. The mask position starts from 0 at the lowest bit. The lowest bit set to 0 shall indicate that properties are not specified (and all bits shall be zero);

<sup>1</sup> Very rare, associated to albinism. Very small minority of individuals world wide are affected.

the lowest bit set to 1 shall indicate that all listed properties have been considered and that a zero value of any property bit indicates an absence of that property.

All reserved bits shall be zero.

**Table 10 — Property flags**

| Description                                                                       | Mask Position |
|-----------------------------------------------------------------------------------|---------------|
| Properties are specified                                                          | 0             |
| Glasses                                                                           | 1             |
| Moustache                                                                         | 2             |
| Beard                                                                             | 3             |
| Teeth visible                                                                     | 4             |
| Pupil or iris not visible (e.g. either or both eyes closed or half closed)        | 5             |
| Mouth open                                                                        | 6             |
| Left Eye Patch                                                                    | 7             |
| Right Eye Patch                                                                   | 8             |
| Dark Glasses (medical)                                                            | 9             |
| Head coverings present                                                            | 10            |
| Feature Distorting Medical Condition (which could impact Feature Point detection) | 11            |
| Reserved by SC 37 for future use                                                  | 12 to 23      |

A “Pupil or iris not visible” flag set to “1” shall indicate non-compliance with the Frontal, Full Frontal, and Token image types.

### 5.5.8 Expression

The (2 byte) Expression Mask is a bit mask of 2 bytes and each bit of the mask position listed in Table 11 shall be set to 1 if the corresponding expression is present, and set to 0 if absent. The mask position starts from 0 at the lowest bit. The lowest bit set to 0 shall indicate that properties are not specified (and all bits shall be zero); the lowest bit set to 1 shall indicate that all listed properties have been considered and that a zero value of any property bit indicates an absence of that expression.

All reserved bits shall be zero.

**Table 11 — Expression flags**

| Description                                                | Mask Position |
|------------------------------------------------------------|---------------|
| Expressions are specified                                  | 0             |
| Neutral (non-smiling) with both eyes open and mouth closed | 1             |
| Smile                                                      | 2             |
| Raised eyebrows                                            | 3             |
| Eyes looking away from the camera                          | 4             |
| Squinting                                                  | 5             |
| Frowning                                                   | 6             |
| Reserved by SC 37 for future use                           | 7 to 11       |
| Vendor specific                                            | 12 to 15      |

## 5.5.9 Pose Angle

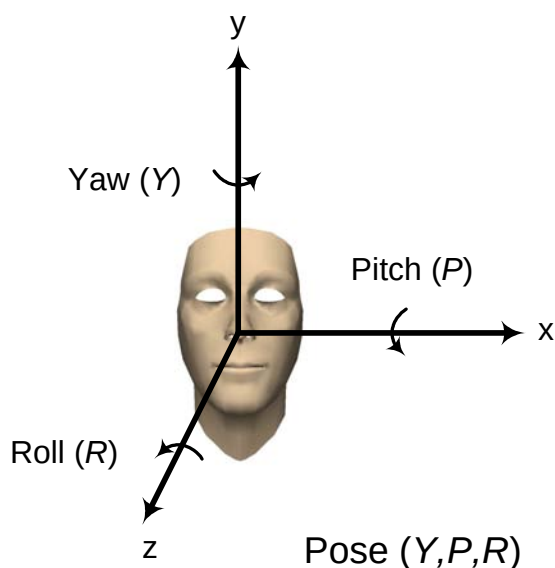


**Figure 5 — The definition of pose angles is with respect to the frontal view of the subject**

The (3 multi-byte) Pose Angle field ( $B_Y$ ,  $B_P$ ,  $B_R$ ) shall represent the estimate or measure pose of the subject in the image. Each byte in the field respectively represents pose angles of yaw, pitch and roll in that order. The pose angle is given by Tait-Bryan angles.

- Yaw angle: rotation about the vertical (y) axis.
- Pitch angle: rotation about the horizontal side-to-side (x) horizontal axis.
- Roll angle: rotation about the horizontal back to front (z) axis.

The angles are defined relative to the frontal pose of the subject, which has angles (0,0,0) as shown in Figure 5. The frontal pose is defined by the Frankfurt Horizon FH (see Annex E) as the xz plane and the vertical symmetry plane as the yz plane with the z axis oriented in the direction of the face sight. The examples are shown in Figure 6.

As order of the successive rotation around the different axes does matter, the encoded rotation angle shall correspond to an order of execution starting from the frontal view. This order shall be given by Roll (about the front axis), then Pitch (about the horizontal axis) and finally Yaw (about the vertical axis). The (first executed) Roll transformation will therefore always be in the image (x,y) plane.

From the point of view of executing a transformation from the observed view to a frontal view, the transformation order will therefore be Yaw, Pitch, and then Roll. Note however that the encoded angle is from the frontal view to the observed view.

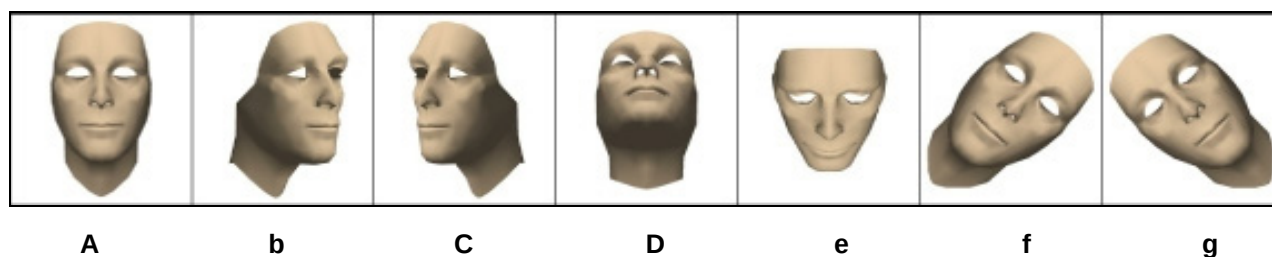


**Figure 6 — Examples of pose angles and their encodings. The pose angles (Y, P, R) of figures a – g are given by (0, 0, 0), (+45, 0, 0), (-45, 0, 0), (0, -45, 0), (0, +45, 0), (0, 0, -45), and (0, 0, +45), respectively. The pose angle encodings, ( $B_Y$ ,  $B_P$ ,  $B_R$ ) as defined in clauses 5.5.9.1 - 5.5.9.3 are subsequently given by (1, 1, 1), (23, 1, 1), (158, 1, 1), (1, 158, 1), (1, 23, 1), (1, 1, 158), and (1, 1, 23), respectively**

### 5.5.9.1 Pose Angle – Yaw

The yaw angle  $Y$  is the rotation in degrees about the y-axis (vertical axis) shown in Figure 5. Frontal poses have a yaw angle of 0 degrees. Positive angles represent faces looking to their left (a counter-clockwise rotation around the y-axis).

