For Publication on the ICAO Website

TECHNICAL REPORT

Portrait Quality

(Reference Facial Images for MRTD)

Version 1.0

April 2018

File: 2018_04_08 Portrait Quality
Author: ISO/IEC JTC1 SC17 WG3

Release Control

Release	Date	Description
0.1	2014-09-05	First draft based on section 3.9_doc9303.part3.7th edition. Draft 11 from 2014.04.01 and the comments from CAN, NZ, UK, D (BKA, BDr)
0.2	2014-09-30	Second draft based on resolution of comments in TF4D meeting in Salamanca, Spain and CAN/AUS comments
0.3	2015-02-16	Content taken from ISO/IEC 19794-5 and sorted into a readable structure, enriched with the content from version 0.02
0.4	2015-05-29	Implementation of the DoC to 0.02 and 0.03 approved in Leiden, additional restructuring and clarifications provided by the editors based on a FR proposal.
0.4+	2015-07-16	Joined meeting AFNOR/DIN on common comments on v0.04
0.5	2015-10-29	Implementation of the Vienna comments and of the sample portraits from Australia.
0.6	2016-05-23	Implementation of the Wellington comments.
0.7	2016-09-01	Implementation of the Paris comments.
0.8	2017-04-01	Implementation of the Sydney comments.
0.9	2017-11-24	Implementation of the Paris comments. Submission to NTWG for approval.
1.0	2018-04-08	Implementation of NTWG comments and Malta CRM results. Final version submitted to TAG for approval.

TABLE OF CONTENTS

1 INTRODUCTION

ICAO Doc 9303 provides the basic functional specification for MRTD's and describes all relevant properties of MRTD's.

The portrait printed on the ICAO compliant MRTD is an essential element of that document and one of the most important information carriers binding the document to the holder. A standardized portrait produced at a high quality helps issuing agencies to screen identity and border agencies to inspect the travel document manually or via automated processing.

After the introduction of the digitally stored image in 2005, ABC (Automated Border Control) systems have been introduced to perform automated comparison of the person and the electronically stored image. Those ABC systems compare, whether it is manually or automated the printed image and/or the electronically stored image and the image taken live while crossing a border.

This document is based on ISO/IEC 19794-5:2005 and ISO/IEC 19794-5:2011 as well as on Doc 9303 statements on portraits. The content of these documents has been rearranged, consolidated, enriched, and improved within ISO/IEC JTC1 SC17 WG3.

As soon as this document is confirmed by ICAO and published as ICAO TR, it is planned to be submitted by SC17 to SC37, and to be integrated into ISO/IEC 39794-5 describing the requirements on MRTD portraits.

This document comes into effect once adopted by the ICAO TAG/TRIP. The implementation strategy will be published in the ICAO Guidance Roadmap for Implementation of New Specifications roadmap.

2 SCOPE

This document describes the requirements and best practice recommendations to be applied for portrait capturing in the application case of enrolment of biometric reference data for electronic MRTD. In this sense, this document is an application profile.

This document:

- shares the lessons learned using the stored and displayed portrait in an MRTD,
- describes how the portraits should be captured that serve as the content of ISO/IEC 19794-5 and ISO/IEC 39794-5 data structures,
- provides the experiences made applying facial recognition technology in ABC gates, manual border control, identity screening, and other applications based on the portraits provided by electronic MRTD's. It also gives guidance on the requirements for capturing and processing portraits contained in MRTD's to support the inspection process,
- provides comprehensive recommendations for portrait capturing including scene, photographic and digital requirements,
- provides requirements for image printing and scanning as well as on digital image processing,
- provides requirements for portraits printed on MRTD's to ensure good visibility for inspection, and
- gives guidance for reader system manufacturers on the use of unified reflection free illumination and view angles.

The following topics are not in scope of this document:

- Image capturing for verification and/or identification applications like ABC, even if many of the requirements listed in this document apply to such images, too,

- Definition of image data formats like JPEG, JPEG2000, PNG,
- Security aspects like digital image electronic signature, Presentation Attack Detection (PAD), and morphing prevention.
- Description of biometric data formats for inclusion into an electronic MRTD which is done in ISO/IEC 19794-5.

For certain criteria, there may be two different levels given in a table form: A minimum requirement and a best practice recommendation. The requirement gives the minimum acceptable values or value ranges in order to reach compliance. The best practice recommendation gives values that will result in better overall performance or quality, and users are encouraged to adopt best practice values whenever possible.

Table 1 – Sample Table summarizing minimum requirements and best practice recommendations.

Criterion	Requirement	The criterion shall be …
	Best Practice	The criterion should be …

3 GENERAL

3.1 Terminology

The key words “shall”, “shall not”, “should”, “should not”, “may”, and “may not” are used in this document as defined in the ISO/IEC Directives Part 2, Annex H, 7th Edition, 2016.

3.2 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.

[R1] ICAO Doc 9303 and ISO/IEC 7501 multipart standard: Machine Readable Travel Documents

[R2] ISO/IEC 10918-1:1994: Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines

[R3] ISO/IEC 19794-5:2005: Information technology -- Biometric data interchange formats -- Part 5: Face image data.

[R4] ISO/IEC 15444-1:2004: Information technology - JPEG 2000 image coding system: Core coding system

[R5] Technical Guideline TR-03121-3: Biometrics for public sector applications, Part 3: Application Profiles and Function Modules, Volume 1: Verification scenarios for ePassport and Identity Card, Version 3.0.1. 2013

[R6] ISO 11664-2:2007(E)/CIE S014-2/E: 2006, CIE Standard Illuminants for Colorimetry

[R7] ISO/IEC 15948:2004: Information technology – Computer graphics and image processing – Portable Network Graphics (PNG): Functional specification

[R8] ISO 12233: Photography – Electronic still picture imaging – Resolution and spatial frequency responses.

[R9] ISO/IEC 14496-2 Information technology – Coding of audio-visual objects Part 2: Visual [MPEG4]

[R10] IEC 61966-2-1: Multimedia systems and equipment – Colour measurement and management – Part 2-1: Colour management – Default RGB colour space – sRGB

[R11] IEC 61966-8:2001. Multimedia systems and equipment – Colour measurement and management – Part 8: Multimedia colour scanners

3.3 Terms and definitions

3.3.1 1:1 application case

biometric process (algorithm) comparing a sample photo with a registered sample of the claimed identity, also known as verification

3.3.2 1:N application case:

biometric process (algorithm) searching an a priori unknown sample photo among N registered samples in a database, also known as identification

3.3.3 ABC

Automated Border Control

3.3.4 Adobe RGB

RGB colour space designed to encompass most of the colours achievable on CMYK colour printers, but by using RGB primary colours on a device such as a computer display

3.3.5 CCD

Charge-coupled device

3.3.6 Chin

central forward portion of the lower jaw

3.3.7 CIC

Contactless Integrated Circuit

3.3.8 CIE Standard Illuminant D65

commonly used standard illuminant defined by the International Commission on Illumination (CIE) that is part of the D series of illuminants trying to portray standard illumination conditions at open-air in different parts of the world

3.3.9 CMOS

complementary metal–oxide–semiconductor

3.3.10 crop factor

ratio of the diagonal of the full frame camera (43.3 mm) to that of a selected camera’s image sensor

Note 1 to Entry: The determination of an appropriate focal length lens for a field of view equivalent to a full frame camera can be made by consideruing the crop factor.

3.3.11 Crown

top of the head ignoring any hair

3.3.12 CSD

camera to subject distance

distance between the eyes plane of a person and the optical center of the camera lens

3.3.13 DG

Data Group as defined in ICAO Doc 9303

3.3.14 DOVID

Diffractive Optically Variable Image Device

3.3.15 Eye to mouth distance

EM

distance between the face centre M and the mouth midpoint (feature point 2.3 from ISO/IEC 14496-2)

3.3.16 Exposure value

EV

number that represents a combination of a camera's shutter speed and f-number, such that all combinations that yield the same exposure have the same EV

3.3.17 EVZ

eye visibility zone

zone covering a rectangle having a distance *V* of at least 5% of the IED to any part of the visible eye ball

3.3.18 Eye centre

centre of the line connecting the inner and the outer corner of the eye

Note 1 to entry: The eye centres are the feature points 12.1 and 12.2 as defined in ISO/IEC 14496-2.

Note 2 to entry: The inner and the outer corner of the eye are defined by ISO/IEC 14496-2. They are the feature points 3.12 and 3.8 for the right eye, and 3.11 and 3.7 for the left eye.

3.3.19 Face centre

M

midpoint of the line connecting the two eye centres

3.3.20 HD

horizontal deviation angle

maximal allowed deviation from the horizontal of the imaginary line between the nose of a person and the lens of the camera

3.3.21 Inter Eye Distance

IED

length of the line connecting the eye centres of the left and right eye

3.3.22 Issuer

organisation that issues MRTDs

3.3.23 LDS

Logical Data Structure (as defined in ICAO Doc 9303)

3.3.24 magnification distortion

image imperfection where the degree of magnification varies with the distance from the camera and the depth of the face

3.3.25 Moiré pattern

artefact resembling a wavy pattern caused by photographing a scene or object containing repetitive details (e.g., lines, dots, etc.) that exceed the sensor resolution of the camera

3.3.26 MP

intensity measurement zone side length, the intensity measurement zones have a square shape and a size of 30% inter eye distance; they are used for measuring the lighting intensity on cheeks, forehead, and chin

3.3.27 MRTD

Machine Readable Travel Document, the term also includes electronic MRTD's, electronic Machine Readable Travel Document using a Contactless Integrated Circuit (CIC)

3.3.28 MTF

modulation transfer function

3.3.29 MTF20

highest spatial frequency where the MTF is 20% or above

3.3.30 parallax

displacement or difference in the apparent position of an object viewed along two different lines of sight, measured by the angle or semi-angle of inclination between those two lines

3.3.31 Photo kiosk

semi-automated system for digitally capturing identity photographs in a bureau-environment; it consists of camera and lighting and usually has a separate panel placed behind the subject to provide the required background but is otherwise open

3.3.32 Photo booth

automated system for digitally capturing identity photographs in either public or office environments; it encloses the subject in a highly-controlled lighting environment and consists of a camera, lighting and peripheral devices such as printers; it has entrances on one or both sides with reflective curtains protecting against ambient light

3.3.33 Portrait

visual representation of the facial image of the holder of the MRTD in printed and electronically stored manner

3.3.34 Presentation Attack

Presentation of an artefact or human characteristic to the biometric capture subsystem in a fashion that could interfere with the intended policy of the biometric system

3.3.35 Presentation Attack Detection

PAD

automated determination of a presentation attack

3.3.36 radial distortion

image imperfection where the degree of magnification varies with the distance from the optical axis

3.3.37 RFID

radio-frequency identification

3.3.38 SFR

spatial frequency response

3.3.39 SNR

signal to noise ratio

3.3.40 sRGB

standard RGB colour space created for use on monitors, printers and the Internet using the ITU-R BT.709 primaries

3.3.41 Subject

person which is to be displayed on the portrait, this person is intended to be the holder of the MRTD

3.3.42 TR

Technical Report

4 OVERVIEW

4.1 Introduction

Portraits appear in several places on and in an MRTD:

- As a printed image on the data page (Zone V as defined in ICAO Doc 9303),
- As a digital image stored in the RFID chip,
- Optionally, as a secondary image on the data and/or observation page, e.g., as a changeable laser image, as a micro-perforation, or as a background print.

All the images used shall be derived from the same captured portrait. However, the technical requirements of each of the images may differ depending on the applied technology.

The intended use for printed portrait is to give a good physical representation of the document holder and to allow for a human comparison of the portrait and the holder of the MRTD. Physical security features and the printing technology may interact or influence the portrait which needs to be considered as part of the comparison process.

The intended use of the portrait image digitally stored in the chip is such that the image can be compared to the printed portrait and the human via manual processes or compared to a live image via automated processes in a 1:1 od 1:N application case. Because of the way the image is stored on the MRTD (see Doc 9303 Part 11), border agencies can confirm that the image has been stored on the MRTD by the issuing authority and remains unaltered or unsubstituted. The digital image is the primary image used for biometric comparison.

The secondary images serve as physical security features protecting the printed portrait. Therefore their appearance shall be the same as the printed portrait. However, size and production technology determine the technical requirements of the portrait derivative used here.

Figure 1 displays the appearance of the portrait as a printed image and as a digitally stored image.

Figure 1 — Portraits to be used on/in an MRTD

Two main types of application processes are considered, those based on:

- submission of printed photographs provided by the citizen to the passport authority, and on
- electronic portrait submission.

There are two sub-types of the second type:

- live capture where the applicant has the photo taken during an interview or application submission, and
- upload, where the image is provided electronically by the applicant, by an enrolment bureau or by an accredited ID photo service.

These two sub-types are subject to the same requirements. Depending on the process type, different Clauses of this document apply, as defined below. Both main types are shown in Figure 2.

The production of the printed as well as of the electronic portraits may be done by automated kiosks, officers of passport authorities, photo booths, or photographers. It is essential that the quality requirements are met. Photographic experts should be consulted before introducing a new enrolment solution.

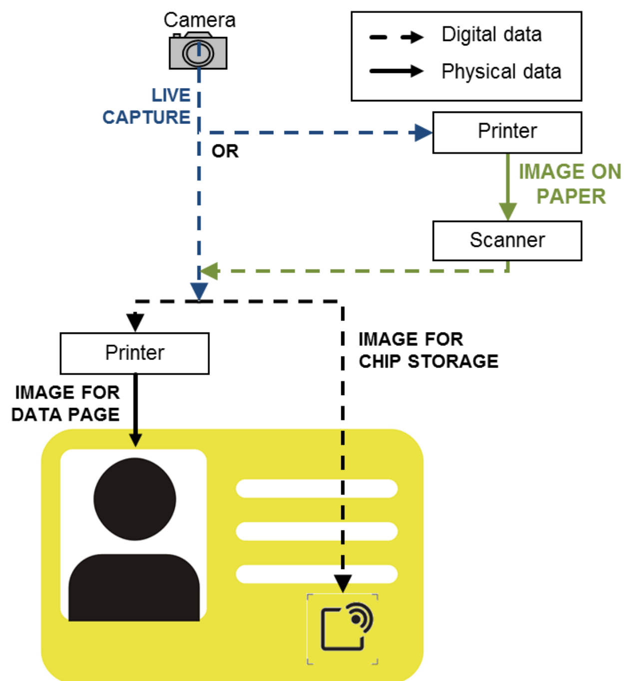


Figure 2 — Portrait enrolment process variations

4.2 Passport application using printed portraits

For a passport application process that uses printed portraits the citizen typically visits a photographer or photo booth to obtain such a portrait. In all cases the citizen receives printed photos, and there is no electronic submission or linkage to an electronically stored image available. Then the citizen submits such a photo to the passport issuing authority as part of their application. To establish a passport application process using printed portraits, Clauses 5, 6, 7, 8, and 9 apply.

4.3 Passport application using electronic portrait submission

Enrolment data providers like photographers, photo booths, or kiosks can be linked electronically to the issuing authorities. For a passport application that uses electronic submission, the intermediate steps dealing with a printed photo are skipped. In most cases, the photo is digitally captured and electronically stored or directly transmitted to the passport issuing authority. There are many ways the portrait may be transferred to the passport issuing authority. Such schemas include direct transmission to the authority, a data carrier submitted by the citizen, and temporary storage on a server and submission of a reference to the uploaded/stored photo provided by the citizen. Live capture where the applicant has the photo taken during an interview or application submission is covered here, too. To establish a passport application process using electronic portraits, Clauses 5, 8, and 9 apply. The IED requirement for electronically submitted images follows the same requirement for chip images as written in Clause 9.2 and Table 11.

4.4 Non-professional photographs

Passport applicants should not be encouraged to submit portraits captured by amateur photographers, or captured on amateur equipment such as mobile phones or tablets, or printed on consumer printers (home-made portraits) as they typically do not achieve the required quality level as specified in Clause 5 (and 6 for printed portraits). If an issuer decides to accept homemade portraits, the issuer shall ensure, based on an appropriate level of expertise, that the printing quality and all of the requirements specified in Clause 5 (and 6 for printed portraits) are maintained, and that the risks of photo manipulation and morphing inherent with such an uncontrolled process are suitably mitigated.

5 ENROLMENT LIVE PORTRAIT CAPTURING

5.1 Introduction

This Clause describes the requirements for the environment that is used for portrait capturing. Additionally, it gives recommendations on best practice. The environment requirements are derived from experiences made in facial recognition applications including ABC gates, and they consider the methods used by professional photographers.

Before introducing new equipment and defining processes for enrolment data capturing, an experienced portrait photographer and/or an optics expert should be asked for advice. The requirements apply to all installations including photo booths and kiosks.

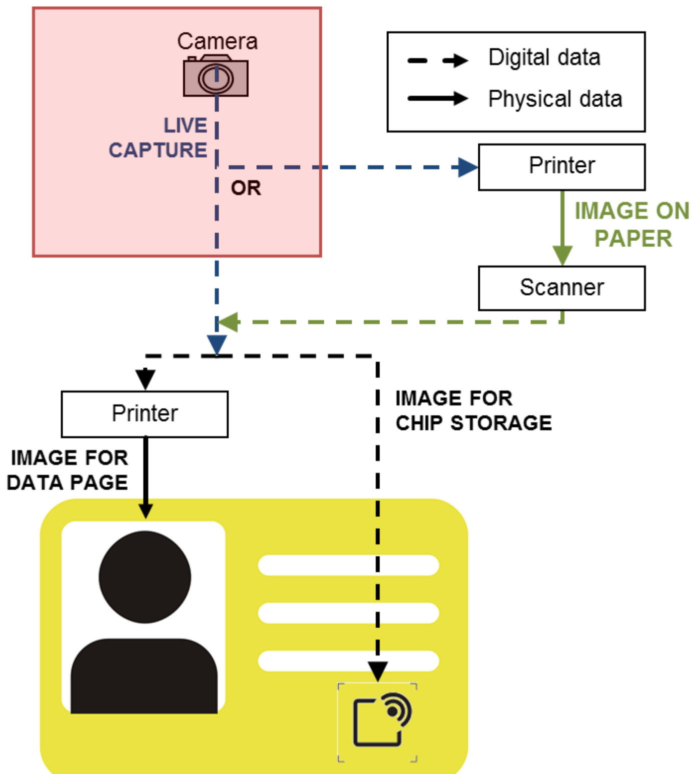


Figure 3 — Content of Clause 5 in the MRTD production process chain (boxed in red).

This Clause specifies requirements for the photograph being captured as well as for the photographic equipment being used. Figure 3 shows the content of Clause 5 in the MRTD production process chain.

NOTE: Pose constraints are very difficult to evaluate on the acquired 2D image, even for experts in this field. Numeric values have been provided in this document to provide for a consistent subject positioning in the full frontal pose.

5.2 Camera and scene

5.2.1 Selection of camera and lens focal length

In addition to choosing an appropriate camera-to-subject distance (CSD), as described in Clause 5.2.2, the selection of a camera and its lens is a major factor affecting the quality of portrait images. To ensure high image quality and a standards-compliant inter eye distance (IED), the camera's sensor must have sufficient pixel dimensions and its lens must be chosen to match its image sensor's physical dimensions.

For example, for a camera using an APS-C sensor (having a crop factor of 1.44), one should consider using a lens of focal length between 50/1.44 and 130/1.44, or roughly 35 to 90 mm.

For portrait photos, photographers using a conventional 35 mm film camera (having a 36 mm × 24 mm frame, with a 43.3 mm diagonal) often selected a normal to moderate telephoto lens, with a focal length between 50 and 130 mm (or an equivalent zoom lens.) For digital cameras employing typically smaller size CMOS or CCD image sensors, the lens selected for portrait photography should have a proportionally decreased focal length.

Figure 4 illustrates the relative sizes of some commonly available image sensors. Table 3 provides the approximate widths, heights, areas, diagonals, and crop factors for these sensors.

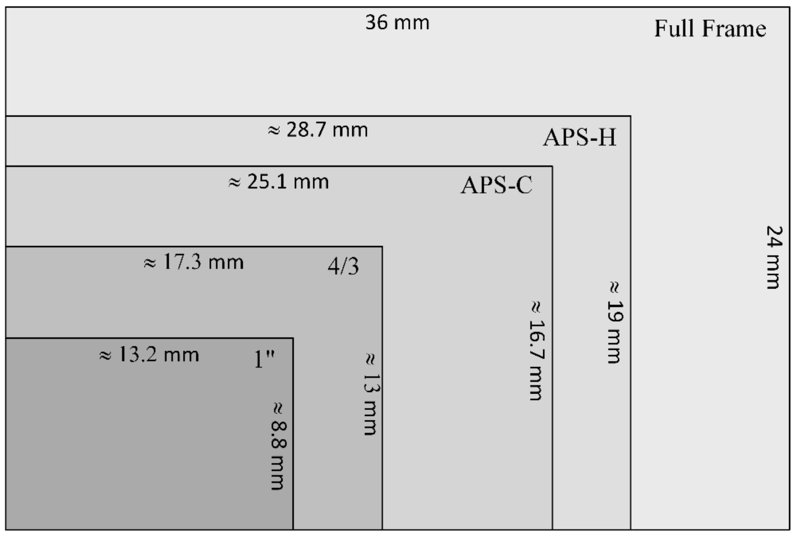


Figure 4 — Typical sensors and their relation in size to the traditional 35 mm full frame

Table 2 —Some image sensor sizes and corresponding crop factors. The dimensions are approximates and serve as examples.

Sensor Type	Width (mm)	Height (mm)	Area (mm^2)	Diagonal (mm)	Crop Factor
35 mm Full Frame	36	24	864	43.3	1
APS-H	28.7	19	545	34.4	1.26
APS-C	25.1	16.7	419	30.1	1.44
Four Thirds System (4/3)	17.3	13	225	21.6	2.00
1 Inch (1")	13.2	8.8	116	15.9	2.73

It might be noted that by gathering more light, a larger image sensor will provide typically lower image noise, and a fixed focal length lens will generally provide a higher image quality than a zoom lens of the same focal length. Moreover, by using a fixed focal length lens, the problem of inadvertent change to the zoom ratio (i.e., the field of view) can be avoided.

Figure 5 illustrates the typical optical arrangement and terminology for portrait image acquisition, as well as some of the variables in the arrangement.