The encoding  $B_Y$  of the yaw angle  $Y$  shall be in degrees as a byte (1 byte) with values from  $-180$  to  $180$  modulo 2.

- If  $180 \geq Y \geq 0$  then  $B_Y = Y/2 + 1$ . The remainder is discarded.
- If  $-180 \leq Y < 0$  then  $B_Y = 181 + Y/2$ . The remainder is discarded.

The maximum value of  $B_Y$  is 180. If the pose angle is not specified, the value of  $B_Y$  shall be 0.

### 5.5.9.2 Pose Angle – Pitch

The pitch angle  $P$  is the rotation in degrees about the x-axis (horizontal axis) shown in Figure 5. Frontal poses have a pitch angle of 0 degrees. Positive angles represent faces looking down (a counter-clockwise rotation around the x-axis).

The encoding  $B_P$  of the pitch angle  $P$  shall be in degrees as a byte (1 byte) with values from  $-180$  to  $180$  modulo 2.

- If  $180 \geq P \geq 0$  then  $B_P = P/2 + 1$ . The remainder is discarded.
- If  $-180 \leq P < 0$  then  $B_P = 181 + P/2$ . The remainder is discarded.

The maximum value of  $B_P$  is 180. If the pitch angle is not specified, the value of  $B_P$  shall be 0.

Note that the definition of zero degree pitch is not obvious.

### 5.5.9.3 Pose Angle – Roll

The roll angle  $R$  is the rotation in degrees about the z-axis (the horizontal axis from front to back) shown in Figure 5. Frontal poses have a roll angle of 0 degrees. Positive angles represent faces tilted toward their right shoulder (counter-clockwise rotation around the z-axis). A roll angle of 0 degrees denotes that the left and right eye centres have identical Y coordinates.

The encoding  $B_R$  of the roll angle  $R$  shall be in degrees as a byte (1 byte) with values from  $-180$  to  $180$  modulo 2.

- If  $180 \geq R \geq 0$  then  $B_R = R/2 + 1$ . The remainder is discarded.
- If  $-180 \leq R < 0$  then  $B_R = 181 + R/2$ . The remainder is discarded.

The maximum value of  $B_R$  is 180. If the roll angle is not specified, the value of  $B_R$  shall be 0.

### 5.5.10 Pose Angle Uncertainty

The (3 multi-byte) Pose Angle Uncertainty ( $U_Y$ ,  $U_P$ ,  $U_R$ ) represents the expected degree of uncertainty of the pose angle yaw, pitch, and roll. Each byte in the field respectively represents the uncertainty of yaw, pitch and roll in that order. The uncertainty is allowed to represent experimental uncertainty specified by each vendor.

The encoding of Pose Angle Uncertainty is given by three bytes ( $U_Y$ ,  $U_P$ ,  $U_R$ ) where each byte  $U_k$  in the field ( $k=Y,P,R$ ) represents 1 degree of uncertainty with minimum and maximum values of 1 and 181 where  $U_k = (\text{uncertainty} + 1)$ . The more uncertain, the value of the uncertainty  $U_k$  shall become larger. If the uncertainty is not specified, then the values of  $U_Y$ ,  $U_P$  and  $U_R$  shall be set to zero (0).

## 5.6 The Landmark Point Block

### 5.6.1 Structure

The optional (8 byte) Landmark Point block specifies the type, code and position of a Landmark Point in the facial image. The number of Landmark Point blocks shall be specified in the Number of Landmark Points field of the Facial Information block. The structure of this block is shown in Table 13.

Landmark Points can be specified as MPEG-4 Feature Points as given by Annex C of ISO/IEC 14496-2 or anthropometric landmarks in two or three dimensions. The description of the anthropometric landmarks [10] and their relation with the set of MPEG4 feature points is discussed in clause 5.6.6 of this part of ISO/IEC 19794.

### 5.6.2 Landmark Point Type

The (1 byte) Landmark Point Type field represents the type of the Landmark Point stored in the Landmark Point block. This field shall be set to 01<sub>HEX</sub> to denote that landmark point is an MPEG4 Feature Point as given by Annex C of ISO/IEC 14496-2 and is represented by the 2D image coordinates. The field shall be set to 02<sub>HEX</sub> to denote that the landmark point is an Anthropometric 2D landmark and is represented by the 2D image coordinates. Finally, the field shall be set to 03<sub>HEX</sub> to denote that the Landmark Point is an anthropometric 3D landmark and is represented by its 3D coordinates. All other field values are reserved by SC 37 for future definition of Landmark Point Types.

**Table 12 — The Landmark Point Type**

| Description                      | Value                                  | Comment                                                                                                                                                                                                                                                                                                                                |
|----------------------------------|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| MPEG4 Feature                    | 01 <sub>HEX</sub>                      | The horizontal and vertical position of the Landmark Point are measured in pixels with values from 0 to Width-1 and Height-1, respectively. The Z coordinate field is ignored.                                                                                                                                                         |
| Anthropometric 2D landmark       | 02 <sub>HEX</sub>                      | The Landmark Point is considered as an anthropometric landmark point in the 2-D image and its coordinates are measured in pixels with values from 0 to Width-1 and Height-1, respectively. The Z coordinate field is ignored.                                                                                                          |
| Anthropometric 3D landmark       | 03 <sub>HEX</sub>                      | X coordinate, Y coordinate and Z coordinate are interpreted as 2 byte values with fixed precision of 0,02 mm ranging from -655,34 mm to 655,34 mm. The Landmark Point is considered as a 3D point in the Cartesian coordinate system.<br><br>EXAMPLE: The value of 10001 corresponds to:<br><br>655,34mm + 10001 x 0,02mm = -455,32mm. |
| Reserved by SC 37 for future use | 04 <sub>HEX</sub> to FF <sub>HEX</sub> | Reserved for future use.                                                                                                                                                                                                                                                                                                               |

### 5.6.3 Landmark Point Code

The (1 byte) Landmark Point Code field shall specify the Landmark Point that is stored in the Landmark Point block.

For the Landmark Point Type 01<sub>HEX</sub> the codes of the Landmark Points in clause 5.6.4, taken from Annex C of ISO/IEC 14496-2 and defined as MPEG4 Feature Points, or the additional eye and nostril Landmark Points in clause 5.6.5 shall be stored in this block.

If the Landmark Point Type is 02<sub>HEX</sub> or 03<sub>HEX</sub>, i.e. anthropometric 2D landmark or anthropometric 3D landmark, the codes of the Landmark Points defined in 5.6.6 shall be stored in this block.

The horizontal and vertical positions of Landmark Points are either texture image coordinates or in the Cartesian coordinate system (see clause 5.10.2.2).