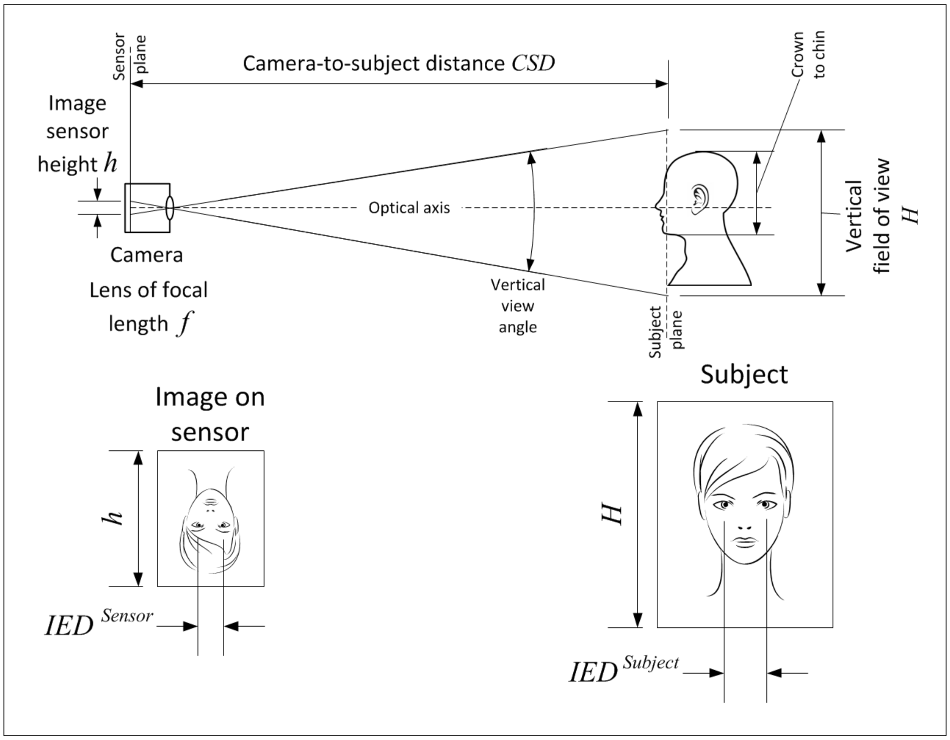


Figure 5 — Illustration of optical arrangement and terminology

For a selected CSD (in millimeters), and camera image sensor with a vertical dimension in millimeters of hmm , a requested field of view of $H_{FieldOfViewmm}$, the focal length f (in millimeters) can be computed using the following relation-ships in order to optimise the requested field of view of the subject into the sensor dimensions :

$$f \cong h_{mm} \frac{CSD}{H_{FieldOfViewmm}}$$

In case of homemade portrait, the lens optimisation is not done due to large camera angle.

For the same camera image sensor with vertical pixel count of hpx, the inter eye distance on the sensor in pixels IED_{px}^{Sensor} may be computed using the following relation-ships, where $IED_{mm}^{Subject}$ is inter eye distance in millimeters on the subject.

$$IED_{mm}^{Sensor} \cong IED_{mm}^{Subject} \times \frac{f}{CSD} \quad \text{and} \quad IED_{px}^{Sensor} = IED_{mm}^{Sensor} \times \frac{h_{px}}{h_{mm}}$$

As an example, one commercially available, digital single lens reflex (DSLR) camera has the following specifications:

Sensor: APS-C, 22.3 mm x 14.9 mm, 5184 × 3456 px, 18 Megapixels

For a CSD of 1200 mm, a typical $H_{FieldOfViewmm}$ of 500mm, a typical IED Subject of about 62 mm:

the calculations below show that the focal length f will be about 50 mm (equivalent to 80 mm full frame).

$$f \cong 22.3mm\,\frac{1200mm}{500mm} \cong 53.5mm \cong 50mm$$

the calculations below show that IED_{px}^{Sensor} will be about 598 pixels, or well above the best practice value suggested in Table 5 in Clause 5.2.4.

$IED_{mm}^{Sensor} \cong\ 62mm \times \frac{50mm}{1200mm} \cong\ 2.58mm$ and $IED_{px}^{Sensor} \cong\ 2.58mm \times \frac{3456\ px}{14.9mm} =\ 598$ pixels

For a sensor of 5 Megapixels (2592 x 1944) with an optimized focal, the IED_{px}^{Sensor} will be about 336 pixels, still well above the best practice value suggested in Table 5 in Clause 5.2.4.

$IED_{px}^{Sensor} \cong\ 2.58mm \times \frac{1944\ px}{14.9mm} =\ 336$ pixels

5.2.2 Magnification distortion and camera subject distance

All images captured by a photographic system will contain image distortion. Every portrait is a compromise between different requirements like camera and lens costs or available space and illumination. In this document requirements and recommendations are given to ensure global interoperability in the sense that the most important properties of every portrait used for MRTD purposes reach the correct quality requirements and therefore ensure similar performance in facial image based authentication applications like border control systems.

The *CSD* requirements are listed in Table 3 and considered in more detail in Clause 5.2.2. For sample portraits illustrating possible effects of the optical system see Figure 7. Table 4 lists different camera subject disctances and their corresponding magnification distortions.

Magnification distortion can only be evaluated by measuring tools. See Annex B. It is not possible to evaluate magnification distortion value from human vision. For information, ears start to be masked around a magnification distortion of 14% or higher.

NOTE: Selfie style portraits are likely not to maintain the minimal distance requirement.

Table 3 – *CSD* requirements and recommendations.

Criterion: *CSD* for 1:1	Requirement	0.7 m ≤ CSD ≤ 4 m
	Best Practice	1,0 m ≤ CSD ≤ 2,5 m
Criterion: *CSD* for 1:N	Requirement	1 m ≤ CSD ≤ 4 m
	Best Practice	1,2 m ≤ CSD ≤ 2,5 m

The camera shall be at the subject's eye-level. The line between camera and center of subject's face shall be horizontal within a maximum *HD* of ± 5°. Height alignment should be done by vertical adjustment of either subject or camera. See Figure 6.

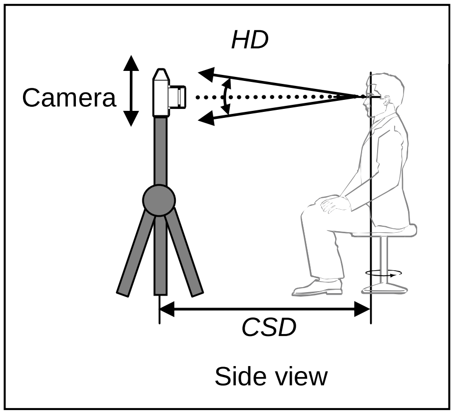


Figure 6 — Alignment of camera and subject

These recommendations and requirements apply for all capturing setups including photo booths and kiosks.

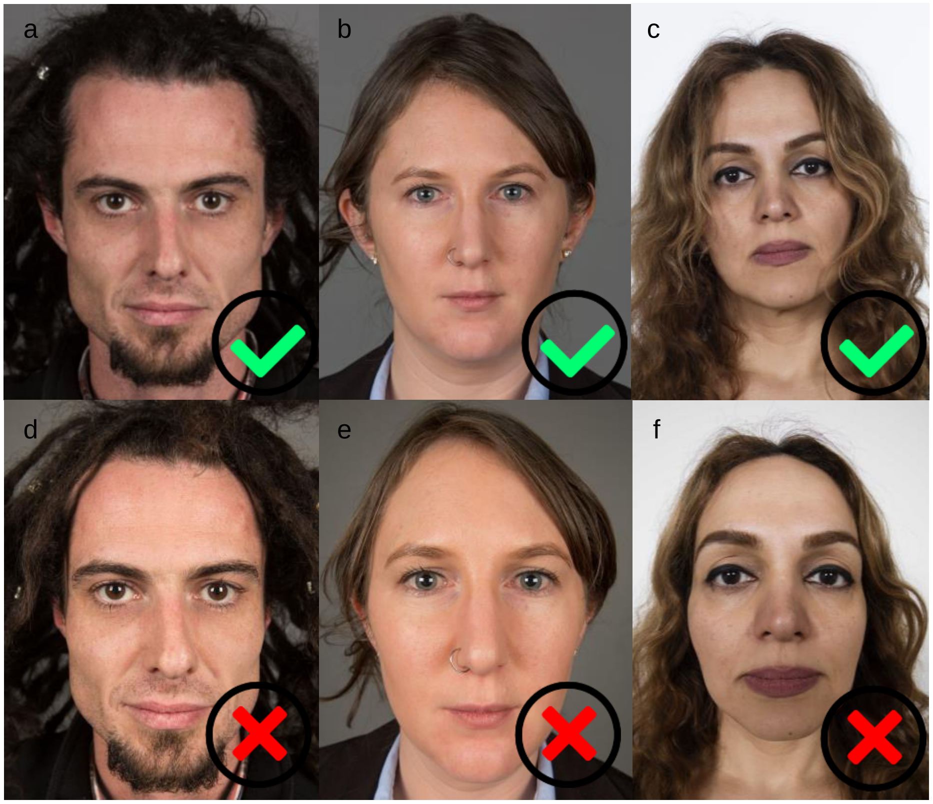


Figure 7 — a), b), and c) Good appearance. d), e), and f) Too strong magnification distortion.

One of the important factors that influence the appearance of the facial features is the distance between subject and camera lens.

The magnification distortion due to camera-subject distance can be noticeable to human examiners, but shall be within defined limits that allow effective face recognition.

Acceptable distortion rate tolerances depend on the performance capacity of state of the art facial recognition technology, and on the capability of typical human inspection staff to recognize people, even those coming from varying ethnic origin.

Figure 8 — Sample portraits taken with a full size sensor camera at focal length 50 mm from distances of a) 30 cm, b) 40 cm, c) 60 cm, d) 70 cm, e) 80 cm, f) 100 cm, g) 150 cm, h) 200 cm, and i) 250 cm. These images have been captured using the enrolment bench described in Annex F. All images have been normalized to a constant IED. The red bars mark the distance between the feature points 10.7 and 10.8 according to ISO/IEC 14496-2:2004 measured in Sub-Figure i).

Rulers at nose and ear may be used to measure the geometric effect to the face, i.e., a millimetre at nose level is larger than a millimetre at ear level on the image of the rulers.

The maximum level of magnification distortion of the capturing process shall be set depending on the appropriate use case (see below).

- **1:1 use case:** At the border, an automatic and/or human facial verification/comparison is progressed. This is the case in most automated border control applications. The maximum magnification distortion rate of the picture in the passport **shall not be greater than 7%** and ideally **should not be greater than 5%**.
- **1:N use case:** At the enrolment or issuance time of the document, a 1:N facial identification is done on a database to help verifying the uniqueness of the identity associated to the new image provided. N is as large as the number of images searched. This use case requires higher quality enrolment. The maximum magnification distortion rate **shall not be greater than 5%** and ideally **should not be greater than 4%**.

The study presented in Annex F has shown that, for a large range of enrolment and verification distances the influence of magnification distortion on automatic face recognition system performance is low.

The magnification distortion is considered to be noticeable if the distance between units on a ruler at the nose tip level measured in pixels is more than 5% larger than the distance between units on a ruler at the outer canthus level measured in pixels. The elevation of the nose compared to the outer canthus of the test subject is assumed to be 50 mm. It is sufficient to measure this properly once whenever a photographic setup is introduced or modified. An example photo is given in Figure 10. Examples of portraits with good appearance and too strong magnification distortion are given in Figure 6. The general case of the optical system is discussed in Annex B.

Figure 9 — a) photo captured with rulers at nose and eye level, b) detail showing the millimeter scales at nose level, c) detail showing the millimeter scales at eye level.

There are several possible strategies for decreasing the magnification distortion. The general assessment of an optical system is discussed in Annex B. Assuming a telecentric lens, the distance between sensor and subject does not introduce any magnification distortion. Real systems need specific considerations and measurements like those described in Annex B. Another strategy to decrease the magnification distortion is to increase the distance between subject and camera, or to fold the optical path. The principle of a folded optical path is illustrated in Figure 10. These strategies are not limitative. Sample images taken with a high quality camera with several magnification distortion rates are given in Figure 8.

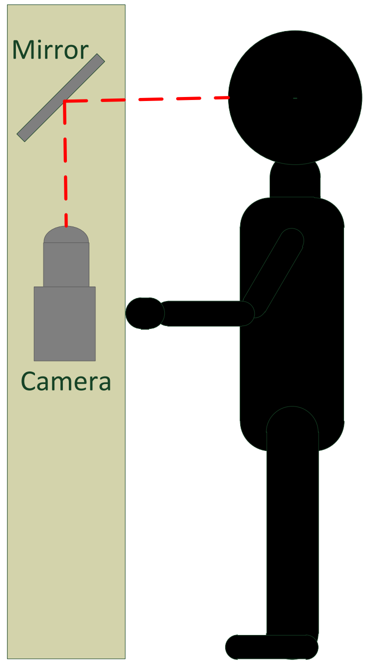


Figure 10 — Principle sketch of a folded optical path.

Table 4 – Camera subject disctance and corresponding magnification distortion.

camera subject distance in m	**magnification distortion Δd/d** **for a standard (i.e. not telecentric) lens**
0,3	16,7%
0,4	12,5%
0,5	10,0%
0,6	8,3%
0,7	7,1%
1	5,0%
1,2	4,2%
1,5	3,3%
2	2,5%
2,5	2,0%
3	1,7%

NOTE: This magnification distortion only applies for standard (i.e. non-telecentric lens) lenses.

Homemade portraits are also affected by magnification distortion tolerance. Issuers who accept homemade portraits should be aware that there is no scientific solution that allows checking the tolerance compliancy. Therefore the acceptance of homemade portrait is not recommended.

The issuer should allow for a transition period in which enrolment systems like, e.g., photo booths and kiosks may be updated to fulfil the magnification distortion requirements considering economic and feasibility reasons. The duration of such transition period is at the discretion of the issuer.

5.2.3 Radial distortion

The radial distortion due to lens properties can be noticeable to human examiners, but shall be within defined limits that allow effective face recognition. In particular, fish eye effects caused by wide angle lenses combined with camera placement too close to the face shall not be present.

Acceptable distortion rate tolerances depend on the performance capacity of state of the art facial recognition technology, and on the capability of typical human inspection staff to recognize people, even those coming from varying ethnic origin.

If the radial distortion is less than 2%, the human eye will not easily perceive it. It is recommended that radial distortion is less than 2.5%.

The general assessment of an optical system is discussed in Annex B.

5.2.4 Pixel count, focus and depth of field

Digital cameras used to capture portraits shall produce images where the vertical and horizontal pixel density is the same.

Live captured portraits of a subject

- Shall be captured in one of the following formats: PNG, JPEG, JPEG2000, RAW formats supported by the camera, lossless formats should be preferred,
- Should be captured at a minimum dimension of 1200 pixels width x 1600 pixels height (cropped image),
- Shall be captured in colour.

One of the four possible encodings shall be used:

- The JPEG Sequential Baseline (ISO/IEC 10918-1) mode of operation and encoded in the JFIF file format (the JPEG file format).
- The JPEG-2000 Part-1 Code Stream Format (ISO/IEC 15444-1), lossy, and encoded in the JP2 file format (the JPEG2000 file format).
- The JPEG-2000 Part-1 Code Stream Format (ISO/IEC 15444-1), lossless, and encoded in the JP2 file format (the JPEG2000 file format).
- The PNG (ISO/IEC 15948:2003) standard. PNG shall not be used in its interlaced mode and not for images that have been JPEG compressed before.

The *IED* in the captured photo shall be at least 90 pixels for legacy applications. If an issuer considers the design of a new passport application process, the new *IED* should be at least 240 pixels. Examples for a new process could be live capturing, digital submission without analogue intermediate steps, or increasing the size of the printed photograph to be scanned.

See Table 5. See Figure 11 for an illustration of the *IED* measurement.

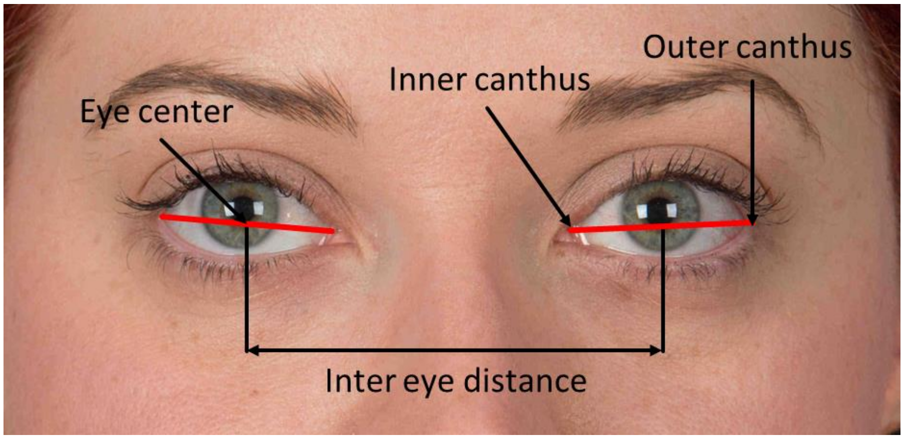


Figure 11 — IED measurement.

NOTE: Be aware that the eye center is not necessarily the center of the pupil.

NOTE: A typical real *IED* (distance measured at the face) is between 60 mm and 65 mm.

Table 5 – *IED* capturing requirements and recommendations.

Criterion: Live Capture IED	Requirement	*IED* ≥ 90 pixels
	Best Practice	*IED* ≥ 240 pixels
Criterion: Scanned Image IED	Requirement	*IED* ≥ 90 pixels
	Best Practice	*IED* ≥ 240 pixels
Criterion: Electronic Submission IED	Requirement	*IED* ≥ 90 pixels
	Best Practice	*IED* ≥ 240 pixels
Criterion: Issuer Repository IED	Requirement	*IED* ≥ 90 pixels
	Best Practice	*IED* ≥ 240 pixels
Criterion: MRTD Chip Storage IED	Requirement	*IED* ≥ 90 pixels
	Best Practice	*IED* ≥ 120 pixels

NOTE: This pixel count is specified for the live captured portrait only. For stored images on the chip, see Clause 9.2 and Table 11.

All images shall have sufficient focus and depth of field to maintain the required level of details. The camera shall be capable of accurately rendering fine contrasted facial details, such as wrinkles and moles, as small as 1 mm in diameter on the face.

The focus and depth of field of the camera shall be set so that the subject's captured image is in focus from nose to ears. In most cases, a depth of field of 15 cm will be sufficient. See Table 6. The background behind the subject may be out of focus. Proper focus and depth-of-field will be assured by either using the camera auto focus function with manual aperture settings or by pre-focusing the lens at the distance of the subject's eyes and by selecting an appropriate aperture (F-stop) to ensure a depth-of-field of the distance from a subject's nose to ears. See Annex E.

Table 6 – Depth of field requirements and recommendations.

Criterion:	Requirement	Nose to ears

Depth of field	Best Practice	15 cm from nose level

EXAMPLE: An aperture of f/8 for an 80 mm lens at a distance of 2,5 m provides a depth of field of 15 cm. An aperture of f/16 for a 50 mm lens at a distance of 1,2 m provides a depth of field of about 18 cm.

A simplified visual compliance check method requires that the individual millimetre markings of rulers placed on the subject's nose and ear facing the camera can be seen simultaneously in a captured test image. See Figure 12. This quality assurance method should be used for quality assurance field checks from time to time. A more systematic test method is described in Annex D.

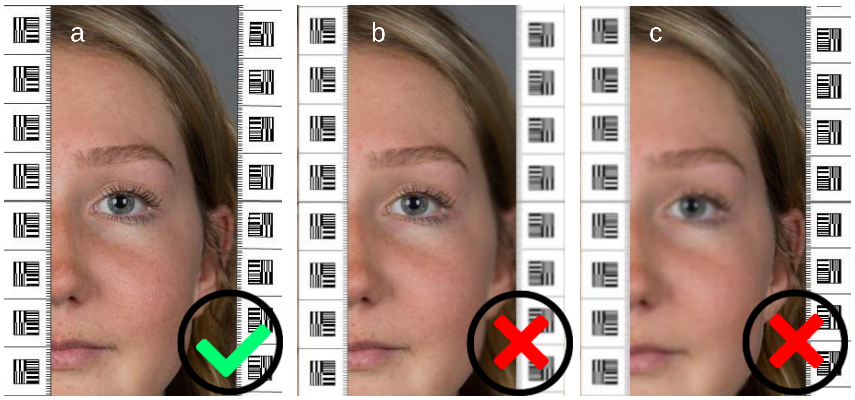


Figure 12 — a) Part of a compliant portrait, b) Image sharp at nose only, c) Image sharp at the ears only.

5.2.5 Background

The background surface behind the subject shall be plain, and shall have no texture containing spots, lines or curves that will be visible in the captured image. The background shall have a uniform colour. There may be gradual changes from light to dark luminosity in a single direction, although this may make it more difficult to remove the background during the document production process (see Clause 8).

A typical background for the scene is grey with a plain, dull flat surface. Plain light coloured backgrounds such as light blue or white may be used as long as there is sufficient distinction between the face/hair area and the background. Camera colour settings should not be shifted depending on the background colour. See Figure 13.

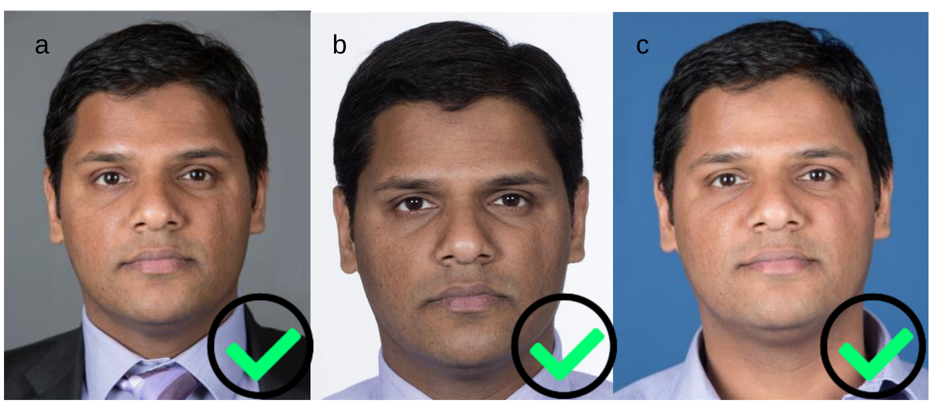


Figure 13 — Compliant portraits with a) Grey, b) Light grey, and c) Blue background.

The boundary between the head and the background should be clearly identifiable around the entire subject with the exception of very large hair volume. See Figure 14. A boundary that is not clearly visible can have a negative impact on the production process which often requires background removal (see Clause 8).