**Table 13 — The Landmark Point block**

| Field                         | Size    | Value                         | Notes                                                                                                                                                                                                                                                                                                 |
|-------------------------------|---------|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Landmark Point Type           | 1 byte  | See clause 5.6.2              | Denotes the type of the Landmark Point.                                                                                                                                                                                                                                                               |
| Landmark Point Code           | 1 byte  | See clause 5.6.3              | Denotes the Landmark Point, e.g. the left eye.                                                                                                                                                                                                                                                        |
| X coordinate,<br>Y coordinate | 2 bytes | See clause 5.6.2,<br>Table 12 | Denotes the coordinate of the landmark point.<br>For Landmark Point Types 01 <sub>HEX</sub> and 02 <sub>HEX</sub> this coordinate denotes the relevant pixel count from upper left pixel starting at 0.<br>For Landmark Point Type 03 <sub>HEX</sub> the value codes the coordinate of a point in 3D. |
| Z coordinate                  | 2 bytes | See clause 5.6.2,<br>Table 12 | Denotes the Z coordinate of the landmark point.<br>For Landmark Point Type 01 <sub>HEX</sub> and Type 02 <sub>HEX</sub> this field is ignored.<br><br>For Landmark Point Type 03 <sub>HEX</sub> the value codes the Z coordinate of a point in 3D.                                                    |

#### 5.6.4 MPEG4 Landmark Points

The normative Figure 7 denotes the Landmark Point codes associated with Feature Points as given by Annex C of ISO/IEC 14496-2. Each Landmark Point Code is represented by a notation A.B using a major (A) and a minor (B) value. The encoding of the Landmark Point Code is given by the (1 byte) value of  $A \times 16 + B$ .

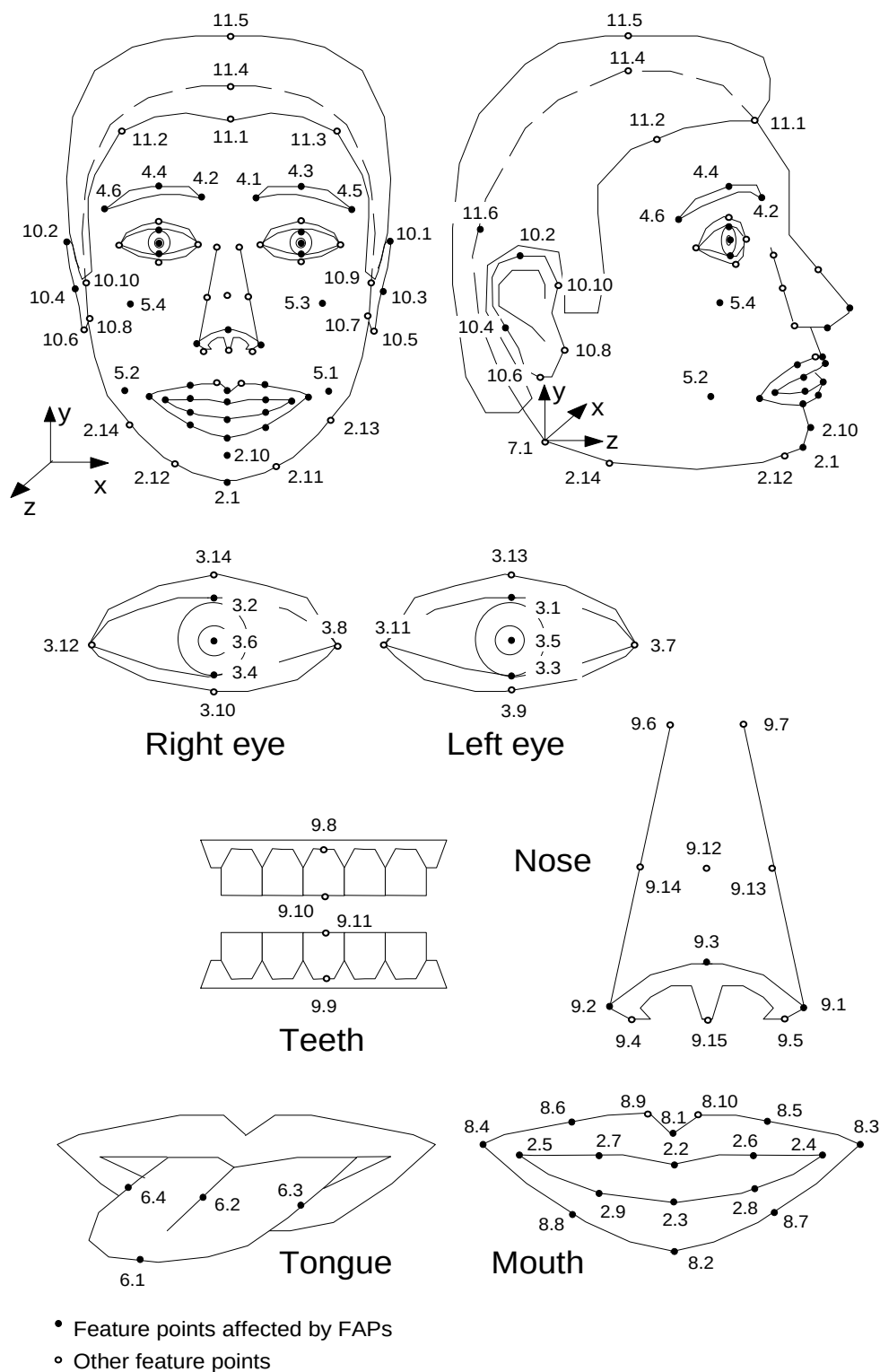


Figure 7 — Feature points affected by Face Animation Parameters specified in ISO/IEC 14496-2.

#### 5.6.5 Eye and nostril centre Landmark Points

The eye centre Landmark Points 12.1 (left) and 12.2 (right) are defined to be the horizontal and vertical midpoints of the eye corners (3.7, 3.11) and (3.8, 3.12) respectively. The left nostril centre Landmark Point 12.3 is defined to be the midpoint of the nose Landmark Points (9.1, 9.15) in the horizontal direction and (9.3, 9.15) in the vertical direction. Similarly, the right nostril centre Landmark Point 12.4 is defined to be the midpoint of the nose Landmark Points (9.2, 9.15) in the horizontal direction and (9.3, 9.15) in the vertical

direction. Both the eye centre and nostril centre Landmark Points are shown in Figure 8 and values given in Table 14.