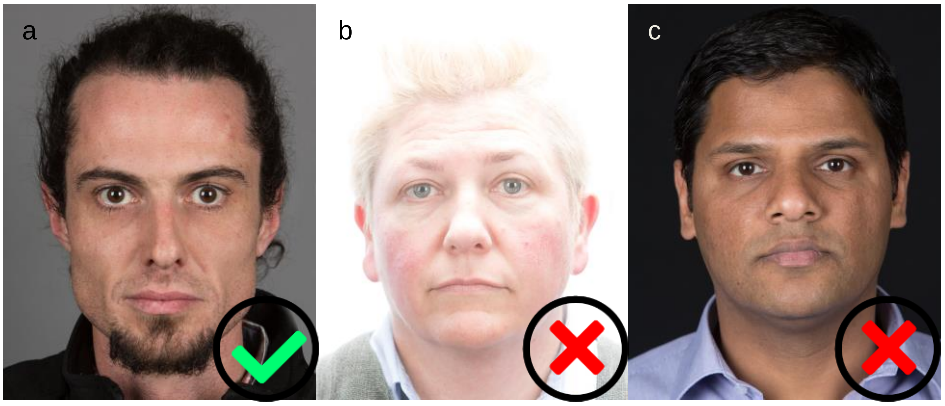


Figure 14 — a) Compliant portrait, b) Image contrast unacceptably low, c) Too low background contrast.

Shadows should not be visible on the background behind the face image. In particular, there shall not be asymmetric shadows. There shall not be any objects visible in the background like supporting persons, chair backs, furniture, carpets, patterned wall papers or plants. For examples see Figure 15.

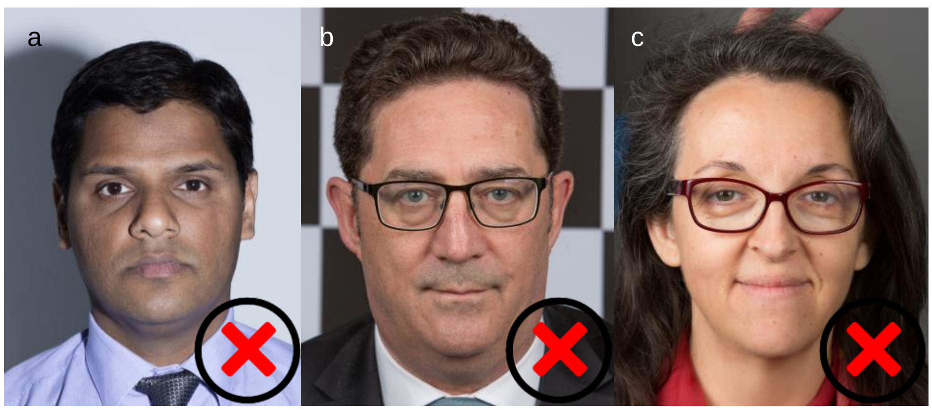


Figure 15 — a) Asymmetric shadow on the left, b) Inhomogenous background, c) Body parts visible behind the head.

5.2.6 Lighting

Portraits shall have adequate and uniform illumination. Lighting shall be equally distributed on the face, in particular symmetrically, i.e., there is no difference between the brightness of the right and left side of the face. There shall not be significant direction of the light from the point of view of the camera.

The measured *EV* at four spots on a subject's face; the left and right cheeks, forehead, and chin, should be the same. An *EV* difference of at most one F-stop or one shutter speed step is acceptable. If one or some of these four spots are covered by hair, e.g., the forehead by the hairstyle or the chin by a beard, these spots can't be evaluated. The appropriate illumination setup of the scene should be verified from time to time. The person being used for these tests should not have a hairstyle covering the forehead or the cheeks, or a beard.

The uniformity measurement should be done as specified below. It is not intended to be used for every single image. See Figure 16 for a visualization of that measurement. Automated quality assurance software, e.g., for registration offices or kiosks, should be implemented accordingly. However, such software should also consider exceptions due to hair on the forehead, beard, facial anomalies and the like.

1. Determine the four measurement zones on the forehead, the cheeks, and the chin. These locations are determined as follows:
   a. Connect the two eye centers (feature points 12.1 and 12.2 from ISO/IEC 14496-2). The IED is the length of the connecting line H. The point M is the midpoint of this line.
   b. Connect M with the mouth midpoint (feature point 2.3 from ISO/IEC 14496-2). EM is the length of the connecting line V. Note that the two lines do not need to be rectangular.
   c. MP, the side length of the four squared measurement zones, is defined to be 0,3 IED.
   d. The centre of the forehead measurement zone F is located at a distance of 0,5 EM upwards from M on V.
   e. The centre of the chin measurement zone C is located at a distance of 1,5 EM downwards from M on V.
   f. The top left corner of the right (from the person) cheek measurement zone R is located at a distance of 0,5 EM downwards from M on V and 0,5 IED to the left of M on H (looking at the person).
   g. The top right corner of the left (from the person) cheek measurement zone L is located at a distance of 0,5 EM downwards from M on V and 0,5 IED to the right of M on H (looking at the person).
2. For all colour channels, measure the mean intensity values MI for the measurement zones F, C, L, and R.
3. For all channels separately, the lowest MI (of F, C, L, and R) in that channel shall not be lower than 50% of the highest MI (of F, C, L, and R).

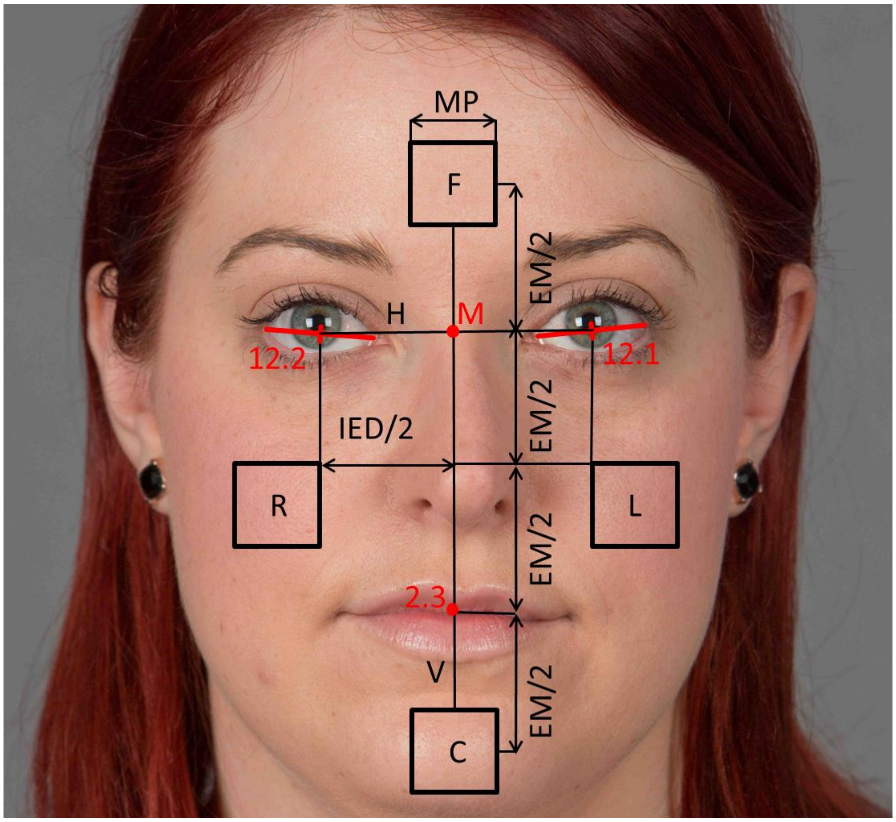


Figure 16 — Location and size of the intensity measurement zones.

The measures for the illumination intensity and requirements to them are listed in Table 7.

Table 7 – Measures for the illumination intensity compliance check.

Term	Description	Requirement
12.1	Feature point at left eye centre	
12.2	Feature point at right eye centre	
H	Line connecting 12.1 and 12.2	
IED	Length of H between 12.1 and 12.2	*IED* ≥ 90 pixels
M	Midpoint of H between 12.1 and 12.2	
2.3	Feature point at mouth centre (with closed mouth the same as 2.2)	
V	Line connecting *M* and 2.3, *V* and *H* do not need to be orthogonal	
EM	Length of *V* between *M* and 2.3	
MP	Side length of the squared measurement zones	*MP* = 0,3 *IED*
F	Forehead measurement zone, located at a distance of 0,5 *EM* upwards from *M* on *V*	
C	Chin measurement zone *C*, located at a distance of 1,5 *EM* downwards from *M* on *V*	
R	Right (from the person) cheek measurement zone *R*, its top left corner is located at a distance of 0,5 *EM* downwards from *M* on *V* and 0,5 *IED* to the left of *M* on *H*	
L	Left (from the person) cheek measurement zone *L*, its top right corner is located at a distance of 0,5 *EM* downwards from *M* on *V* and 0,5 *IED* to the right of *M* on *H*	
MI	Mean intensity value measured for every channel separately	max ≤ 2 × min (per channel)

While it is understood that massive shadows on parts of the face will obscure facial details important for identification, having no shadows at all will result in a non-natural appearance. In such a case, the face will appear flat and without surface features. Appropriate shadows help distinguish the shape of the nose, eye areas, forehead, cheeks, chin and so on. Furthermore, lighting and shadows are necessary to show details around the eyes, wrinkles, and scars. There shall not be extreme dark shadow visible on the face, especially around the nose, in the eye sockets, around the mouth, and between mouth and chin that obscure facial details important for inspection. The brightness shall be nearly the same on both sides of the face, left and right. All features in the face shall be clearly recognizable, and the volume effect especially around nose and eyes shall render the reality. See Figure 17.

EXAMPLE: To comply with this requirement, the illumination elements can be aligned in an angle of approximately 35° off the axis between camera lens and face center. Descriptions of sample illumination layouts are given in Annex A.

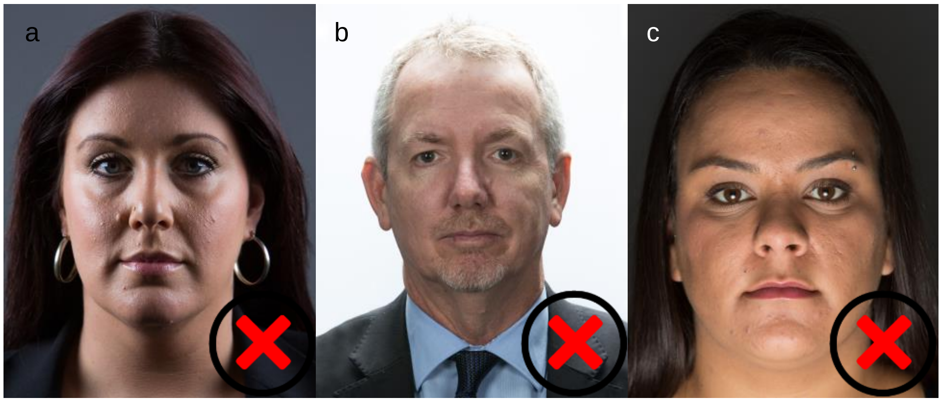


Figure 17 — Non-compliant portraits: a) Side illumination, b) Top illumination, c) Bottom illumination.

Flashes should only be used for indirect illumination. Issuers may exclude the use of flashes. If portraits are captured using flashes, care should be taken to verify that the eyes of the subject are open. As long as the requirements for the portrait from Clause 5 are maintained, one or more flashes or a large surface flash may be used. There shall not be any shadows at the face or in the background of the portrait that obscure facial details important for inspection. Illumination shall not cause any red eye effect visible in the eyes or other lighting artefacts such as spots from a ring flash reducing the visibility of the eyes.

A high colour rendering index is recommended for illumination. See Clause 5.2.9 for details.

The captured image shall contain minimal reflections or bright spots. Diffused lighting, multiple balanced sources or other appropriate lighting methods should be used. A single bare point light source like a camera mounted direct flash shall not be used for imaging. Lamp reflectors or other technologies that provide non-point illumination should be used.

5.2.7 Contrast

For each patch of skin on the person's face, the gradations in textures shall be clearly visible, i.e., being of reasonable contrast. Whites of eyes shall be clearly light or white (when appropriate) and dark hair or facial features (when appropriate) shall be clearly dark. Generally, the portrait shall have appropriate brightness and good contrast between face, hair and background. See Figure 18.

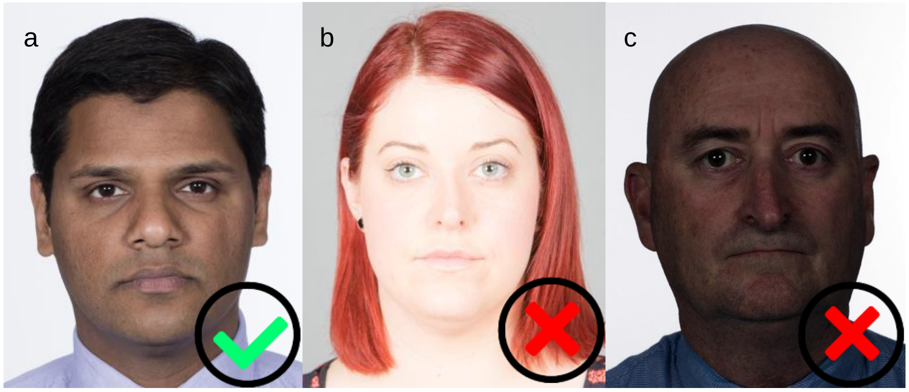


Figure 18 — a) Compliant portrait, b) Over exposure, c) Under exposure.

5.2.8 Dynamic range

The dynamic range of the image should have at least 50% of intensity variation in the facial region of the image. The facial region is defined as the region from crown to chin and from the left ear to the right ear. This recommendation may require an adjustment of the equipment settings on an individual basis when the skin tone is excessively light or dark. In the rectangle between the ISO/IEC 14496-2 feature points

- 2.1: Bottom of the chin,
- 10.9: Upper contact point between left ear and face,
- 10.10: Upper contact point between right ear and face, and
- 11.1: Middle border between hair and forehead,

all colour channels should have at least 50% of intensity variation. As this may be difficult to achieve, best efforts should be made to get as close as possible to that requirement. See Figure 19 for an illustration of the recommended measuring zone and Figure 20 for examples of good and bad quality images.

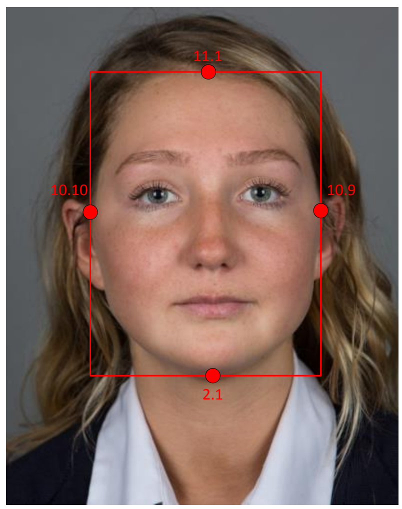


Figure 19 — Illustration of the recommended dynamic range measuring zone.

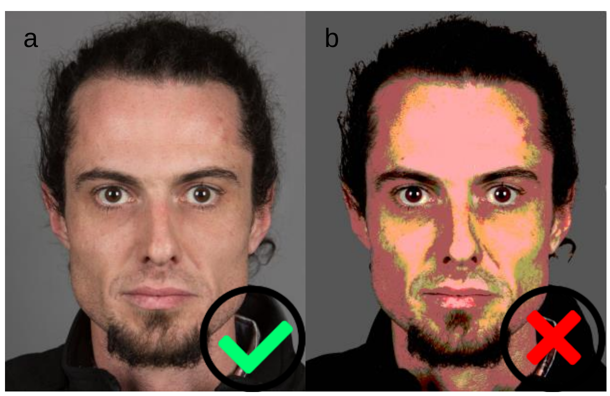


Figure 20 — a) Compliant portrait, b) Too low dynamic range.

5.2.9 Colour

All images should be captured in colour. Newly designed enrolment should capture colour images only.

The captured portrait shall be a true-colour representation of the holder in a typical colour space such as sRGB as specified in IEC 61966-2 [R10]. Other true-colour representations may be used as long as the colour profile is embedded in the image.

The sensor of the camera shall capture the entire visible wavelength, basically the wavelength between 400 nm and 700 nm. It allows rendering correctly the natural colours seen by humans. Unnaturally coloured lighting, i.e., yellow, red, etc., shall not be used. Care should be taken to correct the white balance of image capture devices. The lighting shall produce a face image with naturally looking flesh tones when viewed in typical examination environments. See Figure 21.

Figure 21 — a) Compliant portrait, b) Colour deviations, c) Excessive saturation in hue, saturation, & luminance (HSL) colour space.

The RGB values from the capturing device should be converted to an appropriate RGB space as required by the data format.

Dedicated near infra-red cameras shall not be used for image acquisition.

Colour calibration using an 18% grey background or other method such as white balancing should be applied.

White balance shall be properly set in order to achieve high fidelity skin tones. Quality assurance measurements of light conditions and camera system response should be made when a recommended CIE Standard Illuminant D65 illuminant [R7] or a similar continuous spectrum daylight illuminant and a camera and/or camera control software are used to take pictures. In practice it is necessary to reduce the ambient light emanating from uncontrolled daylight sources, fluorescent or similar light sources and reflections from surfaces.

Imaging fidelity measurements for photo studio and stationary registration office installations may be done either using a light spectrum analyzer to define the spectral characteristics of the illuminants or analyzing measurement target images using software applications.

Annex C contains methodology for measuring colour quality and recommended values.

Colour quality should be measured in terms of colour error using the CIEDE2000 formula (deltaE2000) of a standardized test pattern according to the methodology in Annex C. The average deltaE2000 of all colour patches should not exceed 4 for scanners and 10 for camera systems. The maximum deltaE2000 for any colour patch should not exceed 15 for scanners and 20 for camera systems. Measured CIELAB L*a*b* human skin tone a* and b* values shall be positive as shown in Annex C. See [15, 16] for explanations. Negative a* and b* values are acceptable only for medical reasons.

5.2.10 Noise

The enrolment should be made in a controlled scene; the picture should be captured with high signal-to-noise ratio. Noise is not information contained in the original scene but is created by the electronics due to a too high level of amplification.

EXAMPLE: ISO sensitivity settings at values of ISO 100 and ISO 200 typically reduce noise; for high-quality cameras ISO 400 and ISO 800 may also be used.

EXAMPLE: Noise can be minimized by correct exposure at a low ISO setting.

The Ratio of Signal to Noise (SNR) is one indicator of the overall ability of a collection system to accurately capture a subject's appearance. Unwanted variations in the response of an imaging system (i.e., noise) are inherent in the capture process of a digital representation of a physical scene and arise from the interplay between the system components (e.g., sensor, lens) and the capture environment (e.g., subject illumination). Reducing overall noise to improve the SNR benefits human examiners and automated face analysis systems which rely on high-quality subject images. SNR should be computed as prescribed in ISO 15739:2013 Section 4.7, which incorporates a human visual model to calculate the human observable (i.e., perceived) SNR of the overall collection system.

Commercial software designed for use by photo studios and registration office imaging systems are available with accompanying standard test targets for computing SNR.

5.2.11 Filters

Polarization filters shall not be used in front of the light sources. Linear polarization filters shall not be used in front of the camera lens as they interfere with autofocus cameras and thus reduce or remove skin texture information which might be used by facial image comparison algorithms. Circular polarizing filters decrease reflections that show up in eyeglasses and may be used in front of the camera lens.

5.3 Subject conditions

5.3.1 Pose

The subject should be instructed to look directly at the camera and to keep his or her head erect. Typically people are able to adopt such a position if instructed. Care should be taken to maintain the full frontal pose as well as possible. See Figure 22.

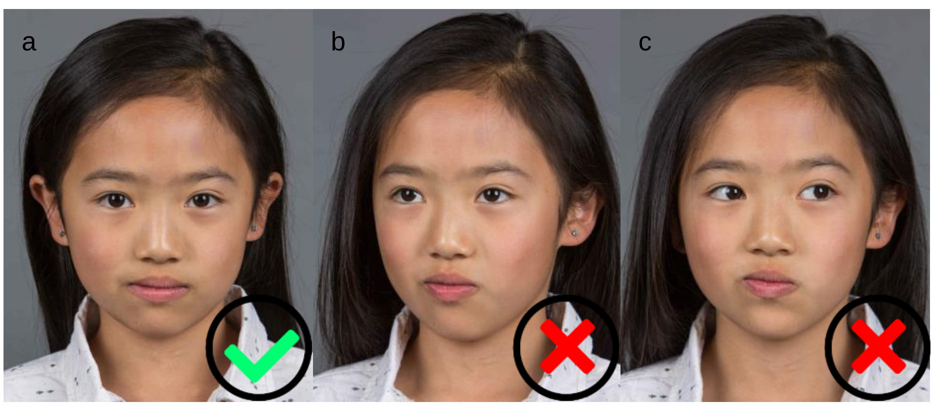


Figure 22 — a) Compliant portrait, b) Head not aligned toward the camera, c) Eyes not aligned toward the camera.

The shoulders shall be square on to the camera, parallel to the camera imaging plane. Portrait style photographs where the subject is looking over the shoulder shall not be used. See Figure 23.

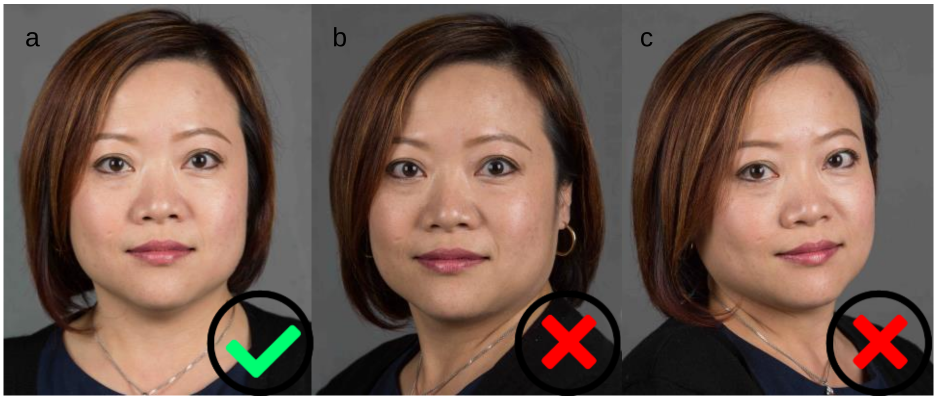


Figure 23 — a) Compliant portrait, b) Shoulders not aligned toward the camera, c) Shoulders not aligned toward the camera, head tipped to the side.

The pitch of the head shall be less than ± 5° from frontal. The yaw of the head shall be less than ± 5° from frontal. The roll of the head shall be less than ± 8°, it is recommended to keep it below ± 5°. Any stronger pose deviation may have negative impact on facial recognition error rates. Therefore, effort should be spent to ensure that all angles are as small as possible. See Table 8. For an illustration of the angles see Figure 24. For samples showing correct pose and pose deviations see Figure 25.

Table 8 – Pose angle requirements and recommendations.