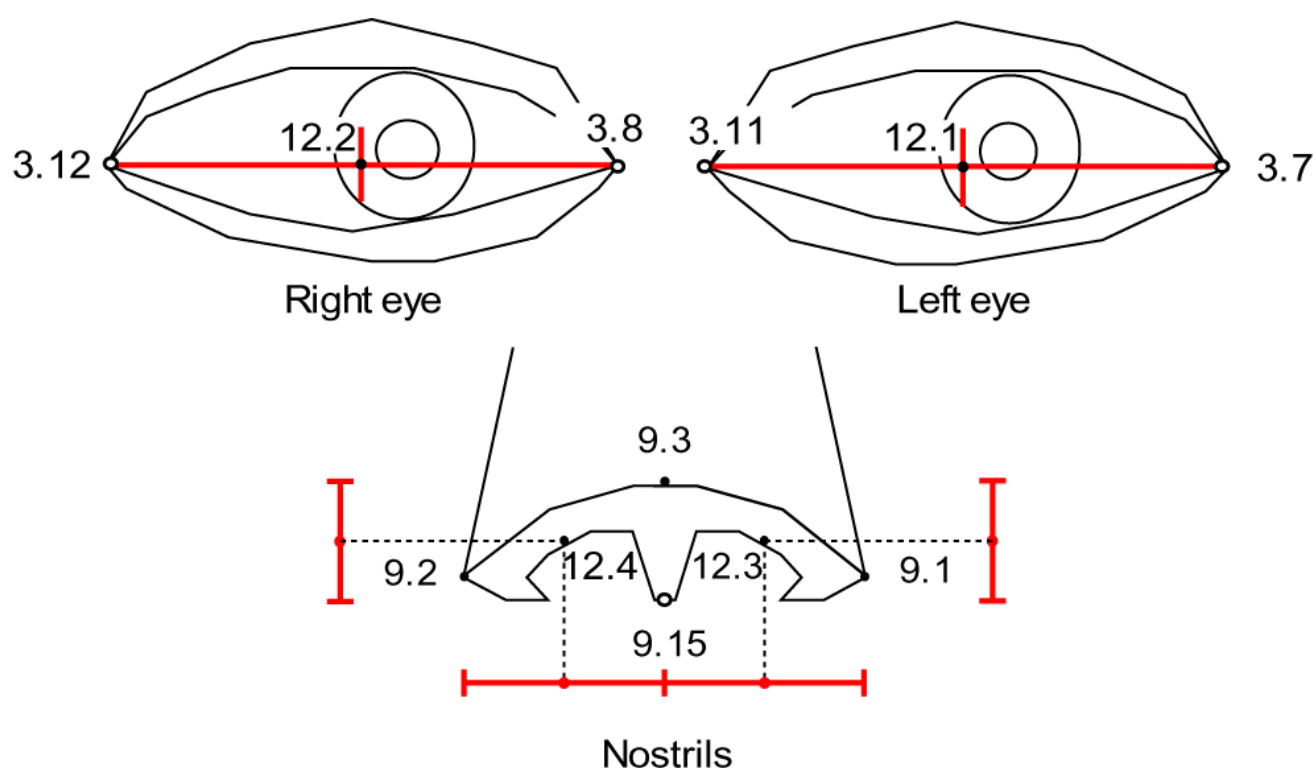


Figure 8 — The eye and nostril centre Landmark Points are defined by midpoints of MPEG4 Landmark Points

Table 14 — Eye and nostril centre Landmark Point Codes

| Centre Landmark Point | Midpoint of Landmark Points |           | Landmark Point Code |
|-----------------------|-----------------------------|-----------|---------------------|
| Left Eye              | 3.7, 3.11                   |           | 12.1                |
| Right Eye             | 3.8, 3.12                   |           | 12.2                |
| Left Nostril          | Horizontal                  | Vertical  | 12.3                |
|                       | 9.1, 9.15                   | 9.3, 9.15 |                     |
| Right Nostril         | Horizontal                  | Vertical  | 12.4                |
|                       | 9.2, 9.15                   | 9.3, 9.15 |                     |

## 5.6.6 Anthropometric Landmarks

Anthropometric landmarks extend the MPEG4 feature model with new points that are used in forensics and anthropology for person identification via two facial images or image and skull over a long time. They also allow specification of points that are in use by criminal experts and anthropologists [10].

Figure 9 and Table 15 show the definition of the anthropometric landmarks. The set of points represents the craniofacial landmark points of the head and face. The latter are used in forensics for “Face to face” and “Skull to face” identification. Some of these points have MPEG 4 counterparts, others not.

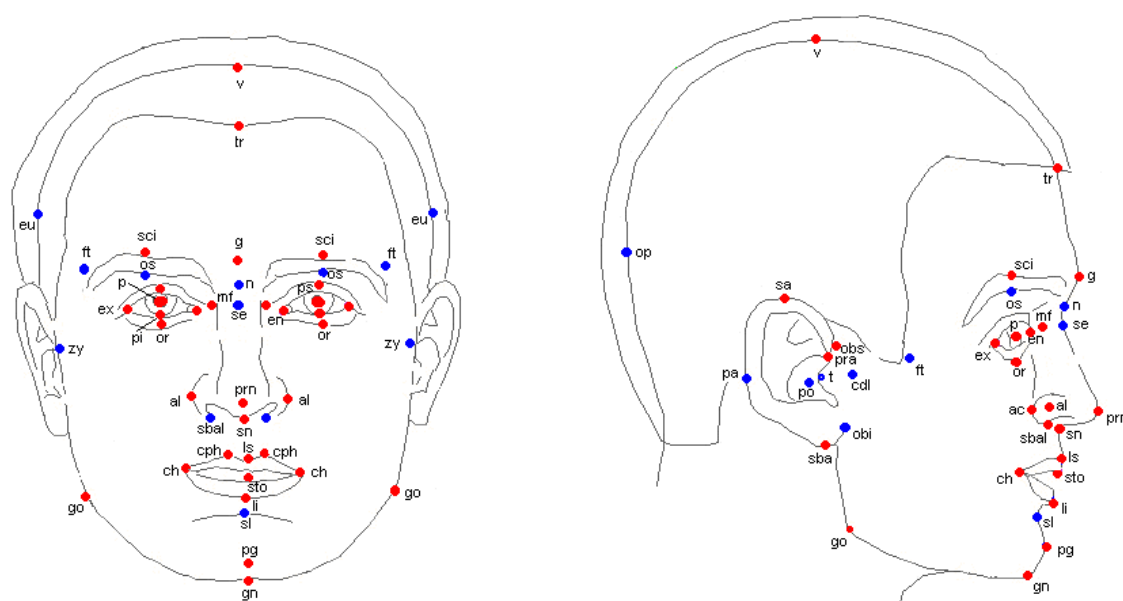


Figure 9: Anthropometric landmarks with (red) and without (blue) MPEG4 counterparts.

Definitions for these points are presented in the Table 15.

Table 15 — Definitions of the anthropometric landmarks

| Point Identifier | Point Code | MPEG4      | Anthropometric point name     | How to point                                                                                                                            |
|------------------|------------|------------|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| V                | 1.1        | 11.4       | vertex                        | The highest point of head when the head is oriented in Frankfurt Horizon. Refer to Annex E for the definition of the Frankfurt Horizon. |
| G                | 1.2        |            | glabella                      | The most prominent middle point between the eyebrows                                                                                    |
| op               | 1.3        |            | opisthocranium                | Situated in the occipital region of the head is most distant from the glabella                                                          |
| eu               | 1.5, 1.6   |            | eurion                        | The most prominent lateral point on each side of the skull in the area of the parietal and temporal bones                               |
| ft               | 1.7, 1.8   |            | frontotemporale               | The point on each side of the forehead, laterally from the elevation of the linea temporalis                                            |
| tr               | 1.9        | 11.1       | trichion                      | The point on the hairline in the midline of the forehead                                                                                |
| zy               | 2.1, 2.2   |            | zygion                        | The most lateral point of each of the zygomatic                                                                                         |
| go               | 2.3, 2.4   | 2.15, 2.16 | gonion                        | The most lateral point on the mandibular angle close to the bony gonion                                                                 |
| sl               | 2.5        |            | sublabiale                    | Determines the lower border of the lower lip or the upper border of the chin                                                            |
| pg               | 2.6        | 2.10       | pogonion                      | The most anterior midpoint of the chin, located on the skin surface in the front of the identical bony landmark of the mandible         |
| gn               | 2.7        | 2.1        | menton (or gnathion)          | The lowest median landmark on the lower border of the mandible                                                                          |
| cdl              | 2.9, 2.10  |            | condylion laterale            | The most lateral point on the surface of the condyle of the mandible                                                                    |
| En               | 3.1, 3.2   | 3.11, 3.8  | endocanthion                  | The point at the inner commissure of the eye fissure                                                                                    |
| Ex               | 3.3, 3.4   | 3.7, 3.12  | exocanthion (or ectocanthion) | The point at the outer commissure of the eye fissure                                                                                    |
| P                | 3.5, 3.6   | 3.5, 3.6   | center point of pupil         | Is determined when the head is in the rest position and the eye is looking straight forward                                             |
| Or               | 3.7, 3.8   | 3.9, 3.10  | orbitale                      | The lowest point on the lower margin of each orbit                                                                                      |
| Ps               | 3.9, 3.10  | 3.1, 3.2   | palpebrale superius           | The highest point in the midportion of the free margin of each upper eyelid                                                             |
| Pi               | 3.11, 3.12 | 3.3, 3.4   | palpebrale inferius           | The lowest point in the midportion of the free margin of each lower eyelid                                                              |

|      |            |             |                                      |                                                                                                                                                                          |
|------|------------|-------------|--------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Os   | 4.1, 4.2   |             | orbitale superius                    | The highest point on the lower border of the eyebrow                                                                                                                     |
| Sci  | 4.3, 4.4   | 4.3, 4.4    | superciliare                         | The highest point on the upper border in the midportion of each eyebrow                                                                                                  |
| N    | 5.1        |             | nasion                               | The point in the middle of both the nasal root and nasofrontal suture                                                                                                    |
| Se   | 5.2        |             | sellion (or subnasion)               | Is the deepest landmark located on the bottom of the nasofrontal angle                                                                                                   |
| Al   | 5.3, 5.4   | 9.1, 9.2    | alare                                | The most lateral point on each alar contour                                                                                                                              |
| Prn  | 5.6        | 9.3         | pronasale                            | The most protruded point of the apex nasi                                                                                                                                |
| Sn   | 5.7        | 9.15        | subnasale                            | The midpoint of the angle at the columella base where the lower border of the nasal septum and the surface of the upper lip meet                                         |
| Sbal | 5.9, 5.10  |             | subalare                             | The point at the lower limit of each alar base, where the alar base disappears into the skin of the upper lip                                                            |
| Ac   | 5.11, 5.12 | 9.1, 9.2    | alar curvature (or alar crest) point | The most lateral point in the curved base line of each ala                                                                                                               |
| mf   | 5.13, 5.14 | 9.6, 9.7    | maxillofrontale                      | The base of the nasal root medially from each endocanthion                                                                                                               |
| cph  | 6.1, 6.2   | 8.9, 8.10   | christa philtri landmark             | The point on each elevated margin of the philtrum just above the vermillion line                                                                                         |
| ls   | 6.3        | 8.1         | labiale (or labrale) superius        | The midpoint of the upper vermillion line                                                                                                                                |
| li   | 6.4        | 8.2         | labiale (or labrale) inferius        | The midpoint of the lower vermillion line                                                                                                                                |
| ch   | 6.5, 6.6   | 8.3, 8.4    | cheilion                             | The point located at each labial commissure                                                                                                                              |
| sto  | 6.7        |             | stomion                              | The imaginary point at the crossing of the vertical facial midline and the horizontal labial fissure between gently closed lips, with teeth shut in the natural position |
| sa   | 7.1, 7.2   | 10.1, 10.2  | superaurale                          | The highest point of the free margin of the auricle                                                                                                                      |
| sba  | 7.3, 7.4   | 10.5, 10.6  | subaurale                            | The lowest point of the free margin of the ear lobe                                                                                                                      |
| pra  | 7.5, 7.6   | 10.9, 10.10 | preaurale                            | The most anterior point on the ear, located just in front of the helix attachment to the head                                                                            |
| pa   | 7.7, 7.8   |             | postaurale                           | The most posterior point on the free margin of the ear                                                                                                                   |
| obs  | 7.9, 7.10  | 10.3, 10.4  | otobasion superius                   | The point of attachment of the helix in the temporal region                                                                                                              |
| obi  | 7.11, 7.12 |             | otobasion infrius                    | The point of attachment of the ear lobe to the cheek                                                                                                                     |
| po   | 7.13, 7.14 |             | porion (soft)                        | The highest point of the upper margin of the cutaneous auditory meatus                                                                                                   |
| t    | 8.1, 8.2   |             | tragion                              | The notch on the upper margin of the tragus                                                                                                                              |

The anthropometric landmark code has the format: A.B. A specifies the global landmark of the face to which this landmark belongs such as nose, mouth, etc. B specifies the particular point. In case a Landmark Point has two symmetrical entities (left and right) the right entity always has a greater and even minor code value. Hence, all Landmark Points from the left part of the face have odd minor codes, and from the right part – even minor codes. Both A and B are in the range from 1 to 15. Hence, the code  $A*16 + B$  is written to the 1 byte Landmark Point Code field.

### 5.6.7 Anthropometric 3D landmark

The error of an anthropometric 3D landmark point location should be no greater than 3 mm. The point shall withstand from the nearest point on the surface no further than 3 mm. The point on the surface is a vertex, or a point on an edge, or a point on a face of the surface.

### 5.6.8 Z coordinate

This field is not used if the Landmark Point Type is equal to MPEG4 feature or anthropometric 2D landmark. In case the Landmark Point Type equals anthropometric 3D landmark this field along with the horizontal and vertical positions denotes the coordinates of the landmark point in the 3D Cartesian coordinate system. The metric coordinates of 3D landmarks shall be obtained by multiplying the X, Y, and Z coordinates by a fixed scale of 0,02 mm. Note, that the Landmark Point Type field codes the type of the Landmark Point and determines the interpretation of the Z coordinate.