Criterion: Pose angle	Requirement	Pitch ≤ ± 5°, yaw ≤ ± 5°, roll ≤ ± 8°
	Best Practice	Pitch ≤ ± 5°, yaw ≤ ± 5°, roll ≤ ± 5°

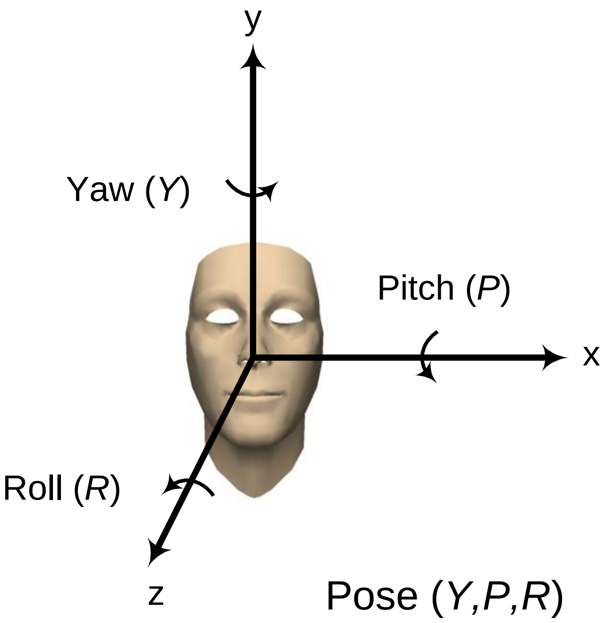


Figure 24 – Definition of pose angles with respect to the frontal view of the subject.

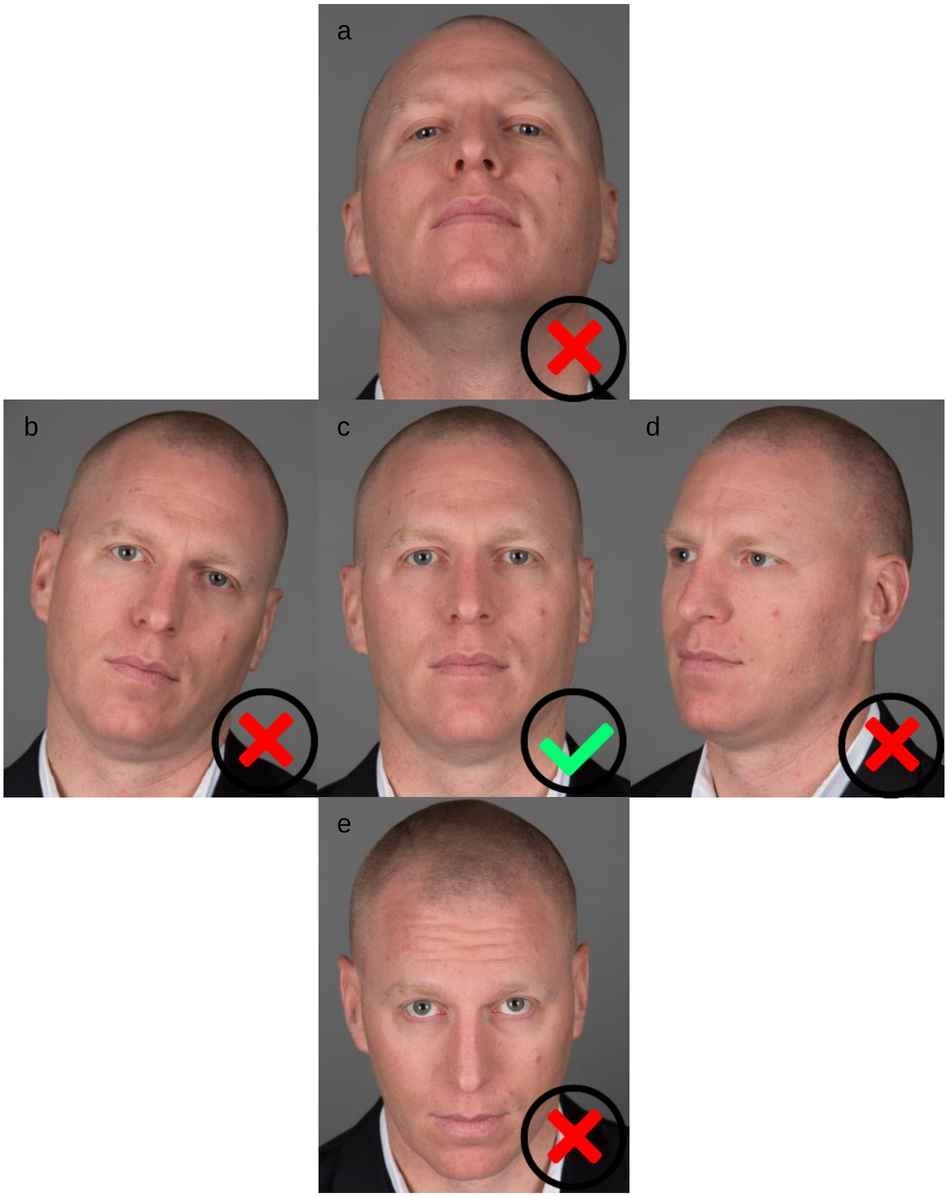


Figure 25 —a) Too large pitch up, b) Too large roll angle, c) Compliant portrait, d) Too large yaw angle, e) Too large pitch down.

5.3.2 Expression

The face shall have a neutral expression; in particular the person shall not smile. The mouth shall be closed; the teeth shall not be visible. A smile is not allowed, even with closed jaw. The eyebrows shall not be raised. Squinting and frowning shall not be visible. See Figure 26.

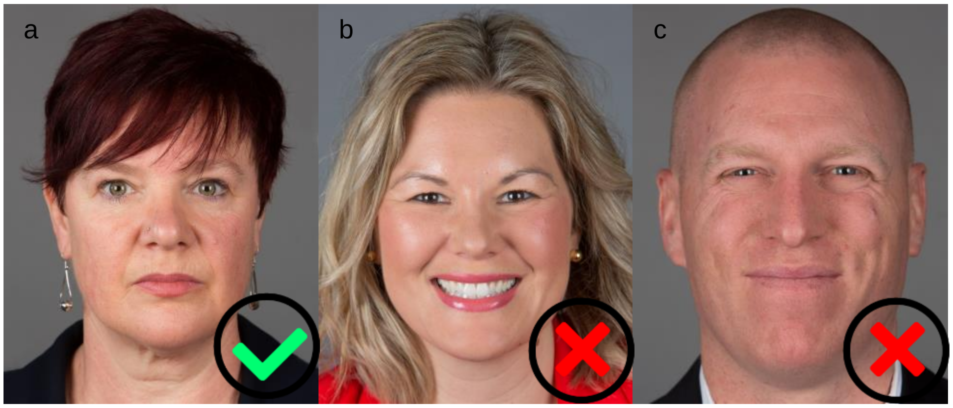


Figure 26 — a) Compliant portrait, b) Smiling person, c) Person smiling with closed mouth.

The mouth is considered to be closed if the distance A between the inner borders of the lips (distance between feature points 2.2 and 2.3) is less than 50% of the thickness of the lower lip B (distance between the feature points 2.3 and 8.2). See Figure 27.

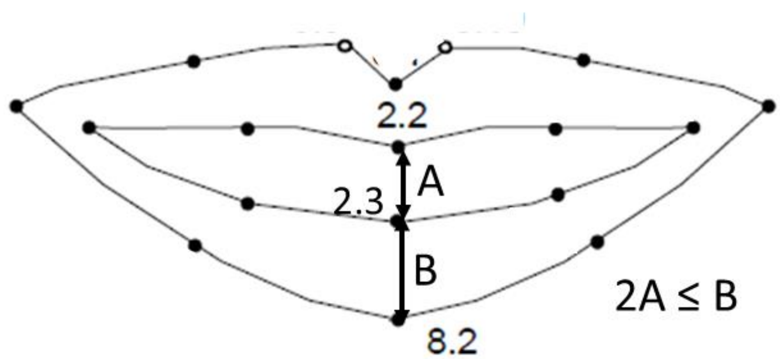


Figure 27 — Definition of a closed mouth

5.3.3 Eye visibility

Both eyes shall be opened naturally, but not forced wide-opened. Pupils and irises, including iris colour, shall be completely visible, although there may be exceptions due to ethnicity or other individually specific reasons. The eyes shall look into the camera unless there are medical conditions preventing this. There should not be strong shadows in the eye-sockets. See Figure 28.

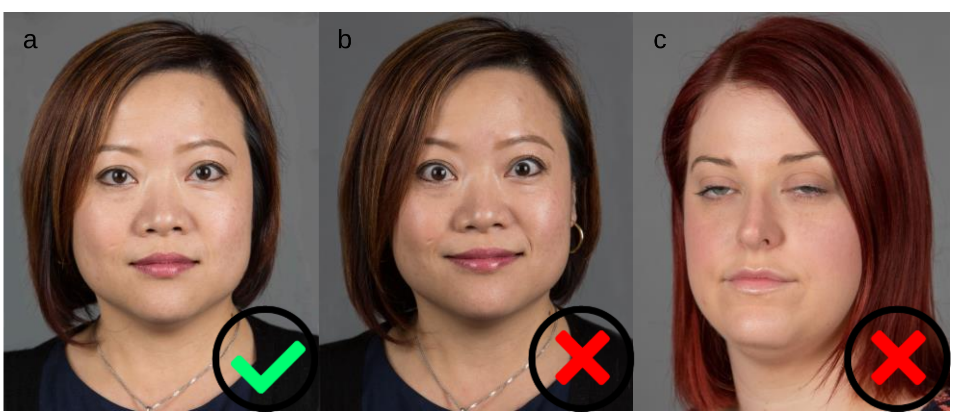


Figure 28 — a) Compliant portrait, b) Unnaturally wide opened eyes, c) Eyes not fully opened.

Any lighting artefacts present on the region of the eyes shall not obscure eye details such that identification becomes difficult. Lighting artefacts shall not be larger than 15% of the area of the iris. If there are unacceptable reflections, the illumination should be relocated appropriately. The pitch shall not be increased by moving the head forward.

Examples of setups preventing or at least reducing lighting artefacts are given in Annex A.

The Eye Visibility Zone (*EVZ*) is defined as the covering rectangle having a distance *V* of at least 5% of the *IED* to any part of the visible eye ball. See Figure 29. The *EVZ* shall be visible and unobscured.

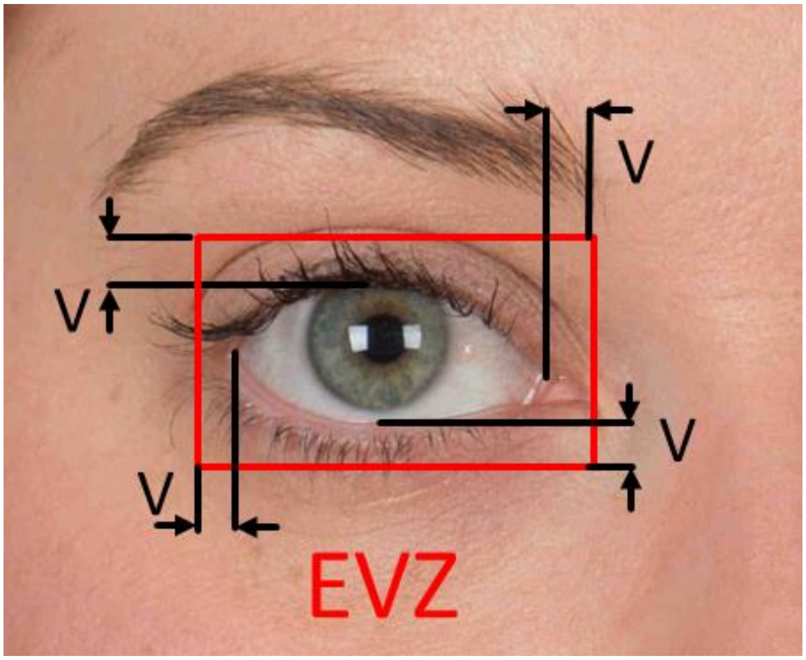


Figure 29 — Illustration of the *EVZ* indicating the distance *V* of the covering rectangle to the visible parts of the eye ball.

Contact lenses changing the appearance of the iris including the size and the shape shall not be worn. The pattern of the lens shall not exceed the limbus.

5.3.4 Accessories: Glasses

If glasses are permitted by the issuer, subjects may wear glasses during image capture if they typically do so. Glasses other than those worn due to ametropia shall not be worn. Reading glasses shall not be worn during image capture. The lens area of glasses shall be made of fully transparent material. Tinted glasses, sunglasses, and glasses with polarization filters shall not be worn. An exception applies when the subject asserts a medical reason to retain glasses which are not fully transparent. If glasses are worn that tint automatically under illumination, they shall be photographed without tint by tuning the direct illumination or background lighting. In cases where the tint cannot be reduced the glasses shall be removed or the subject should be asked to use other glasses. See Figure 30. A circled yellow "P" indicates compliance depending on the acceptance policy of the issuer.

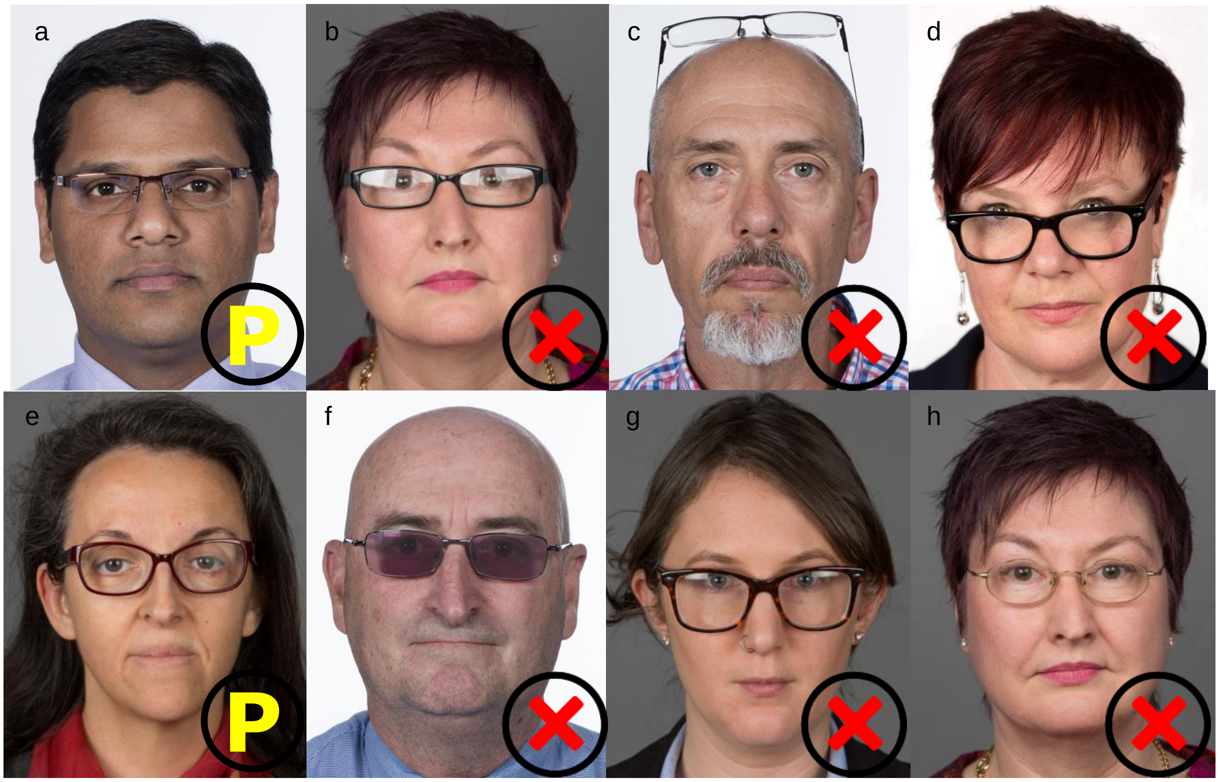


Figure 30 — a) Compliant portrait, b) Strong reflections on glasses, c) Glasses worn on the head, d) Frame crossing the eyes, e) Compliant image, f) Sunglasses, g), h) Frames partially covering the *EVZ*.

Any lighting artefacts present on the region of the glasses shall not obscure eye details such that identification becomes difficult. Glasses may be repositioned to eliminate lighting artefacts, but frames shall not obscure eye details. The pitch shall not be increased by moving the head forward.

Rims and frames of glasses shall not obscure the eyes as well as the *EVZ*. The irises of both eyes shall be visible to the same extent as without glasses. Frames should not be thicker than 5% of the *IED* (typically 3-4 mm). A person wearing heavier frames should be asked to use other glasses or to remove their glasses.

5.3.5 Accessories: Head coverings

The region of the face, from the crown to the base of the chin, and from ear-to-ear, shall be clearly visible. Special care shall be taken in cases when veils, scarves or head covering cannot be removed for religious reasons to ensure these coverings do not obscure any facial features and do not generate shadow. Head coverings shall not be accepted except in circumstances specifically approved by the Issuing State of the MRTD. Such circumstances may be religious, medical or cultural. If head coverings are allowed, they shall be firm fitting and of a plain uniform colour with no pattern or no visible perforations and the region between hair lines, both forwards of the ears and chin including cheeks, mouth, eyes, and eyebrows shall be visible without any distortion or shadows. The elliptically shaped region between the following facial feature points as defined in ISO/IEC 14496-2 shall be visible without any intensive shadows:

- 2.1: Bottom of the chin,
- 10.9: Upper contact point between left ear and face,
- 10.10: Upper contact point between right ear and face, and
- 11.1: Middle border between hair and forehead.

An issuer may or may not require that the ears are visible. The capture process should minimize shadows and obscuration of features in the facial region. This might involve adjustment of the head coverings. See Figure 31.

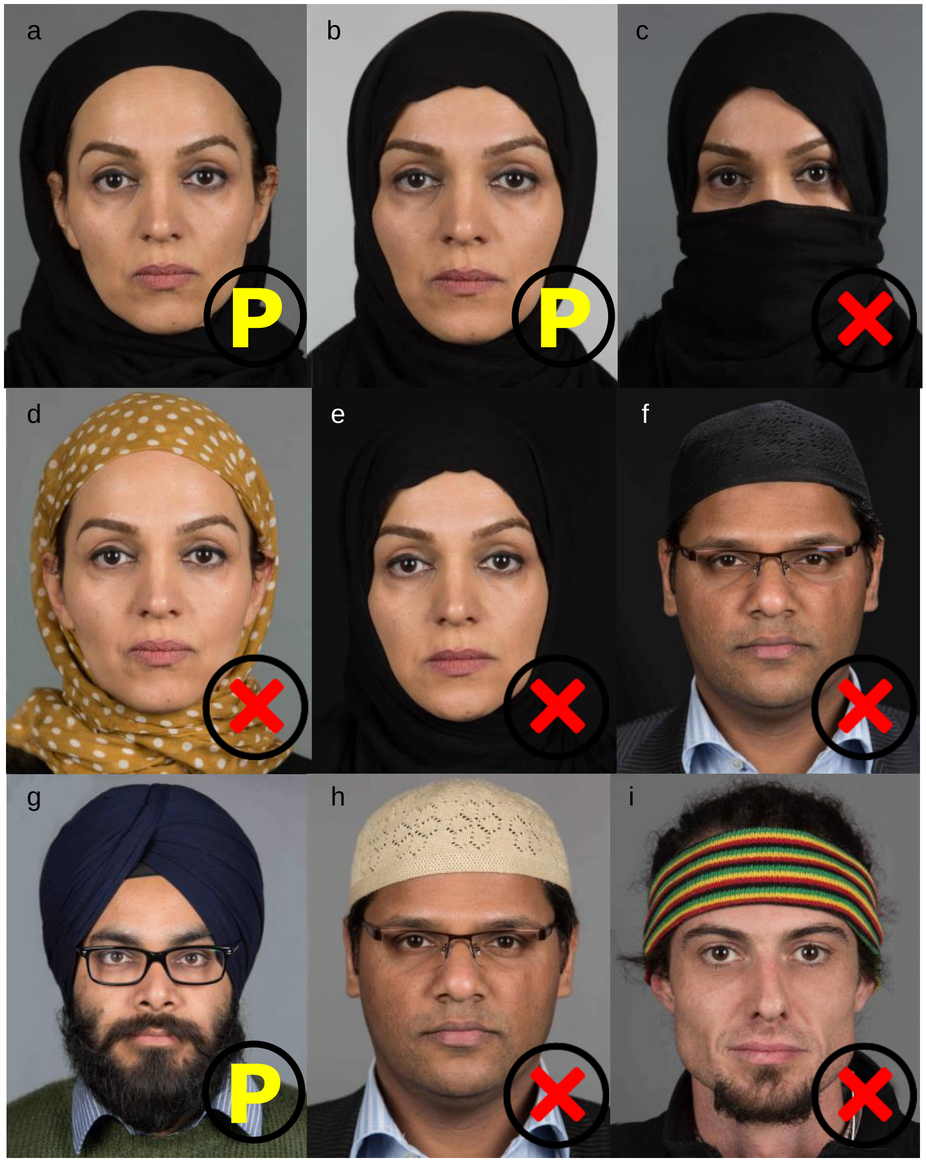


Figure 31 — a), b) Compliance depends on issuer policy, c) Face not completely visible, d) Non-uniform head covering, e), f) Low background contrast, g) Compliance depends on issuer policy, h) Visible perforation, i) Non-religious head coverage.

5.3.6 Accessories: Facial ornamentation

Facial ornamentation which obscures the face shall not be present. Concerning facial ornaments not obscuring the face, the issuer may use its discretion as to the extent to which facial ornaments may appear in the portrait. In any case, only permanently worn facial ornaments may appear in the portrait. See Figure 32.

Figure 32 — a), b), c): Compliant facial ornamentation.

5.3.7 Style: Makeup, hair style

People usually try to look better than normal in an ID photo. In some extreme cases an excessive use of make-up affects computerized as well as human facial recognition capabilities. Therefore the subject should only wear typical every day make up.

There shall be no dirt visible on the face in a captured portrait. The hair of the subject shall not cover any part of the eyes. The hair should not cover any part of the *EVZ*. See Figure 33. Eye patches shall not be worn unless required for a medical reason.

Figure 33 — a) Compliant portrait, b) Hairs and glasses close to the *EVZ*, c), d) Hair partly covering the *EVZ*.

5.4 Portrait dimensions and head location

The head shall be centred in the final portrait as described in this Clause. The referenced feature points are defined in ISO/IEC 14496-2. See Figure 34, Figure 35, and Table 9.

The image width *A* to image height *B* aspect ratio should be between 74% and 80%. The imaginary line *H* is defined as the (almost horizontal) line through the eye centres of the left (feature point 12.1) and the right eye (feature point 12.2).

The centre of *H* is the face midpoint *M*. The horizontal distance M_h between the left image border and *M* shall be between 45% and 55% of *A*. The vertical distance M_v between the top image border and *M* shall be between 30% and 50% of *B*.The mouth centre (feature point 2.3) and *M* define the imaginary (almost vertical) line *V*. Note, that *V* and *H* are not necessarily perpendicular.

The head width *W* is defined as the distance between the two imaginary lines parallel to the line *V*; each imaginary line is drawn between the upper and lower lobes of each ear (feature points 10.2/10.6 for the right and 10.1/10.5 for the left ear). The *W* to *A* ratio shall be between 50% and 75%. This constraint is more important than including the entire hairline in the photograph for subjects with large hair volume.

The head length *L* is defined as the distance between the base of the chin (feature point 2.1) and the crown (feature point 11.4) measured on the imaginary line *V*. If these feature points are not exactly located at *V*, the vertical projection of them to *V* shall be used. The *L* to *B* ratio shall be between 60% and 90%.

Often, the location of crown, chin, or ears can't be determined precisely. In such a case, a good guess shall be made.

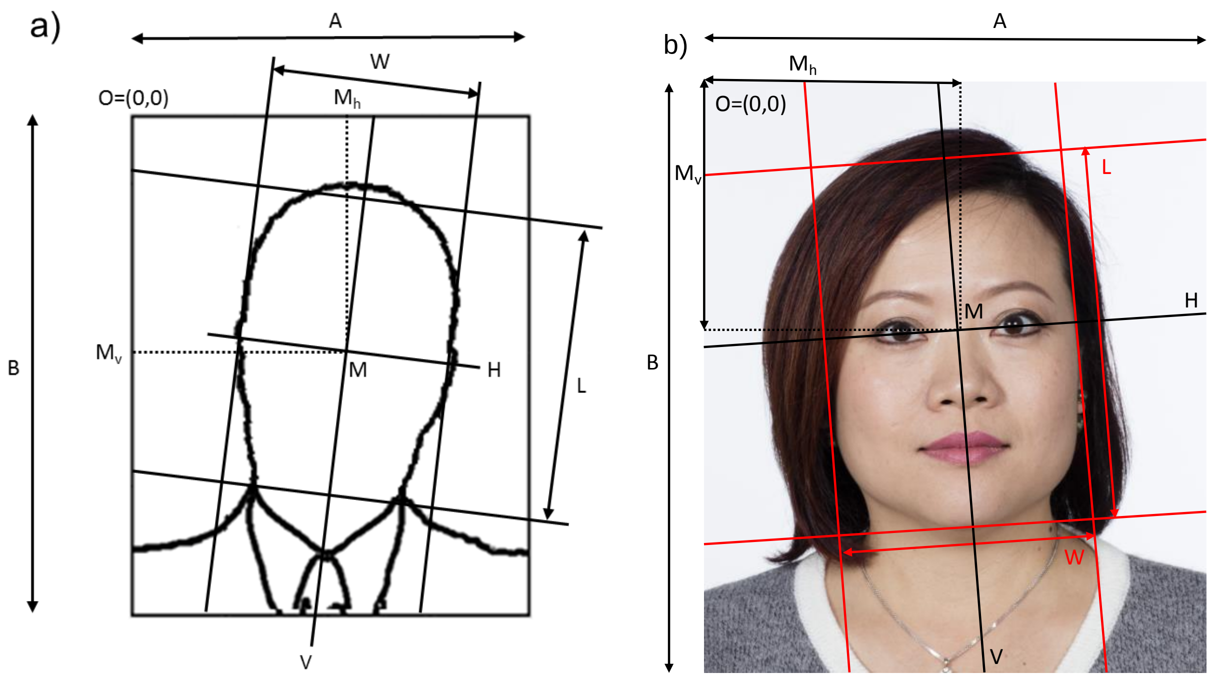


Figure 34 — a) Abstract geometric characteristics of a portrait, b) Applied to a sample.

Figure 35 — a) Compliant portrait, b) Invisible crown, c) Invisible chin, d) Invisible crown and chin.

Table 9 — Geometric portrait requirements

Term	Description	Requirement
A	Image width	
B	Image height	$74\% \leq A/B \leq 80\%$
H	Line through the centres of the left (feature point 12.1) and the right eye (feature point 12.2)	
M	Face centre (midpoint of *H*)	
M_h	Distance from the left side of the image to *M*	$45\% \leq M_h/A \leq 55\%$
M_v	Distance from the top of the image to *M*	$30\% \leq M_v/B \leq 50\%$
V	Line through mouth centre (feature point 2.3) and *M*	
W	Head width: Distance between the two imaginary lines parallel to the line *V*; each imaginary line is drawn between the upper and lower lobes of each ear (feature points 10.2/10.6 for the right and 10.1/10.5 for the left ear)	$50\% \leq W/A \leq 75\%$
L	Head length*:* Distance between the base of the chin (feature point 2.1) and the crown (feature point 11.4) measured on the imaginary line *V*, if these feature points are not exactly located at *V*, the vertical projection of them to *V* shall be used	$60\% \leq L/B \leq 90\%$

Figure 36 shows a typical example of a portrait. In Figure 36 a), the intersection of the two rectangles marks the region where the point *M* shall be located. In Figure 36 b) the smaller rectangle in the right portrait shall be completely included in the head; the head itself shall be completely included in the larger rectangle. Note that the locations of these two rectangles do not depend on the location of *M,* the rectangles can be moved freely as long as they stay parallel to the borders of the image. Figure 37 gives samples where the faces do not fit into the larger rectangle or do not fill the smaller rectangle.

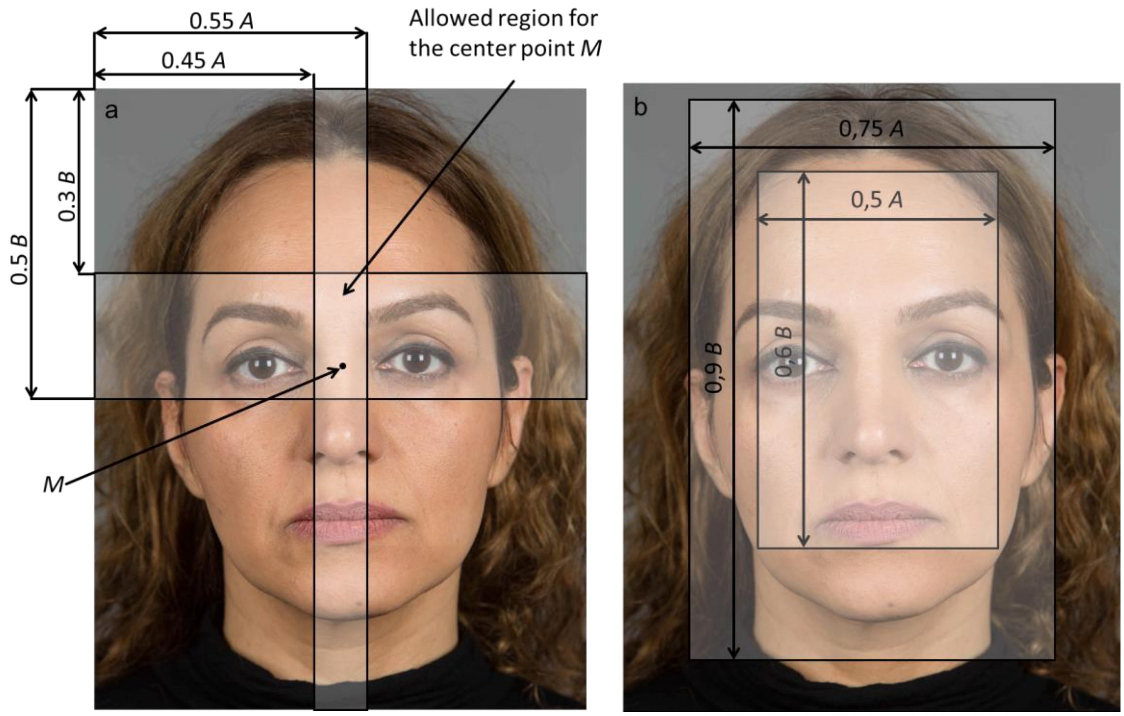


Figure 36 — Sample portraits with the respective minimal and maximal head dimensions: a) true location and allowed region for the center point *M*, b) minimal size (inner rectangle) and maximal size (outer rectangle, note that the position including the relative position of the two rectangles is not fixed) of *W* and *L* depending on *A* and *B*.

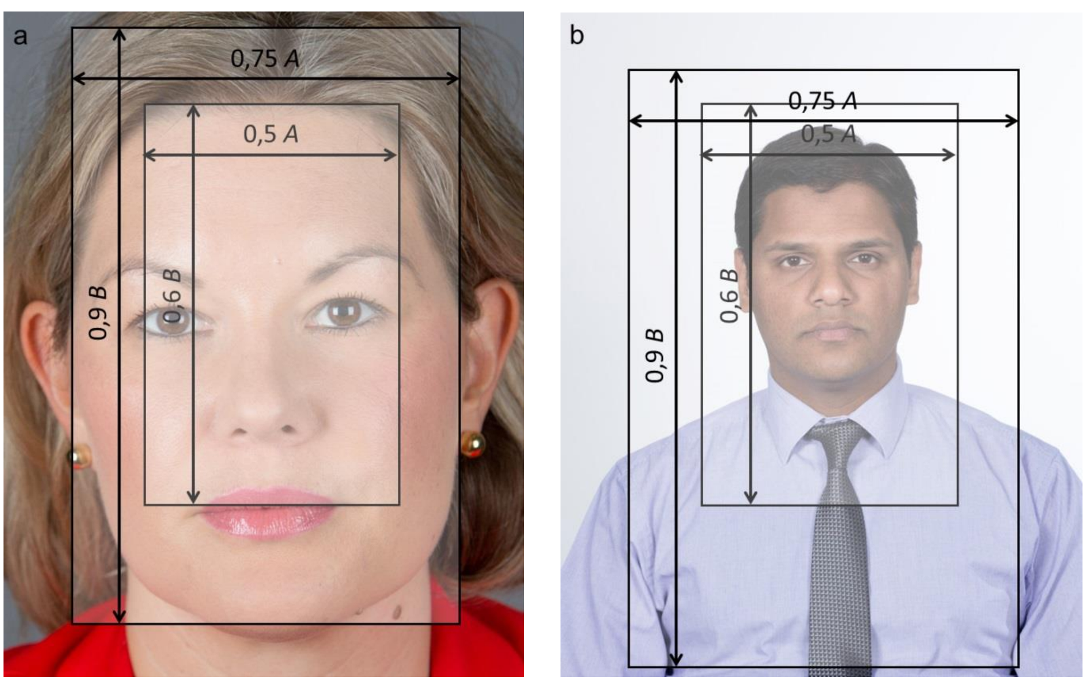


Figure 37 — Sample portraits not complying with minimal and maximal head dimensions: a) face too large and doesn't fit into the larger rectangle, b) face too small and does not cover the entire smaller rectangle.

5.5 Children

5.5.1 General

This Clause specifies additional guidance for capturing portraits of children. Care should be taken to capture such images according to the specifications in Clause 5.2, 5.3, and 5.4, however, sometimes this is not possible or would cause huge discomfort. Therefore, some requirements from these Clauses may be relaxed for children as specified below. See Figure 38, Figure 39, and Figure 40 for sample images.

5.5.2 Children below one year

Deviating from the specifications in Clause 5.2, babies under one year should be in an upright position, but it is acceptable to capture the portrait with the baby lying on a white or plain light-coloured blanket which conforms to the requirements in Clause 5.2.5 and 5.2.9. Alternatively, the baby may be placed in a baby seat as long as the background behind the head of the baby conforms to the requirements above and no portions of the baby seat are visible in the portrait.

Deviating from the specifications in Clause 5.3.3, it is not necessary that babies under one year have their eyes open.

Hands, arms, and other body parts of an assisting person used to support the positioning of the subject, e.g., parents supporting their child, shall not be visible in the image. Shadows of these assistant parts shall not be visible on the portrait or in the background.

5.5.3 Children below six years

Deviating from the specifications in Clause 5.3.1, children aged six and under shall face the camera within an angle of ±15° in pitch, yaw, and roll. Deviating from the specifications in Clause 5.3.2, children aged six and under do not need to have a neutral expression. For infants under the age of six, images are acceptable as long as the infant is awake, has his or her eyes open, there are no other people or objects in the photo and the background is uniform and portrait meets the colour requirements in Clause 5.2.9.

5.5.4 Children below eleven years

Deviating from the specifications in Clause 5.4 for children of up to eleven years, *L/B* shall be between 50% and 90%. Furthermore, M_y/B shall be between 30% and 60%.

Figure 38 — a), b), c), d) Compliant child portraits.

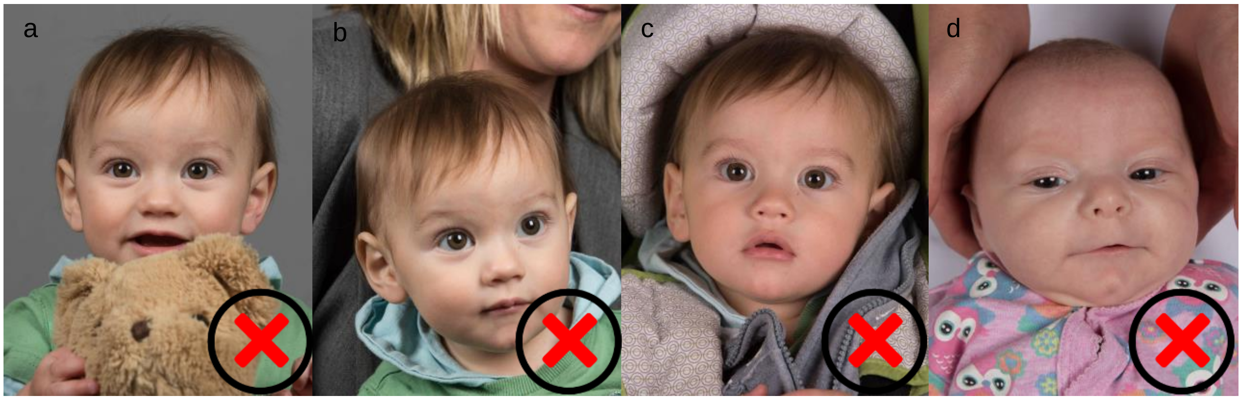


Figure 39 — a) Toy, b) supporting person ,c) Non-neutral background, d) Supporting hands.

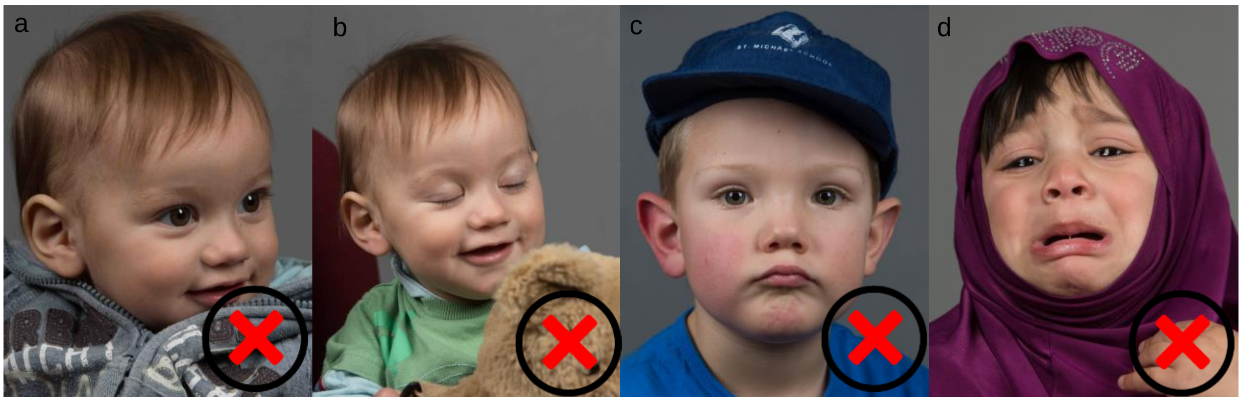


Figure 40 — a) Not looking into the camera, b) Eyes closed, c) Cap, d) No neutral pose.

6 IMAGE PRINTING FOR PORTRAIT SUBMISSION

6.1 Introduction

This Clause specifies the requirements for printed images submitted with document application. Many issuing states use a printing/re-scanning procedure for document application. This approach is feasible. However, caution should be taken to ensure quality according to this TR. If a new design of the application process is considered, digital submission should be taken into account as the preferred technology whenever possible. See Figure 41.

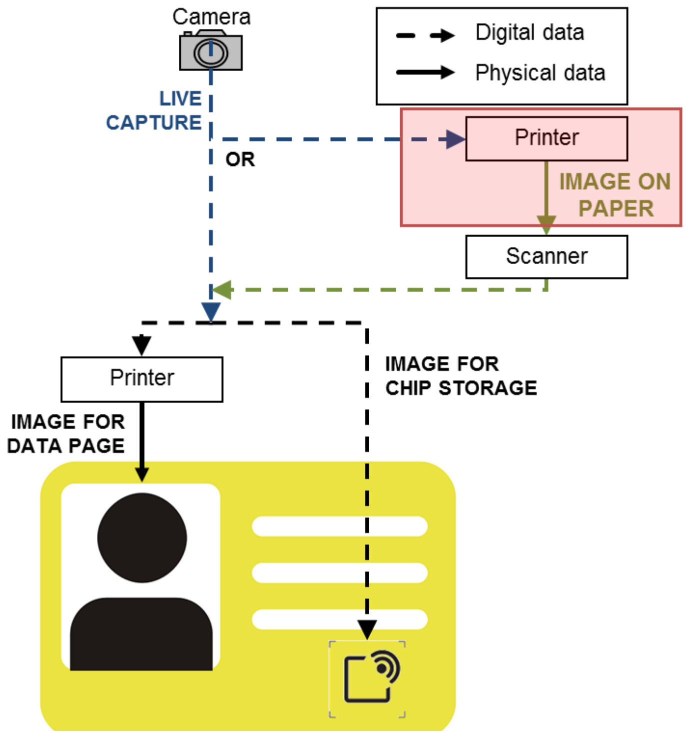


Figure 41 — Content of Clause 6 in the passport production process chain (boxed in red).

The physical portrait shall yield an accurate recognizable representation of the subject. The quality of the original captured image should at least be comparable to the minimum quality acceptable for paper photographs (resolution comparable to 6 – 8 line pairs per millimetre). To achieve this comparable image quality in a digital reproduction, careful attention shall be given to the image capture, processing, digitization, compression and printing technology and the process used to produce the portrait. The printing process shall maintain the width to height ratio of the original image.

6.2 Print resolution

The printing process should produce a smooth image that is capable of accurately rendering fine contrasted facial details, such as wrinkles and moles. All flesh tones from both light- and dark-complexioned subjects should be printed accurately and limited hot spots or shadow drop-outs apparent. Smooth facial details should be rendered without noticeable posterization or contouring.

6.3 Saturation and colour

With the exception of glare or glints caused by small areas of possible specular (mirror-like) reflection, only a small portion of the printed image should be saturated in white or black. Excluding the background area, using luminosity, the number of fully saturated 0 value pixels shall be less than 0.1%, and the number of fully saturated 255 value pixels shall be less than 0.1%.

No portions of the background or the subject's garments should be printed fully white and details should be apparent in dark shadow regions.

Printed photos shall be colour images having balanced colour channels. It may be assumed that the capture device (digital camera or scanner) is correctly white balanced.

6.4 Paper properties and portrait size

The photograph shall be on photo-quality paper. Examples of such paper are the following. Other technologies with similar properties are acceptable, too:

- Instant photographic standard gloss,
- Dye sublimation photographic semi gloss,
- Silver halide photographic semi gloss, or
- Drylab photographic inkjet bases standard gloss.

The photograph paper shall have a low roughness, non-structured surface (no pearl- or silkscreen effect). Submitted portraits should have a minimum width of 35 mm. The IED should be at least 10 mm.

Newly designed application processes still relying on printed portrait submission should consider using larger photo sizes, as, e.g., 7 cm by 10 cm. Larger photos reduce the risk of quality losses in the process chain. However, a switch to larger photos will have process implications to be considered.

6.5 Moiré or visible dot patterns

Digitization of printed photos may introduce artefacts, such as moiré, and certain printing processes may exacerbate the generation of such artefacts. The printing process employed should allow accurate face recognition when its prints are scanned with a document scanner at a spatial sampling rate of 120 pixels per centimeter (300 pixels per inch) in each axis.

If a printed photo has been produced through a periodic half-toning process, scanning the photo will almost invariably introduce moiré patterns. Thus, those printers, such as inkjet and laser printers, which inherently employ half-toning to simulate continuous tones, should use non-periodic (or dithered) half-toning methods. Furthermore, the printing process should not produce dot patterns visible to the unaided eye.

6.6 Photo template

It is often useful to provide a transparent template to a person responsible for photo quality evaluation. The template would display the limits of head size and rotation (roll) and, when superimposed on the photo, could assist in the determination of whether a printed photo is compliant to the requirements. Samples of such tools are given in Annex B.

7 SCANNING OF SUBMITTED PORTRAITS

7.1 Introduction

This Clause contains requirements to the scanning of a printed portrait to be submitted with a MRTD application. See Figure 42.

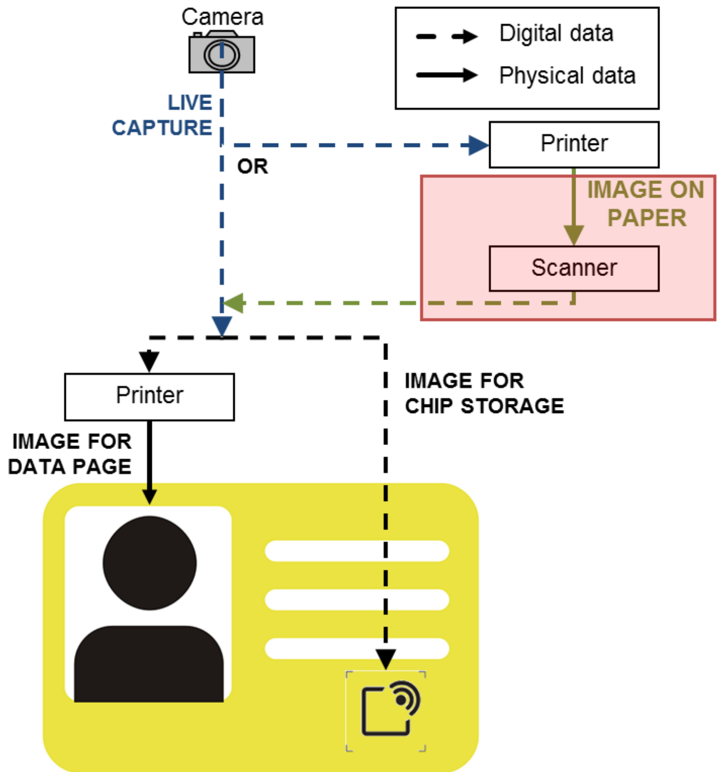


Figure 42 — Content of Clause 7 in the passport production process chain.