## 5.7 The Image Information Block

### 5.7.1 Data Structure

The (11 byte) Image Information block is intended to describe digital properties of the facial image, one is included for each facial image included in the record. The structure of this block is shown in Figure 2. One Representation data block shall follow this block.

### 5.7.2 Face Image Type

The Face Image Type field shall represent the type of the facial image stored in the Image Information block and, if applicable, the 3D Data block according to Table 16.

Note that all Frontal Image Types are either Full Frontal, Token Frontal, Post-processed Frontal or one of the respective 3D Full Frontal or Token Frontal Image Types. Therefore a separate Frontal Value is not required.

**Table 16 — Face Image Type codes**

| Description                      | Value                                  |
|----------------------------------|----------------------------------------|
| Basic                            | 00 <sub>HEX</sub>                      |
| Full Frontal                     | 01 <sub>HEX</sub>                      |
| Token Frontal                    | 02 <sub>HEX</sub>                      |
| Post-processed Frontal           | 03 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 04 <sub>HEX</sub> to 7F <sub>HEX</sub> |
| Basic 3D                         | 80 <sub>HEX</sub>                      |
| Full Frontal 3D                  | 81 <sub>HEX</sub>                      |
| Token Frontal 3D                 | 82 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 83 <sub>HEX</sub> to FF <sub>HEX</sub> |

The Basic Face Image Type is defined in clause 6. The Frontal, Frontal/Full, Frontal/Token, Frontal/Post-processed are defined in clauses 7, 8, 9, and 10, respectively. Face Image Types use the notion of inheritance. For example, the Frontal Face Image Type inherits all of the requirements of the Basic Face Image Type - the Frontal Face Image type obeys all normative requirements of the Basic Face Image Type. The inheritance structure of currently defined image types is shown in Figure 10.

If a 2D record that is compliant to the Basic, Full Frontal or Token Frontal requirements, respectively, contains 3D data, this is indicated by the highest bit of the Face Image Type set to one, resulting in the Face Image Type codes 80<sub>HEX</sub> to 82<sub>HEX</sub>.

Normative requirements for the Basic, Frontal, Full Frontal, Token Frontal and Post-processed Face Image Types are given in clauses 6, 7, 8, 9, and 10 respectively. The 3D Image types are defined in clauses 11, 12, and 13, respectively.

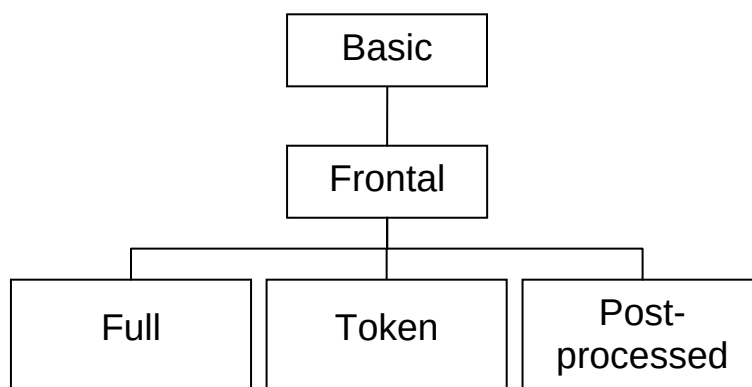


**Figure 10 — 2D Face Image Types and their inheritance map**

### 5.7.3 Image Data Type

The (1 byte) Image Data Type field denotes the encoding type of the Image Data block. Either JPEG (ISO/IEC 10918-1 and ITU-T Rec. T.81, [18]) or JPEG2000 (ISO/IEC 15444-1) or PNG (ISO/IEC 15948) shall be specified.

For lossless compression PNG or JPEG2000 lossless shall be used. For lossless representation of images using more than 8 bits per channel PNG or JPEG2000 lossless shall be used. For lossy representation of images using more than eight bit per channel JPEG2000 shall be used. Note that an “Unspecified” Value cannot be encoded.

**Table 17 — Image Data Type codes**

| Description                      | Value                                  |
|----------------------------------|----------------------------------------|
| JPEG                             | 00 <sub>HEX</sub>                      |
| JPEG2000 lossy                   | 01 <sub>HEX</sub>                      |
| JPEG 2000 lossless               | 02 <sub>HEX</sub>                      |
| PNG                              | 03 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 04 <sub>HEX</sub> to FF <sub>HEX</sub> |

### 5.7.4 Width

The (2 byte) Width field shall specify the number of pixels in the horizontal direction.

### 5.7.5 Height

The (2 byte) Height field shall specify the number of pixels in the vertical direction.

### 5.7.6 Spatial Sampling Rate Level

For specific application domains different minimal spatial sampling rates of the interchange data may be required. For example using higher spatial sampling rate images allow for specific human as well as machine inspection methods that depend on the analysis of very small details. The (1 byte) Spatial Sampling Rate Level field (see Table 18) provides information on the number of pixels in the image across the width of the head. (The Width of Head (CC) is defined in Figure 14)

**Table 18 — Spatial Sampling Rate Level codes**

| Width of Head, CC                | Spatial Sampling Rate Level            |
|----------------------------------|----------------------------------------|
| CC ≤ 180                         | 00 <sub>HEX</sub>                      |
| 180 < CC ≤ 240                   | 01 <sub>HEX</sub>                      |
| 240 < CC ≤ 300                   | 02 <sub>HEX</sub>                      |
| 300 < CC ≤ 370                   | 03 <sub>HEX</sub>                      |
| 370 < CC ≤ 480                   | 04 <sub>HEX</sub>                      |
| 480 < CC ≤ 610                   | 05 <sub>HEX</sub>                      |
| 610 < CC ≤ 750                   | 06 <sub>HEX</sub>                      |
| 750 < CC                         | 07 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 08 <sub>HEX</sub> to FF <sub>HEX</sub> |

NOTE Interocular distance in pixels will be approximately half of the head widths.

### 5.7.7 Post-acquisition Processing

While the alteration of face image data is discouraged, there are cases when no alternative may exist:

- legacy database of  $\frac{3}{4}$  frontal face images shall be rotated to full frontal prior to biometric comparison,
- from a frontal image artificial non-frontal facial images at predetermined non-frontal poses are automatically generated (multi-view images) using an implicit head model or similar. These images

can be beneficial during the comparison process or a manual review process as they show a more similar pose than the original frontal image,

- a single image is to be age progressed and used for verification of a passport holder,
- a short video stream is super-resolved to a single face image for comparison against a watchlist.