7.2 Properties of the submitted portrait

Submitted printed portraits shall comply with Clause 6.

Multiple scan/print steps are not allowed in an application process. If the portrait has been printed for submission and is subsequently scanned, all remaining production steps shall be digital.

A submitted printed portrait shall have been captured within the time frame defined in Clause 9.1.

If printed portraits are submitted, evidence on the capturing date should be requested. This may be the printed manufacturing date on the back side of the photo, or a dated invoice of the photographer. The complete card should be provided if the portrait is part of a photo card (e.g., a 10x15 print containing 2x2 images).

The submitted portrait shall be clean, not bent, not scratched, not folded and not damaged. There shall be no ink marks or creases on the printed portrait.

7.3 Pixel count and MTF

The finally scanned images shall have a pixel count as specified in Clause 9.2.

MTF20 should occur at 4,7 cy/mm or higher for scanners. The scanner’s MTF should be the same in both axes. Image enhancement processing using either built-in hardware or software-based image sharpening generally should not be used to boost the MTF.

EXAMPLE: The optical properties of the image can be maintained if the digital camera original image MTF20 should occur at approximately 80% or higher of the Nyquist frequency when using test patterns shown in Annex D. The size of a freckle/mole that should be detectable in face photos is 2 to 3 mm. Rulers make good fiducial markers to make measurements on the image.

The MTF analysis should be done using the appropriate target from ISO 12233. See Annex D.

EXAMPLE: A typical printed image with 10 mm IED should be scanned at a sampling rate of at least 300 ppi.

The MTF will be limited by the size of the paper photo and the resolution (fineness of detail) therein. To obtain higher resolution from scanned images the issuer should consider increasing the size requirement for printed portraits.

Particular care shall be taken in the acquisition process in order to avoid any kind of image dimensional stretching in any direction.

The width to height ratio of the final image is defined by the application process of the issuer, a typical value is 7:9. Necessary modifications shall be made by cropping and shall not be made by stretching.

7.4 Colour, sharpness, and saturation

The scanned portrait shall have the same colour as the submitted one. The human eye shall not be able to detect differences between the portrait and scanned result when viewed on a colour corrected display device and under daylight conditions. The portrait shall have appropriate brightness and contrast that show skin tones naturally.

The number of quantization levels should be at least 256 levels per colour, with three colours per pixel. The scanned image shall comply with the colour requirements in Clause 5.2.9.

Since red-green-blue (RGB) colour space and its derivatives are inherently device-dependent, the scanner’s output shall be converted to one of the well-defined, device-independent colour spaces as given in Clause 5.2.9.

Saturation occurs when significant numbers of pixels have values that are at the limits of quantization, i.e., at the levels of 0 or 255, if quantization of eight bits per colour is employed. Acceptable scanned face images should not have a significant number of pixels in saturation in the facial region.

The scanned portrait shall be centred, clear and in sharp focus with no shadows. It shall not have visible compression artefacts.

8 IMAGE PRINTING FOR MRTD PRODUCTION

8.1 Introduction

This Clause describes the requirements for images to be printed on the data page of a MRTD. The portrait printed on the data page shall be derived from the same digital image source as the image stored electronically in the MRTD. However, due to the influence of printing technologies as well as to the application of several security features to the portrait and to the data page, the image may not be exactly the same. Examples for possible deviations are the printer resolution, removed background in the printed portrait, image enhancements, dithering of grayscale content, or guilloches occurring in the print. See Figure 43.

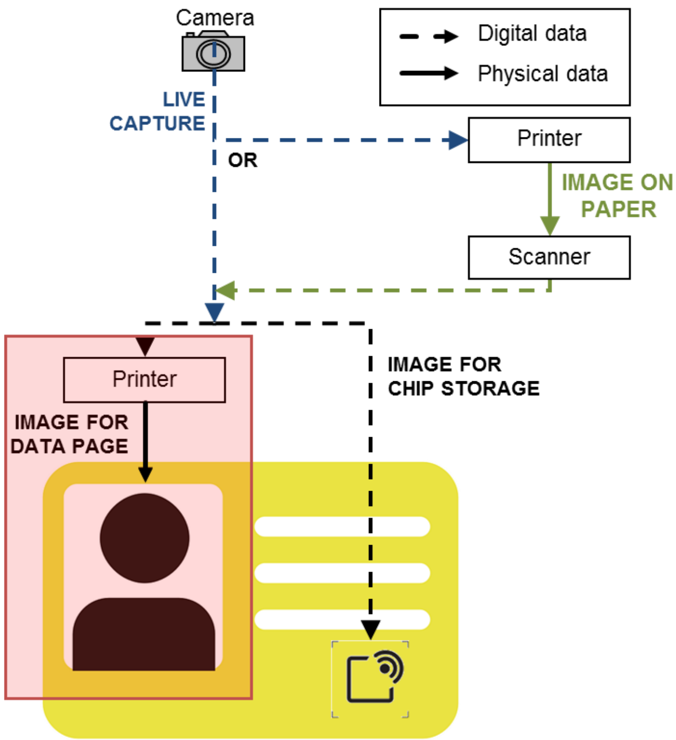


Figure 43 — Content of Clause 8 in the passport production process chain.

The implementation of the portrait on or into the MRTD should be done considering the properties of the different materials and technologies in use. It is possible that the printing technology itself introduces specific features into the printed portrait.

The digital reproduction shall yield an accurate recognizable representation of the subject. To achieve such image quality in a document data page, careful attention shall be given to the processing, compression and printing technology and the process used to produce the portrait. Printed portraits have specific features which depend on categories of printing technologies.

The primary printed image on the MRTD may be either greyscale or colour.

Any face printing process should produce a smooth image that is capable of accurately rendering fine facial details, such as contrasted wrinkles, contrasted moles, and contrasted scars, as small as two millimeters in diameter on the face positioned anywhere in the printed image area. Such details shall be detectable when viewed with the naked eye at a distance of 0.3 m.

All flesh tones from both light- and dark-complexioned subjects should be printed accurately and no hot spots or shadow drop-out should be apparent. Smooth facial details should be rendered without posterization or contouring.

8.2 Size

The portrait dimensions should be the same as specified in Clause 5.4. Necessary modifications shall be made by cropping and shall not be made by stretching. In cases where the background has been removed from the image, the correct width or height of the printed image may be impossible to determine. In such cases the height-to-width ratio is considered to be maintained if the ratio between *IED* and *EM* of the printed image is the same as of the portrait.

8.3 Tonal range

The tonal range of the printed image shall not interfere with facial details important for human identification when making a comparison of the printed image to the document holder.

8.4 Moiré or visible dot patterns

Moiré or dot patterns in the printed image should be minimized. Any such patterns in the printed image shall not interfere with facial details important for human identification when making a comparison of the printed image to the document holder.

8.5 Portrait placement in an MRTD and coexistence with security printing

The printed portrait shall be centred within Zone V as defined in ICAO Doc 9303.

If present, a digitally printed reproduction shall coexist with background security treatment(s) located within Zone V, i.e., the background security printing shall not interfere with proper viewing of the displayed portrait, and vice versa, yet still offer protection to the displayed portrait.

9 IMAGE STORAGE IN THE CHIP

9.1 Introduction

This Clause specifies properties of the portrait to be electronically stored in a MRTD. Figure 44 shows the content of Clause 9 in the MRTD production process chain.

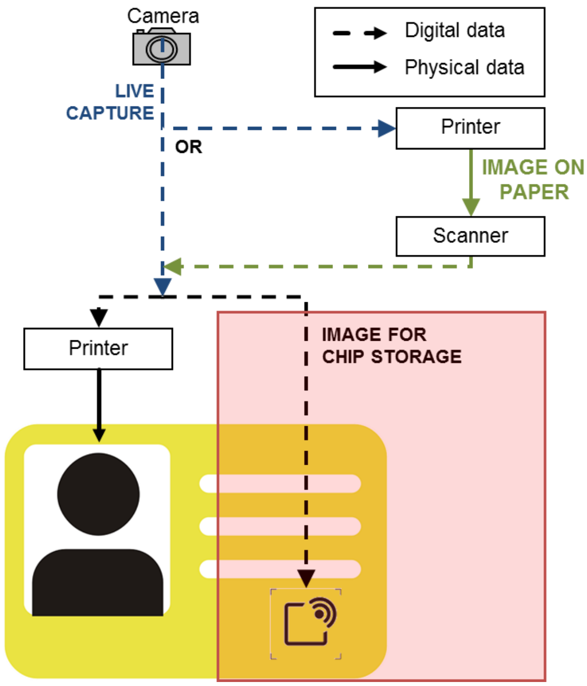


Figure 44 — Content of Clause 9 in the MRTD production process chain (boxed in red).

The requirements and recommendations given in this Clause shall ensure that the photographic requirements given in Clause 5.2 are retained in the portrait that is finally stored in a MRTD, with the exception of the pixel count. The minimal requirements specified in this Clause apply to the image finally stored in DG2.

A submitted portrait shall have been captured within the last six months before application. Portraits with a capture time dating back more than three months should not be accepted. Issuers should consider the use of the metadata encoded with the digital image, to assure that the photograph is recent. See Table 10.

Table 10 – Capture time requirements and recommendations.

Criterion: Capture time	Requirement	At most six months before application.
	Best Practice	At most three months before application.

9.2 Data format

Portraits of a subject to be stored in the MRTD chip

- Shall be stored in one of the following formats: JPEG, JPEG2000,
- Should have a minimum IED of 90 pixels, preferably of 120 pixels,
- Shall be in colour.

These specifications provide adequate spatial sampling rate for use on the MRTD while maintaining an adequate quality for human and machine facial recognition purposes.

Table 11 – *IED* requirements and recommendations for the chip image.

Criterion: IED	Requirement	*IED* ≥ 90 pixels
	Best Practice	*IED* ≥ 120 pixels

NOTE: The pixel count specified in Clause 5.2.4 applies to the originally captured portrait and not to the imaged to be stored in a passport. The processing steps between capturing and passport production might lead to information losses. It is therefore recommended that a higher resolution version of the image be stored in the issuer's repository.

One of the three possible encodings shall be used:

- The JPEG Sequential Baseline (ISO/IEC 10918-1) mode of operation and encoded in the JFIF file format (the JPEG file format).
- The JPEG-2000 Part-1 Code Stream Format (ISO/IEC 15444-1), lossy, and encoded in the JP2 file format (the JPEG2000 file format).
- The JPEG-2000 Part-1 Code Stream Format (ISO/IEC 15444-1), lossless, and encoded in the JP2 file format (the JPEG2000 file format).

The coordinate origin shall be at the upper left given by coordinate (0,0) with positive entries from left to right (first dimension) and top to bottom (second dimension).

9.3 Property Mask

The positions of the Property Mask in the header structure of ISO/IEC 19794-5 should be set for:

- (Medical) dark glasses
- Head coverings
- Left and right eye patches
- Glasses
- Feature distorting medical conditions which could impact feature point detection
- Subject height

9.4 Post-acquisition processing

No other post processing than

- In-plane rotation and/or
- Cropping and/or
- Down sampling and/or
- White balance adjustment and/or
- ICC colour management transformation and/or
- Processing RAW images into the target encoding (once) and/or
- Compression as described in Clause 9.5.

shall be applied on the captured image to create the portrait. Any processing shall maintain the requirements given in Clause 4 and Clause 5. Any processing shall render skin and hair colours realistically enough to allow straightforward human identification of the MRTD holder. The face images shall not be modified locally, e.g., for removal of scars, pimples or other skin impurities or to modify the shape or location of the nose, the eyes, the eyebrows or any other facial landmarks. The image shall not be modified locally by editing clothes (e.g., a turban).

In particular, any image processing targeting at background removal shall not be implemented. If necessary, the MRTD issuing authority may remove or alter the image background later in the MRTD production process. See Clause 8.

9.5 Compression

Captured portraits of a subject should not sacrifice image quality by overly compressing the image.

For maximum effect in human and automated facial recognition, the raw image or an image with limited compression should be retained. The JPEG compression ratio shall not exceed 15:1.

EXAMPLE: In many cases, such images have a size of at least 12 kBytes for JPEG and JPEG2000 for storage in the chip of an electronic MRTD. The upper limit is defined by the available storage space available on the chip and reading time requirements.

The image used for the printing process on the MRTD and for storage in the chip will give better results if not compressed beyond a ratio creating visible artefacts on the image when viewed at 100% magnification – where a single pixel in an image file is displayed by a single pixel on a monitor or viewing device. This allows for electronic judging of whether an image is overly compressed.

Lossy compressions can only be applied once per each of the following steps:

- one initial compression by the camera itself,
- one compression done by the photographer or citizen, and
- one compression done by the issuer.

JPEG 2000 enables compressing to a target file size. If using JPEG compression must be done iteratively to the target file size while reverting back to the original image rather than successive compressions.

Annex A

(informative)

Setup examples for portrait capturing

A.1 Introduction

The implementation of the portrait acquisition setup should be done considering the properties of the different technologies and environments. Some portrait acquisition setups are detailed below. This Annex lists examples in no particular order and does neither recommend technologies mentioned here nor exclude technologies not mentioned here.

A.2 Studio environment with one single light

A single light and multiple reflector panels illuminate the subject's face uniformly. The light with a reflector should be placed approximately 35° above the line between the camera and the subject and be directed toward the subject's face at a horizontal angle of less than 45° from the line. A reflector panel should be placed on the subject's opposite side to prevent shadows on the face. Optionally, an additional reflector may be placed below and in front of the subject's face to illuminate the area around the chin. See Figure A.1.

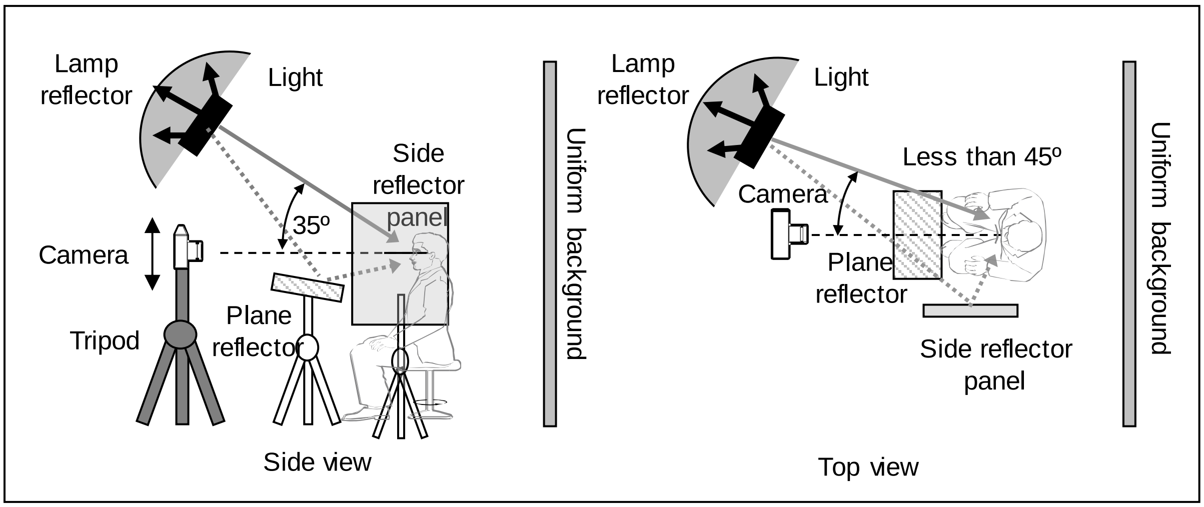


Figure A.1 — Example of a setup with one single light

A.3 Studio environment with two lights

Two lights with reflectors should be placed approximately 35° above the line between the camera lens and the subject. Both lights should be placed within 45° of the line between the camera lens and the subject. The optional plane reflector in front of the subject supplies additional light around and below the subject's chin. See Figure A.2.

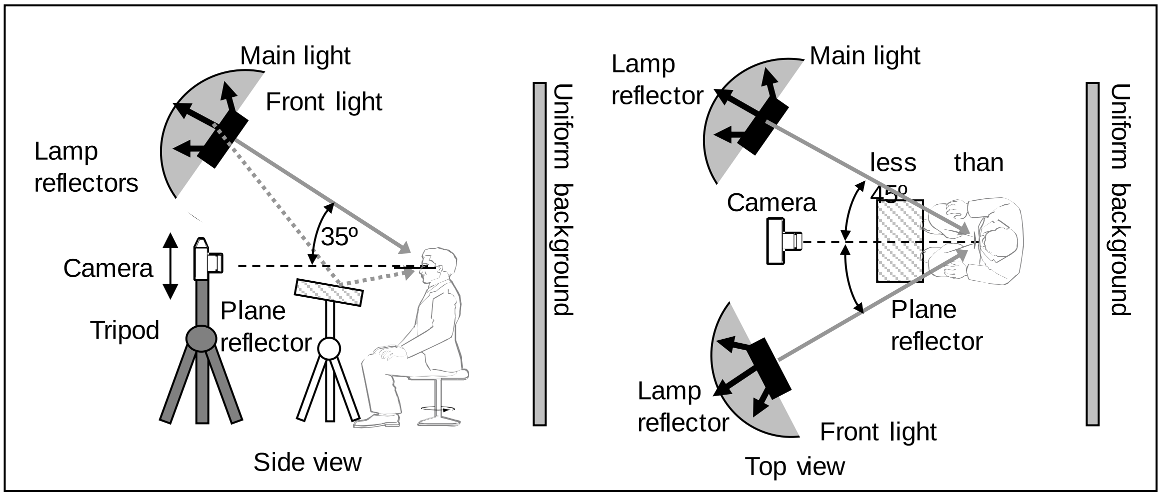


Figure A.2 — Example of a setup with two front lights

A.4 Studio environment with two lights and background illumination

A background light is added to the arrangement in Clause A.3 to eliminate shadows visible on the background behind the head. The background light should be targeted to the background and be placed directly behind and below the subject. See Figure A.3.

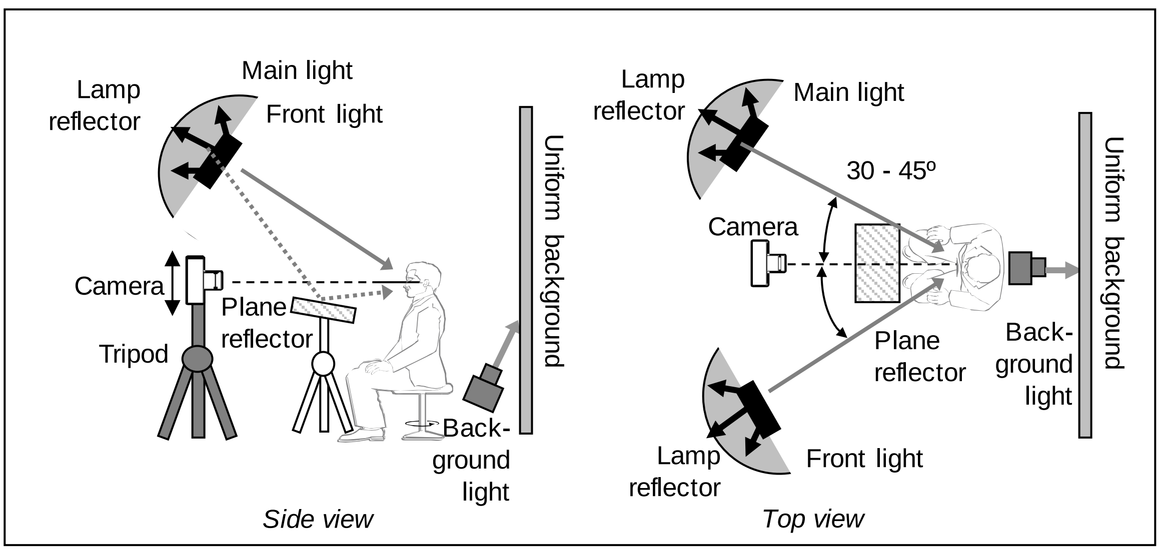


Figure A.3 — Example of a setup with two front lights and one background light

A.5 Photo booth environment

The requirements for a good portrait also apply for a photo booth and an operator should make their best effort to get as close as possible to the recommendations given in this document. In studio environments, the human element is capable of checking the quality assurance; this capability should be replaced by automated quality assurance technology in photo booths and kiosks.

In a photo booth, multiple lights should be positioned symmetrically behind a diffuser panel above and aside of the camera to provide uniform lighting on the subject's face and to eliminate glare and shadows visible on the face. See Figure A.4 and A.7. A background light should be placed on the ground between the background and the subject. The front lights should be placed at an angle of approximately 35° above the line between the camera and the subject's head to prevent reflection artefacts on the subject's glasses. The inside walls should be white and serve as reflectors. Directly behind the subject should be no directly reflecting material. The interior lights of the booth should be kept switched on during operation to reduce red eye effects. Direct or indirect lighting from below and in front of the subject should be used to eliminate shadows around the chin. An opaque curtain should be used and be closed during capturing to eliminate external light effects.

Proper positioning of the subject and control of the subject's pose may be improved through feedback provided to the subject via a mirror or a live-video monitor. A height-adjustable seat or camera should be provided to allow the subject to face the camera. See Figure A.5. Alternatively, the camera may be movable to adjust the height to the head position. See Figure A.6.

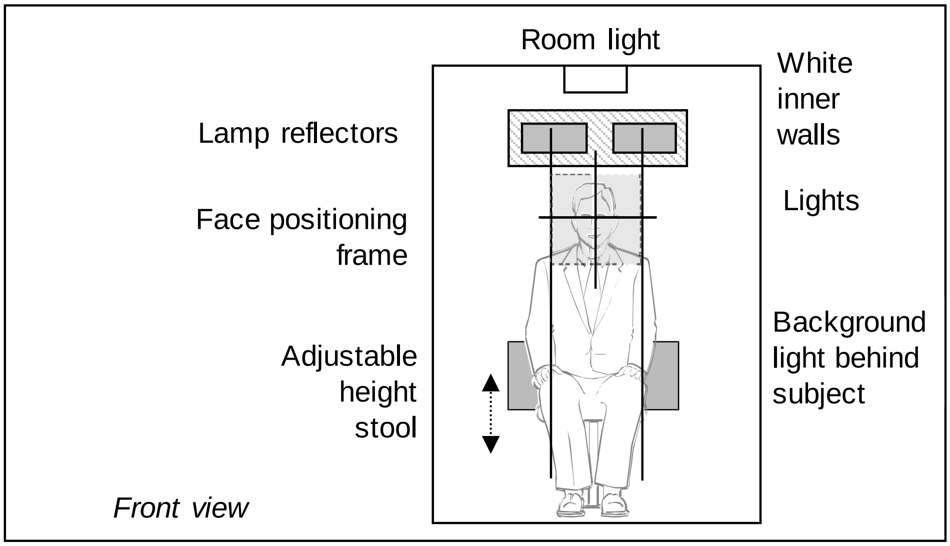


Figure A.4 — Example of a photo booth setup: Front view.