The (2 byte) Post-acquisition Processing bit field allows the specification of the kind of post processing that has been applied to the original captured image. Each bit of the mask position listed in Table 19 shall be set to 1 if the corresponding processing has been applied, and set to 0 if not. The mask position starts from 0 at the lowest bit. All bits set to zero indicates that no post-acquisition processing has been applied at all. All reserved bits shall be zero.

On the one hand a captured image typically needs some post-processing so that the resulting representation conforms to the clauses of this part of ISO/IEC 19794, especially for the frontal image type. On the other hand these processing steps should be minimal and not distort the characteristics of the original image. The right column in Table 19 clearly states, what post-acquisition processes can be applied without having to store the resulting representation in the post-processed image type as defined in clause 10.

**Table 19 — Post acquisition Processing**

| Post-acquisition Processing      | Mask Position | Use of post-processed image type mandatory |
|----------------------------------|---------------|--------------------------------------------|
| Rotated (in-plane)               | 0             | No                                         |
| Cropped                          | 1             | No                                         |
| Downsampled                      | 2             | No                                         |
| White balance adjusted           | 3             | No                                         |
| Multiply compressed              | 4             | Yes                                        |
| Interpolated (upsample)          | 5             | Yes                                        |
| Contrast stretched               | 6             | Yes                                        |
| Pose corrected                   | 7             | Yes                                        |
| Multi View Image                 | 8             | Yes                                        |
| Age progressed                   | 9             | Yes                                        |
| Super-resolution processed       | 10            | Yes                                        |
| Reserved by SC 37 for future use | 11 to 15      |                                            |

### 5.7.8 Cross Reference

The (1 byte) Cross Reference Data Type field denotes inter-dependencies when multiple representations are stored in the interchange record. This is of particular interest in the case post-processing has been used (see clause 10). Then representations that are of type Post-processed shall code the ordinal number of the representation that they have been derived from, in the Cross Reference Field. The first representation of the interchange format has the code 01<sub>HEX</sub>.

**Table 20 — Cross Reference**

| Description                                        | Value                                         |
|----------------------------------------------------|-----------------------------------------------|
| Image Type is Basic, Token Frontal or Full Frontal | 00 <sub>HEX</sub>                             |
| Image Type is Post Acquisition Processed           | Ordinal number of the original representation |

Example: There are four representations in the overall record. The second representation has been post-processed and resulted in the fourth representation. Then, the fourth representation shall have Cross Reference set to 2, all other Records shall have set Cross Reference to 0.

### 5.7.9 Image Colour Space

The (1 byte) Image Colour Space field indicates the colour space used in the encoded Image Information block according to the values in Table 21. The values of 80<sub>HEX</sub>-FF<sub>HEX</sub> are vendor specific. Application developers may obtain the values for these codes from the vendor.

**Table 21 — Colour Space codes**

| Colour Space                     | Value                                  |
|----------------------------------|----------------------------------------|
| Unspecified                      | 00 <sub>HEX</sub>                      |
| 24 bit RGB                       | 01 <sub>HEX</sub>                      |
| YUV422                           | 02 <sub>HEX</sub>                      |
| 8 bit greyscale                  | 03 <sub>HEX</sub>                      |
| 48 bit RGB                       | 04 <sub>HEX</sub>                      |
| 16 bit Greyscale                 | 05 <sub>HEX</sub>                      |
| Other                            | 06 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 07 <sub>HEX</sub> to 7F <sub>HEX</sub> |
| Vendor specific                  | 80 <sub>HEX</sub> to FF <sub>HEX</sub> |

## 5.8 The Representation Data block

The Representation Data consists of the Image Data Block, the 3D Information Block and the 3D data Block.

## 5.9 The Image Data Block

### 5.9.1 Data structure

The (variable byte) Image Data block shall consist of two fields as shown in Table 22.

### 5.9.2 Image Data Length

This four byte field shall indicate the length of the image data in bytes.

### 5.9.3 Image data

This variable length field shall contain the image data encoded by the JPEG or JPEG2000 or PNG standards

**Table 22 —image data structure**

| Field                | Size     | Value                                   | Notes                                                                        |
|----------------------|----------|-----------------------------------------|------------------------------------------------------------------------------|
| Length of Image data | 4 bytes  | $K \leq \text{Length} \leq 2^{32} - 53$ | Length of "Image data". K is minimum JPEG or JPEG 2000 or PNG header length. |
| Image data           | Variable | See Table 17                            | Either JPEG or JPEG2000 or PNG                                               |

## 5.10 The 3D Information Block

The 3D Information block consists of the following fields and sub-blocks:

The Length of 3D Data Representation, the Coordinate System Type, the Texture Projection Matrix, Scale, Offset, the 3D Representation Type, the 3D Supplemental Data, a field reserved for future use, the 3D Capture Device Technology Identifier, the 3D Capture Device Vendor Identifier, the 3D Capture Device Type Identifier, the 3D to 2D Image Temporal Synchronicity, the 3D to 2D-Texture Temporal Synchronicity, the 3D Acquisition Time, the 2D-Texture Acquisition Time, the Texture Map Type and finally the Texture Map Spectrum.

### 5.10.1 Length of 3D Data Representation

This (4 byte) field codes the length of the 3D Information and 3D Data block including the optional fields and blocks, if they are present.

### 5.10.2 Coordinate System Type

#### 5.10.2.1 General

Originally, 3D data is acquired in a device dependent coordinate system. Based on the knowledge about several device parameters the 3D data can be transformed in metric world coordinates with two disadvantages:

- the regular structure of the device dependent data gets lost (e.g. leading to varying distances between data points)
- to obtain a regular structure in world coordinates, one has to interpolate. The original data is not preserved. This may be sufficient for many applications, but this part of ISO/IEC 19794 is intended to be able to store the original data, too.

Thus the standard features several ways to store 3D data in different representations. All representations support a Cartesian coordinate system. The range data representation additionally supports a cylindrical coordinate system. Note that the coordinate system may be further restricted for different Face Image Types (see clauses 11 to 13, and Annex B.7).

The transformation to metric world coordinates is described by appropriate scaling factors and implicit rules (e.g. as used in the 3D anthropometric landmark type) defined in this part of ISO/IEC 19794 (see clause 5.10.2.2 to 5.10.2.3),

The (1 byte) Coordinate System Type field specifies the coordinate system of the 3D data by using the following values.

**Table 23 — The Coordinate System Type**

| Description                      | Value                                  |
|----------------------------------|----------------------------------------|
| Cartesian coordinate system      | 00 <sub>HEX</sub>                      |
| Cylindrical coordinate system    | 01 <sub>HEX</sub>                      |
| Reserved by SC 37 for future use | 02 <sub>HEX</sub> to FF <sub>HEX</sub> |

The different coordinate systems are defined as follows:

#### 5.10.2.2 Cartesian Coordinate System

In the Cartesian coordinate system the point of origin of the sensor data typically is used as the point of origin of the coordinate system.