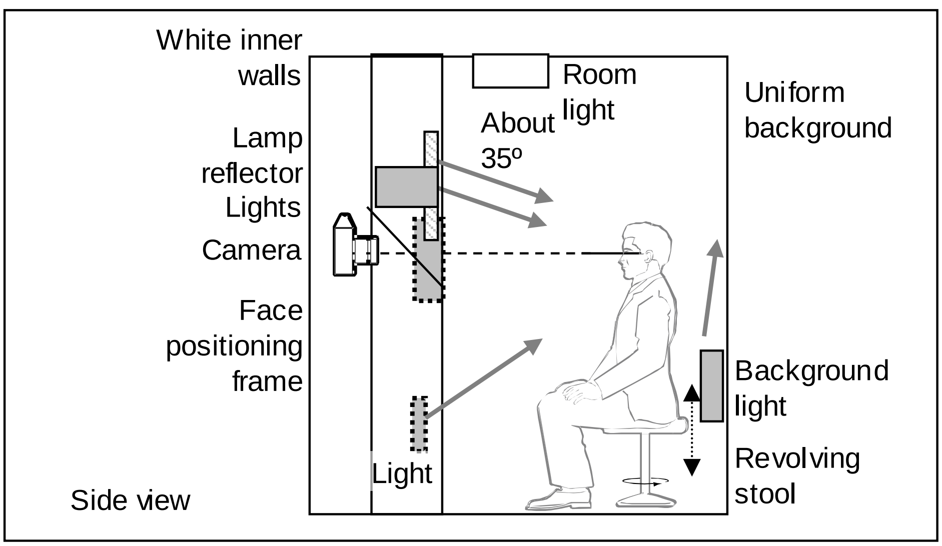


Figure A.5 — Example of a photo booth setup: Side view.

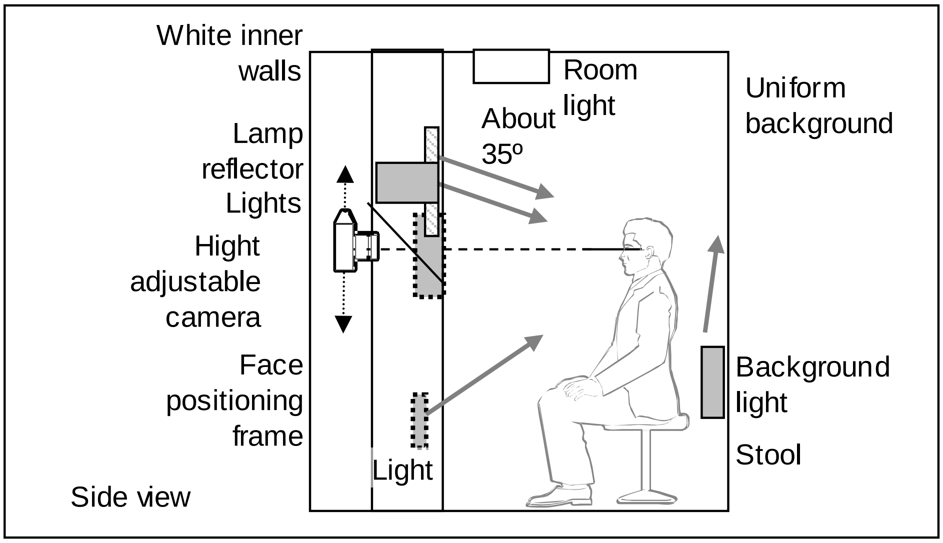


Figure A.6 — Example of a photo booth setup: Side view with height adjustable camera.

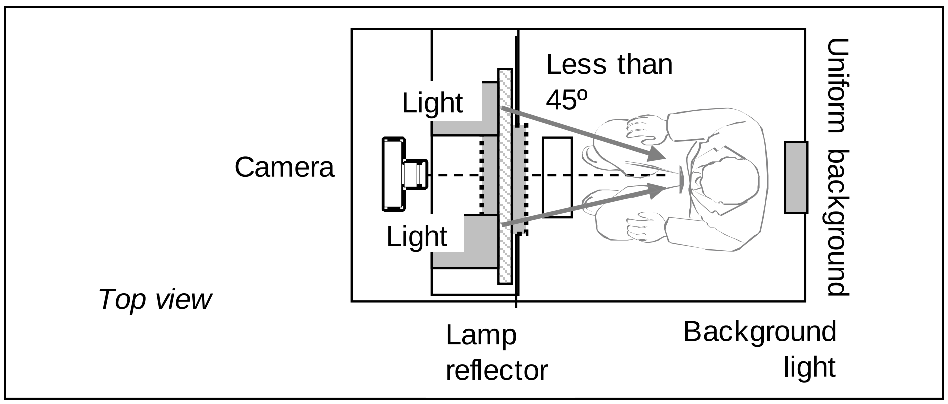


Figure A.7 — Example of a photo booth setup: Top view.

A.6 Registration office environment

The requirements for a good portrait also apply in a registration office environment and an operator should try to get as close as possible to the recommendations given in this document. One should have in mind that such an easier setup regularly leads to portraits of suboptimal quality and should therefore not be the preferred solution.

The subject and the background should be illuminated by two diffuse light sources that are mounted in a console with a small footprint, so that it fits into a typical registration office environment. The console may be mounted on the floor or on the wall. Flash light should not be used, at most in combination with appropriate permanent illumination. The main illumination during capturing should be that of the capturing system. Illumination mainly by the room illumination from the roof lights, the window or desk light is not acceptable, as well as direct sun light. Even if the office conditions might require much easier setups, the principle requirement of a uniform illumination remains valid.

A revolving and height-adjustable chair or stool with an additional cushion for smaller persons should be provided to allow the subject to face the camera and adjust his head to the proper height. See Figure A.8.

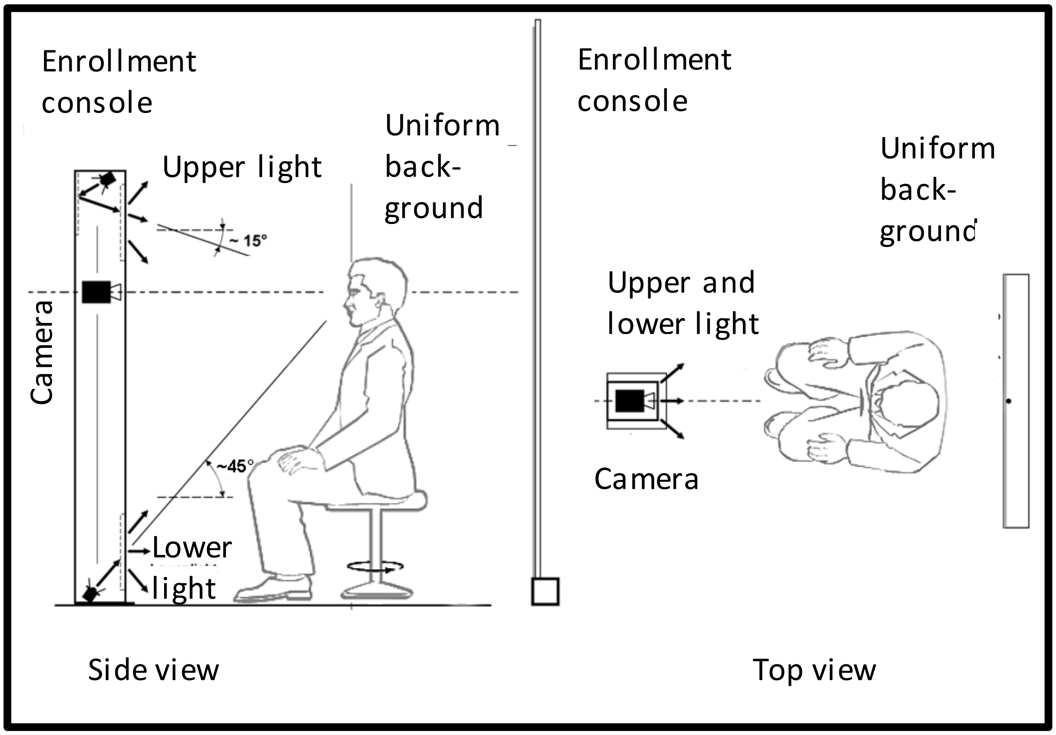


Figure A.8 — Example of a registration office environment setup

Feedback should be provided to the subject via a second live-video monitor facing to the subject for positioning and behaviour guidance. An image preview should be offered to allow a subject to choose from a selection of portraits or to be recaptured if necessary before the final portrait is submitted for further processing.

However, empirical data from production environments indicates that subjects who see the live view enter into a “vanity mode”. This can significantly reduce the throughput of the process and the quality of the captured biometric data. As an alternative to live view, visual, graphical, or verbal instructions should be provided to the subject to reach optimal face and body posture.

A.7 Setup with flash

If carefully applied, flashes may be used. In this case the quality requirements especially with respect to shadows, homogenous illumination, and absence of reflections in the eyes have to be maintained. See Figure A.9. The given distance measures are examples.

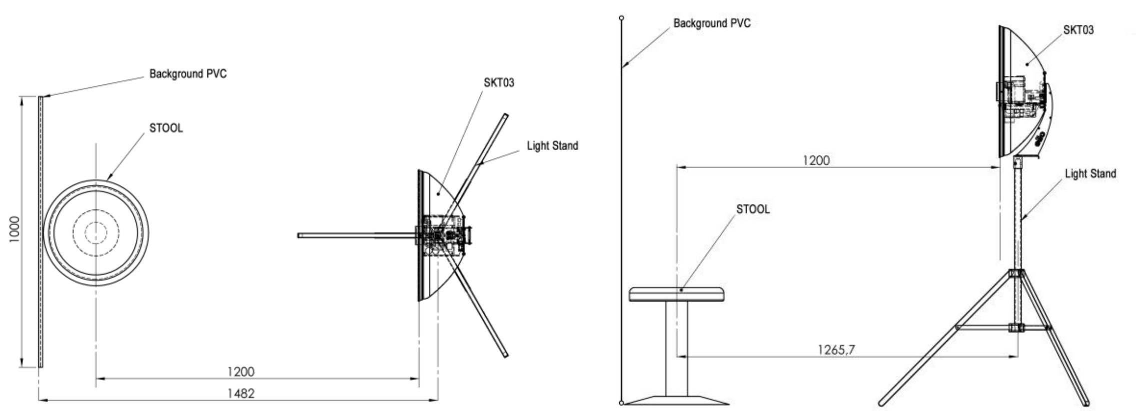


Figure A.9 — Example of a setup with flash, top and side view. The given measures are examples.

Annex B

(informative)

Measuring magnification and radial distortion in a portrait capture setup

B.1 General

This Annex deals with two kinds of distortion.

The first kind is the magnification distortion, a geometrical effect of the optical perspective. Some optical systems image objects of the same size differently depending on the distance between object and sensor. Such magnification distortion always appears in human vision.

The second kind is the radial distortion (barrel distortion, pincushion distortion, moustache distortion) caused by optical properties of a lens. This feature doesn't exist for human vision.

Both magnification and radial distortion may influence the performance of automated face recognition systems as well as of human recognition. Therefore, this Annex describes how to measure various types of distortion using two targets.

B.2 Magnification distortion target construction

To build the magnification distortion target, take Figure B.1, print it on A4 size paper, and fold it into a T shape. One should look at the foot of the T, the head bar is on the opposite side away from the viewer. The length of the T leg is the typical eye to nose distance of a Caucasian (50 mm). In a photograph of the magnification distortion target taken as described in Clause B.4, the relative size of an object near to the eye level and near to the nose tip level of a human can be measured. Additionally, it can be observed if the image is sharp enough at the entire facial region including nose and eyes looking at the visibility of the 0,5 mm and 1 mm wide markers in the squares. See Figure B.2.

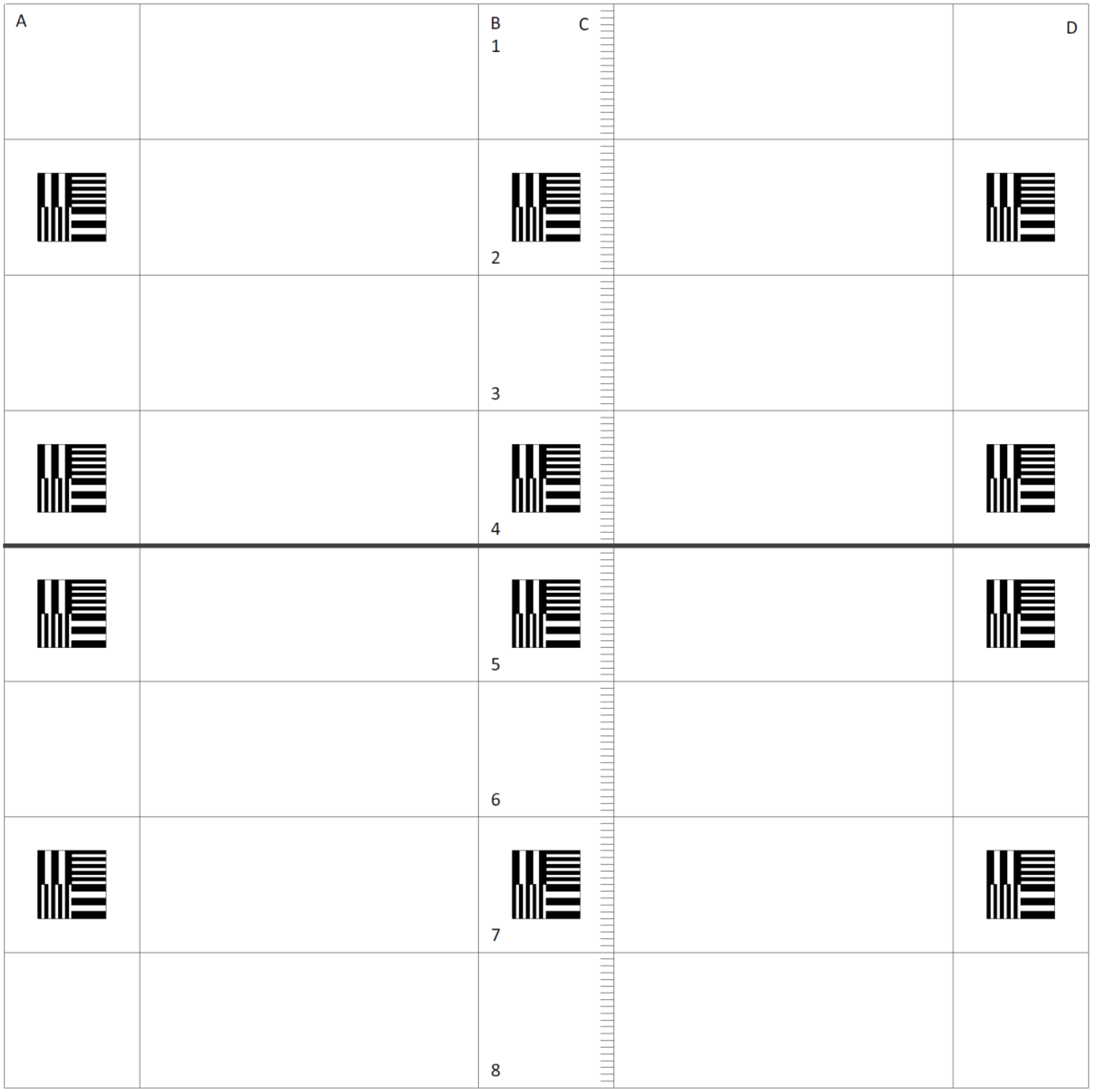


Figure B.1 — Magnification distortion target.

Figure B.2 — Ready to use simple magnification distortion target. One can see the millimetre markers on the nose level, and the markers in 2 cm distance on nose and eye level as well as the resolution targets with one and two line pairs per millimetre.

To precisely and repeatedly measure magnification distortion, fabrication of a rigid magnification distortion target is recommended.

To keep the magnification distortion target pointing at the camera it might be necessary to attach the folded target to a support. For target support material it is recommended to use white foamboard with a thickness of 5 mm or 2 mm aluminium. The target support structure construction is also suitable for 3D printing using plastic materials in light colour. Paper based materials should not be used.

This magnification distortion target should be used to measure magnification distortion in the nose region of the face. The magnification distortion measuring is described in Clause B.4.

B.3 Radial distortion target construction

A radial distortion target is composed of evenly-spaced, horizontal and vertical lines, forming a net-like pattern.

Start building the radial distortion target by cutting the back plane and supporting panels.

This can be done by stacking four foamboard pieces together to form a 50 mm × 160 mm × 20 mm support slab. See Figure B.3. If thinner aluminium or thicker foamboard is used, then change the dimensions accordingly. Paper based materials should not be used for the support. Foamboard or similar material is recommended for the box material. The size of the visible white board is 200 × 200 mm. To maintain the base material flatness requirements use 2 mm aluminium or 5 mm foamboard.

Glue the four foamboard pieces together. Use strong glue that does not melt board material. Make sure that the slab has 90 degree corners. In order to keep the slab in correct shape use supports while gluing the slab together piece by piece.

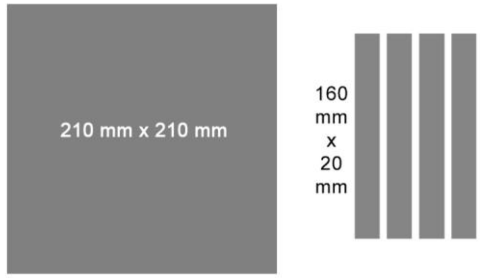


Figure B.3 - Magnification distortion target support parts and dimensions for 5 mm thick foamboard construction.

The printed radial distortion target is glued on the back pane foamboard. Figure B.4 contains the target support board back plane radial distortion target print version for A4 size printing.

Print the target on an A4 paper. Check the size of the grid before cutting. The grid size is 150 mm x 150 mm (exact size is 151,1 mm x 151,1 mm due to the line width in use). Cut out the radial distortion target along the outmost border line so that the border line stays intact or cut out the upper and lower part of the A4 to form a 200 mm x 200 mm size paper target. Glue the radial distortion target at the center of the 210 mm x 210 mm foamboard.

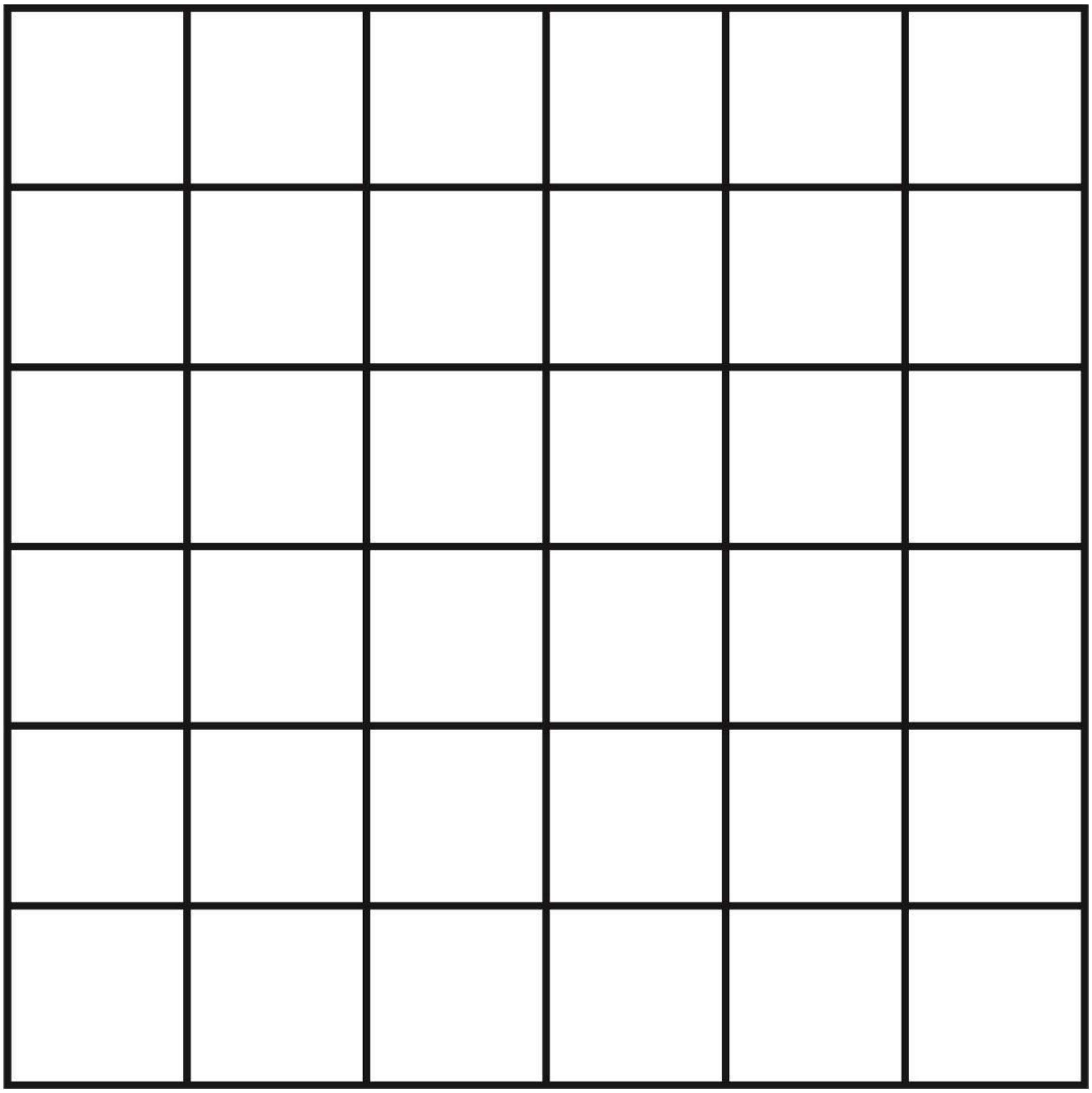


Figure B.4 - The radial distortion target version for A4 size printing.

The radial distortion target should be used to measure barrel or pincushion distortion. The radial distortion measuring is described in Clause B.6.

B.4 Target photography

Place the target at the typical location of the subject's head in the given photographical setup or glue it in a target support board setup. Take a picture of the target in the intended setup, from the chosen distance, with the intended focal length, with the intended aperture.

Target is placed at an appropriate distance from the lens following the guidelines set in this document. Before taking photographs camera and lights are set following the recommendations of this document.