The transformation from Cartesian coordinates to metric Cartesian coordinates is derived as follows:

$$X = x * ScaleX + OffsetX;$$

$$Y = y * ScaleY + OffsetY;$$

$$Z = z * ScaleZ + OffsetZ.$$

NOTE 1 for certain Image Types the origin of the Cartesian coordinate system shall be the nose, i.e. the Prn landmark as defined in Table 15.

NOTE 2 for certain Image Types the pose of the head is restricted. The frontal pose is defined by the Frankfurt Horizon FH (see Annex E) as the XZ plane and the vertical symmetry plane as the YZ plane with the Z axis oriented in the direction of the face sight.

A strong relation between anthropometric landmarks and the coordinate system is still established by

- the anatomical alignment requirements of the corresponding 2D image and
- the alignment between the 3D range data and the corresponding 2D image after applying the Texture Projection Matrix.

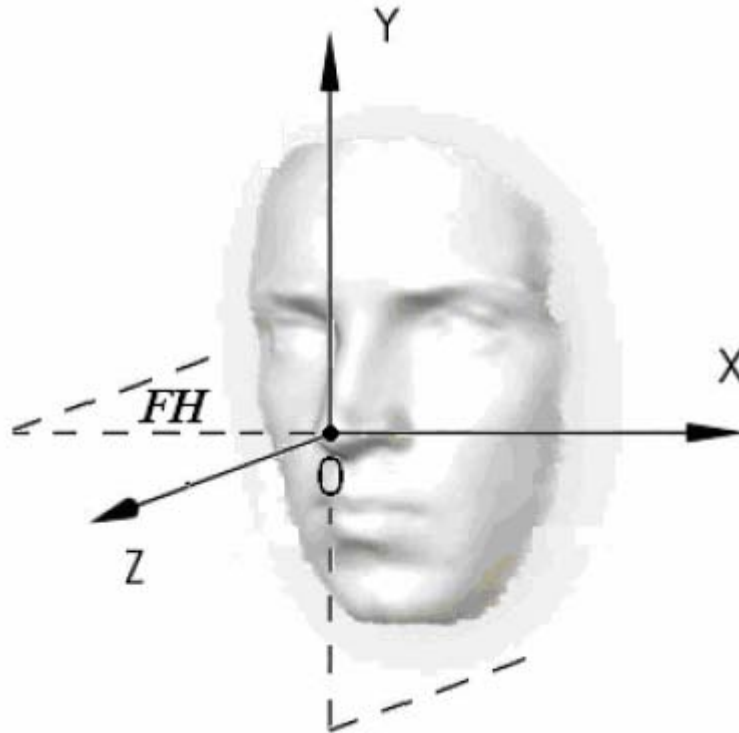


Figure 11 — A sample of a Cartesian coordinate system

#### 5.10.2.3 The Cylindrical Coordinate System

A point in the Cylindrical Coordinate System is given by  $(\alpha, h, r)$ . The angle  $\alpha$  and the  $h$ -axis are defined in a way that they form a clockwise coordinate system.

The transformation from cylindrical coordinates to metric Cartesian coordinates is derived as follows:

$$X = r * ScaleZ * \sin(\alpha * ScaleX) + OffsetX;$$

$$Y = h * ScaleY + OffsetY;$$

$$Z = r * ScaleZ * \cos(\alpha * ScaleX) + OffsetZ;$$

ScaleX, ScaleY, ScaleZ, OffsetX, OffsetY and OffsetZ are the necessary constants for the transformation. ScaleX has the physical unit of rad (degree radian). ScaleY, ScaleZ, OffsetX, OffsetY and OffsetZ are given in the physical unit mm (millimetre). Note, that large values of ScaleX, ScaleY or ScaleZ indicate a low spatial sampling rate in the respective dimension.

Typically, the point of origin of the sensor data is used as the point of origin of the cylindrical coordinate system.

NOTE for certain Image Types the origin of the cylindrical coordinate system shall be the nose, i.e. the Prn landmark as defined in Table 15.

A strong relation between anthropometric landmarks and the coordinate system is still established by

- the anatomical alignment requirements of the corresponding 2D image and
- the alignment between the 3D data and the corresponding 2D image after applying the Texture Projection Matrix.

The transformation from cylindrical coordinates to Cartesian coordinates is done by applying the transformation denoted in clause 5.10.2.3 and then inverting the transformation given in 5.10.2.2.

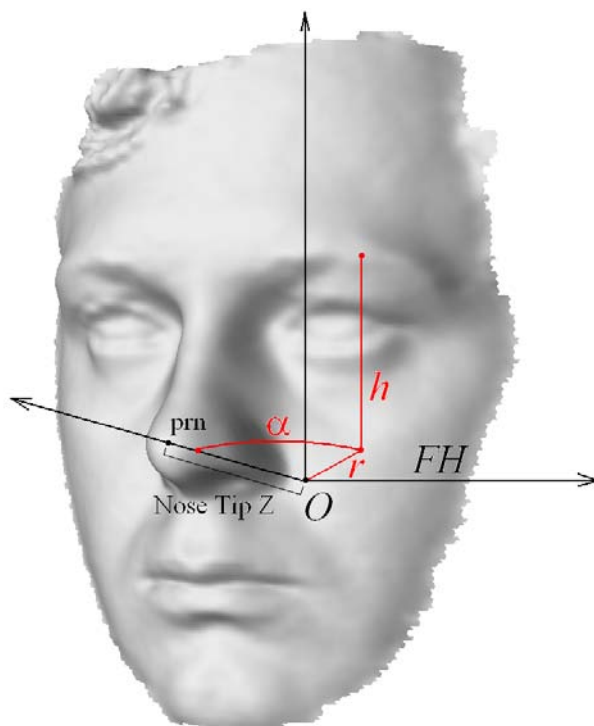


Figure 12 — A sample of a cylindrical coordinate system. FH is the Frankfurt Horizon as defined in Annex E.

### 5.10.3 Texture Projection Matrix

The Texture Projection Matrix  $P$  (3x4 float, 48 bytes) is required to map the 3D data onto the 2D texture image of the Image Data block. The matrix shall be stored row by row starting from the left top.

One can project a point in 3D space  $[X, Y, Z]^T$  on the texture image of the Image Data block by multiplying the Texture Projection Matrix  $P$  with the so called homogeneous 3D coordinates of the 3D point [11].

$$[x, y, w]^T = P * [X, Y, Z, 1]^T$$

Homogeneous 3D coordinates are a vector of four values  $[X, Y, Z, 1]^T$ . Here  $X, Y, Z$  are the coordinates of a point in the metric cartesian coordinate system. The multiplication results in  $[x, y, w]^T$ , the so called homogeneous 2D coordinates with the auxiliary coordinate  $w$ . One obtains the resulting 2D image pixel coordinates of the texture image in the Image Data block by dividing the first two coordinates of the 2D