It is easy to use a tripod or similar support to keep the target properly aligned. Touch (hook and loop) fasteners may be used to temporarily hold the target for photography. A small patch of fastener tape attached to the target support is able to hold light weight targets described in this document.

B.5 Magnification distortion measuring

If the distance between the foremost point of a subject and the optical center of a standard lens (i.e not telecentric) is D and the height of a structure S_1 in the front of the face, e.g., the nose is h_{s1}, then the camera to subject distance (CSD) is assumed to be $\Delta D+D$ and a structure S_2 of the height $\Delta h_{s1}+h_{s1}$ at eye level, would virtually appear to have the same size as S_1. S_1 virtually seems to be larger than it is in reality due to the magnification distortion. See Figure B.5.

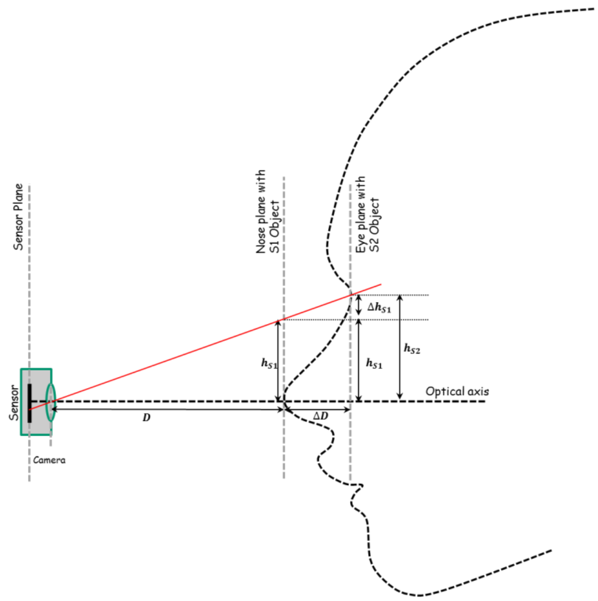


Figure B.5 – Illustration of the magnification distortion effect.

The above figure holds: $(\Delta h_{s1}+h_{s1})/h_{s1} = (\Delta D+D)/D$. This leads to the definition of the magnification distortion factor:

$$K_{magnification} = \Delta D/(D+\Delta D) \times 100\% = \Delta h_{s1}/(h_{s1}+\Delta h_{s1}) \times 100\%$$

where ΔD is the depth of the measured object,

$D+\Delta D$ is the camera-subject distance,

h_{s1} is the height of a structure S_1 in front of the face.

The magnification distortion factor $K_{magnification}$ is the relative enlargement of an object at nose level compared to an object at eye level. The value of ΔD is assumed to be 50 mm as the typical distance between the nose tip and the eye level of an adult Caucasian. Table B.1 shows the (computed) absolute enlargement of an object with the size $\Delta D = 50\ mm$ (like a nose) seen from several distances and the corresponding relative enlargements $K_{magnification}$.

Table B.1 – Illustration of the magnification distortion effect.

CSD in mm	Δh_{s1} **in mm at** h **= 50 mm**	**magnification distortion** $K_{magnification}$
700	3,57	7,14%
1000	2,50	5,00%
1200	2,08	4,17%
1500	1,67	3,33%
2500	1,00	2,00%
3000	0,83	1,67%

Figure B.6 shows details of the photograph of the magnification distortion target taken in the intended photographic setup, from the chosen distance, with the intended focal length, and with the intended aperture. In order to determine the magnification distortion, measure a size $h_{s1} + \Delta h_{s1}$ at nose-tip level and the difference to the corresponding size at eye level Δh_{s1}, and calculate $K_{magnification} = \Delta h_{s1}/(h_{s1}+\Delta h_{s1}) \times 100\%$.

Using the magnification distortion target from Figure B.2, the following experimental data have been collected. Figure B.6 shows details from the captured image, Table B.2 shows the measured and computed results. Note that the observations almost exactly match the computations from Table B.1.

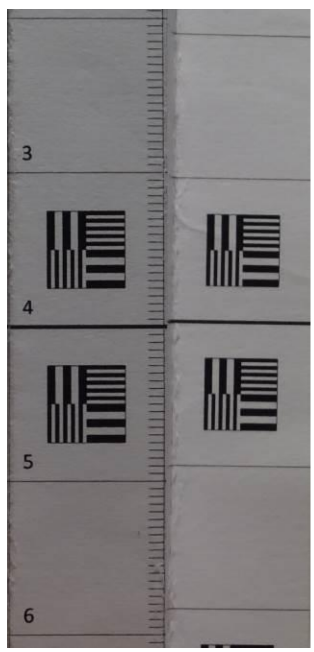


Figure B.6 – Example measurement of the magnification distortion effect.

Table B.2 – Experimental measures.

CSD	Target size $h_{s1}+\Delta h_{s1}$	h_{s1}	Δh_{s1}	$K_{magnification} = \Delta h_{s1}/(h_{s1}+\Delta h_{s1}) \times 100\%$
750 mm	80 mm (3216 pixels)	74,7 mm (3004 pixels)	5,3 mm	6,6%
1050 mm	80 mm (2631 pixels)	76,3 mm (2508 pixels)	3,7 mm	4,6%
1250 mm	80 mm (2160 pixels)	76,8 mm (2073 pixels)	3,2 mm	4,0%
1550 mm	80 mm (1742 pixels)	77,4 mm (1686 pixels)	2,6 mm	3,2%

B.6 Radial distortion measuring

Radial distortion of the camera lens causes information about the object to be misplaced but not lost. Camera lens measurement related radial distortion measurement is important as some of the ABC and kiosk cameras are not using highest quality lenses. For focal lengths shorter than about 30 mm, the distortion seems dominantly barrel, fish eye lens being the ultimate example. Cell phone cameras and fixed focus lenses have short focal lengths.

In practice measurements of the radial distortions are done either manually or using measurement software. If doing the measurements manually then it is important to zoom in the measured area and to use an image processing software ruler tool for measurements. See Figure B.7.

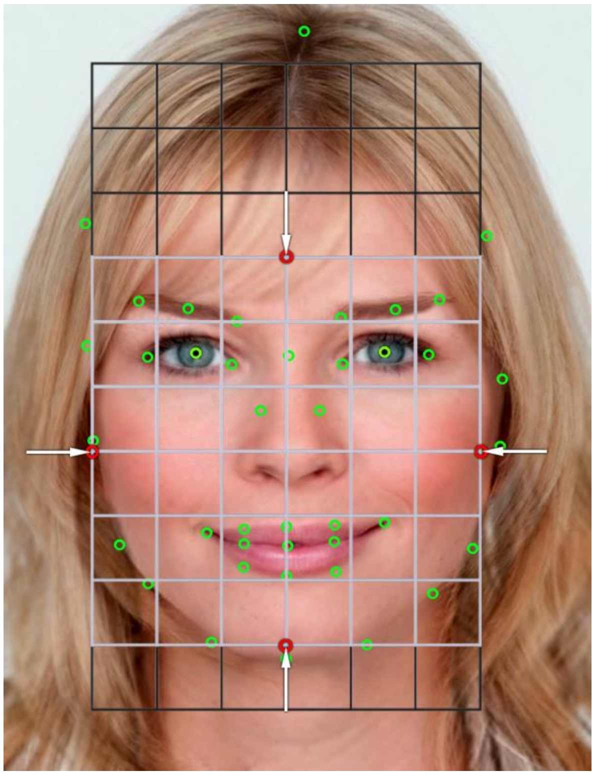


Figure B.7 – The printed grid is marked above in light blue to show the recommended alignment and size when compared to a portrait. Measurement vector starts from the grid corner point pixel location and ends at the respective corner point of the grid. The green circles are for reference purposes only to show how the recognition program may set certain landmark points on the actual portrait. Guidance for the measurements is shown in Figure B.8.

Radial distortion is a type of geometrical aberration that causes a difference in magnification of the object at different points in the image. Various points are misplaced relative to the central point of the image. The barrel or pincushion distortion K_{radial} is calculated by:

$$K_{radial} = (PD - AD) / PD \times 100,$$

where *AD* is the actual distance and *PD* is the photographic distance from the centre point of the image. Figure B.8 gives an example of barrel distortion non-distorted and distorted frame intersection points shown with dots and arrows. Use guidelines to find the corner points of the undistorted image shown here as a grid drawn through the marked points with black lines.

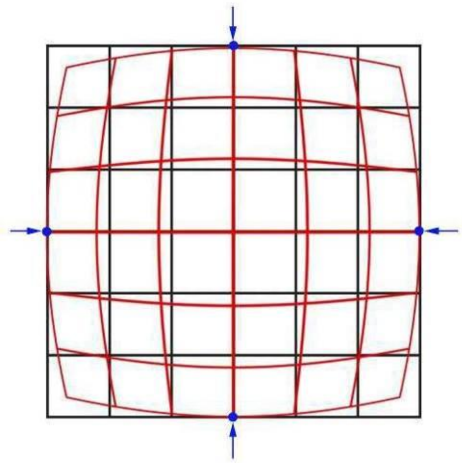


Figure B.8 - Barrel distortion non-distorted and distorted frame intersection points shown with dots and arrows. Use guidelines to find the corner points of the undistorted image shown here as a grid drawn through the marked points with black lines.

Distortion, represented by a percentage, may be either positive or negative. A positive percentage represents pincushion distortion, whereas a negative percentage represents barrel distortion. Figure B.9 illustrates barrel and pincushion distortion compared to an ideal grid picture in a perfectly square non-distorted image.

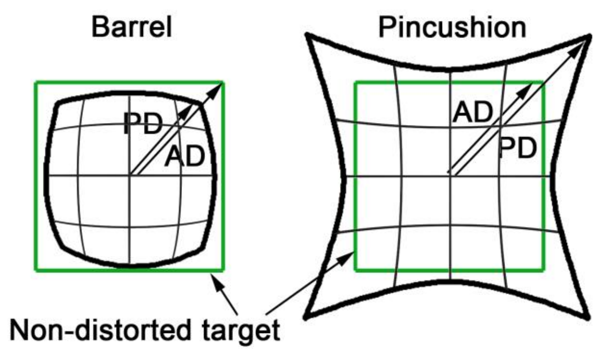


Figure B.9 - Barrel and pincushion distortion

In practice it is difficult to define the exact location of the non-distorted target location on the distorted photographic image. By drawing a cross through the middle point of the target image it is possible to find four intersecting points, which in turn may be used to locate the non-distorted target frame for error calculation purposes. This makes the measurements of different systems compatible but the values are not absolute as the real distortion zooming error behaviour of different lenses are not the same.

Annex C

(informative)

Colour test

C.1 Colour tests according to ISO 11664-4:2008(E)/CIE S 014-4/E:2007

The IEC 61966-8 [1], ANSI IT8.7/2 [2] or similar ISO 12641-1:2016 [3] test chart or other compatible and well documented colour test chart (e.g., IEC 61966-8 or ANSI IT8.7/2 of Figure C.1 and C.2) containing skin colour patches should be used to measure colour accuracy, and dynamic range (indirect method). Colour quality is determined by converting the image to CIE *L*a*b** [5] space and measuring the distance between the colours captured and known colour values of the chart patches.

Figure C.1 — IEC 61966-8 colour test chart.

Patterns with sufficiently large patches are needed to measure noise (e.g., ISO-15739, ISO-14524). See Figure C.2.

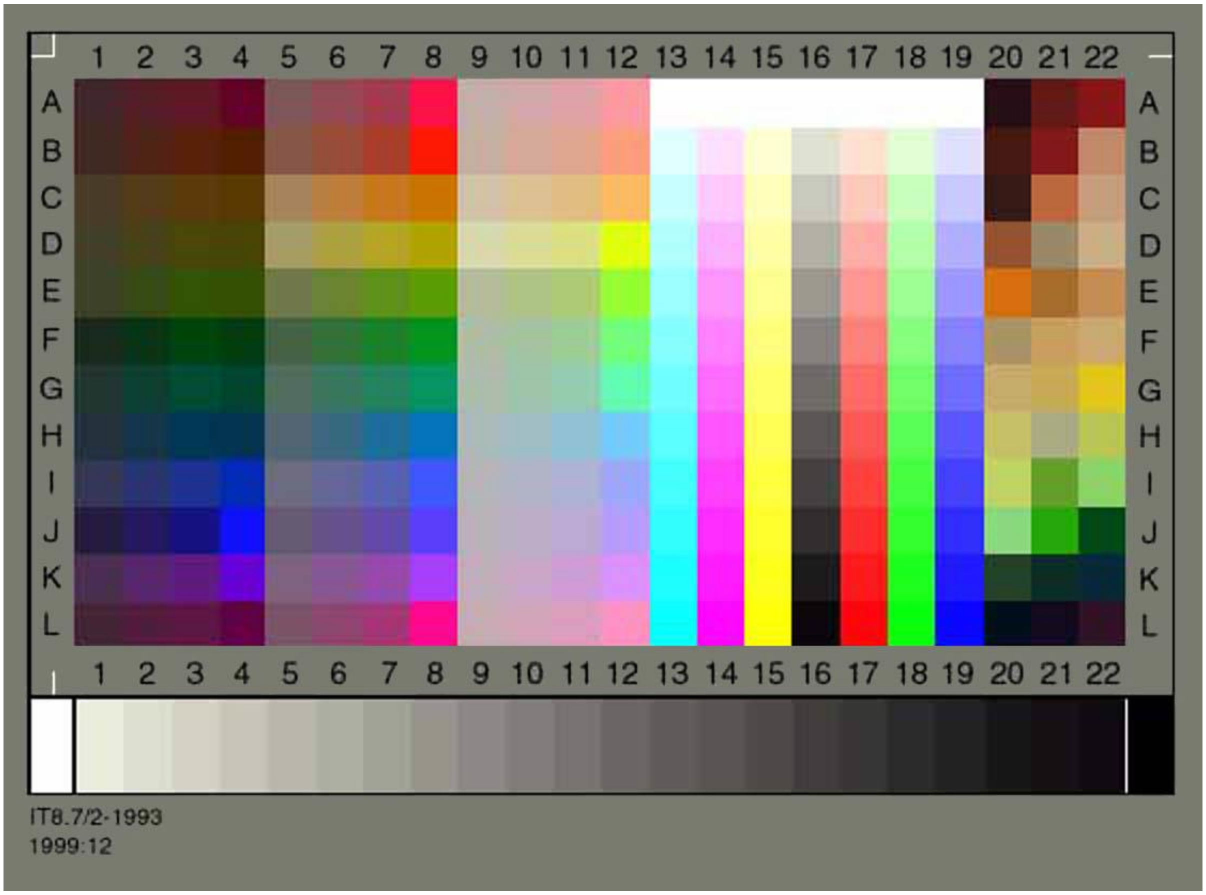
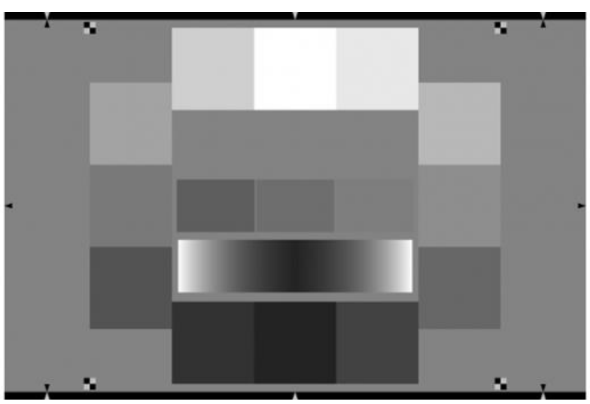


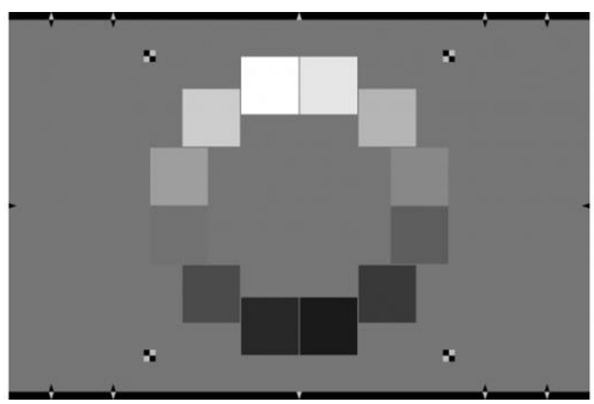
Figure C.2 — ANSI IT8.7/2 test chart.

C.2 Image colour quality

Human examiners and face recognition systems rely on high-quality skin tone presentation. Colour is a subjective psychological phenomenon, and human perception of colour depends on the context in which a perceived object is presented (i.e., chromatic adaptation). Therefore, the colour test should measure the entire gamut, as an examiner needs the surrounding colours to perceive colours in portions of the face (e.g., lips, hair, eyes, makeup) correctly.



ISO-15739
305 mm × 457 mm



ISO-14524
483 mm × 304 mm

Figure C.3 — Standard test patterns with sufficiently large swatches to measure signal to noise.

Cameras should be white balanced and scanners colour managed to ensure high fidelity colour reproduction across the entire gamut. This is necessary as the digital camera software is preprocessing the internal raw image format and may distort the image colour when JPEG is used as source format for image analysis.

In order to achieve good sample fidelity, there shall be no saturation (e.g., over- or underexposure) on the measurement target. All RGB channels of the image should have at least 7 bits of intensity variation (i.e., span a range of at least 128 unique values) in the test target patch region of the image. This is required to get as near as possible to a L^* level of 50, which in turn ensures a wide sRGB [6] gamut will be available for the analysis.

C.3 Measurements and analysis

In order to assure the required image quality system, installers shall make quality assurance measurements of light conditions and camera system response when a recommended CIE Standard Illuminant D65 high quality illuminant or similar continuous spectrum daylight illuminant and camera including camera control software are used to take pictures. In practice, it is also required to reduce the ambient light pollution emanating from uncontrolled daylight sources, fluorescent or similar light sources and reflections from surfaces.

Portrait images are typically stored in sRGB IEC 61966-2.1 [6], a device-independent colour space designed for consistent display across a wide range of commercial devices. However equal distances in spaces like sRGB do not represent equally perceptible differences between colour stimuli. To address this, in 1976 CIE created the LAB colour space whose coordinate system is based on non-linear transformations which attempt to capture perceptual distance. When computing colour error, images are typically converted from the sRGB colour space to L*a*b* [5], a colour space engineered to approximate the way the human visual system perceives colour. Algorithmically, the transformation from sRGB to L*a*b* values is accomplished by taking advantage of colour-matching functions developed for the CIE 1931 standard colourimetric system.

Fixed registration office imaging systems should be calibrated using manual or automated methods described in this document. For photo booth and mobile registration office imaging automatic white balance setting procedures and automatic quality analysis should be used. In a mobile environment, the use of advanced manual measurements may fail due to time and user training constraints. However, portraits shall not be captured without adequate colour balance.

Variations in human skin color have been measured using visible reflectance spectroscopy and the device-independent colour space (CIELAB) [15, 16]. Skin colour values can be expressed along the three dimensions of the CIELAB colour space: Lightness scaled from 0 (black) to 100 (white) along L*, and the opponent color axes a* and b* representing from Red through Green from positive to negative values along a* and similarly Yellow through Blue along b*. NOTE When the face image is expressed in sRGB colour space, then the gamut shrinks and moves upwards (in a*b* positive direction), and therefore higher L* values may produce higher a* and b* values than shown in research papers.

In a study [15] that used the above technique to measure the variation in the skin colour on the cheeks and foreheads of 960 Caucasian, Chinese, Kurdish, and Thai individuals, the expected mean and standard deviations of the variations found was as follows: L* mean 58.21 sigma 4.23; a* mean 11.45 sigma; 2.38 and b* mean 15.91 and sigma 2. However, as the study population was not representative of the variation in the skin across all ethnicities, wider variations in skin colour should be expected. As studies [15,16] have shown that values lower than 5 for a* and 10 for b* have not been measured for facial skin colors, these limits can be used to provide a warning of possible problems in captured skin colour.

NOTE: These studies did not take into account the possible impact on skin color due to dermatological conditions.

C.4 Coordinate calculation

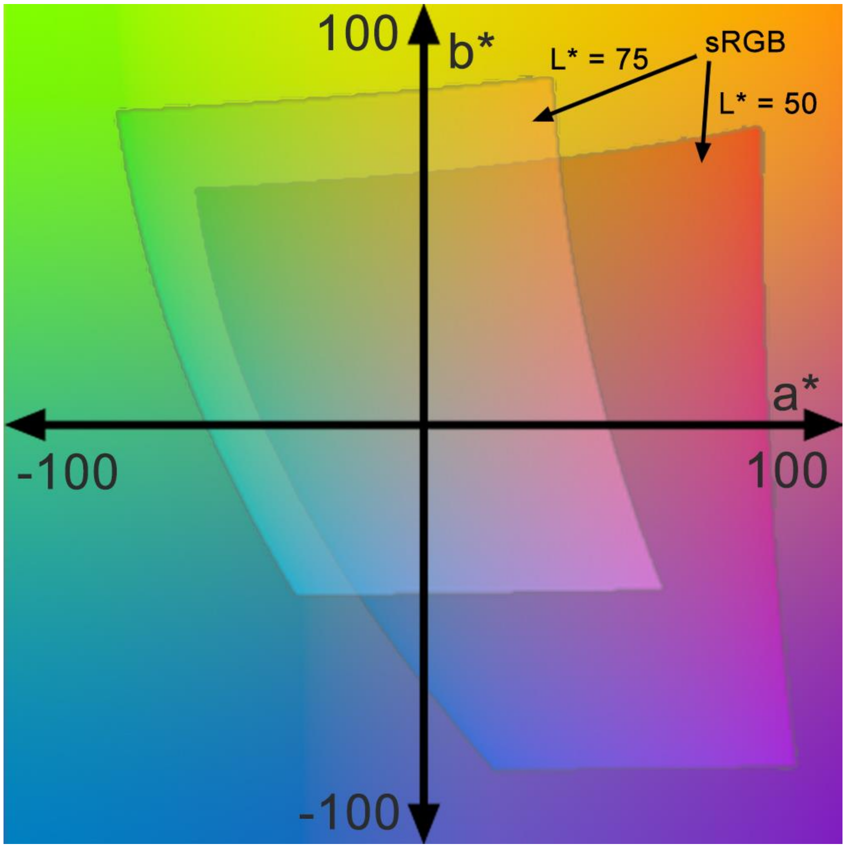


Figure C.4 — CIE Chroma (a*b*) luminance (L*) levels 50 and 75 show the sRGB [6] gamut compared to the entire a*b* area. Human skin tones should be located in the upper right hand sector.

In CIE L*a*b* the non-linear relations for L*, a* and b* are intended to mimic the logarithmic response of the human eye [4]. Colour quality is determined by converting the image to CIE L*a*b* [5] space and measuring the distance between the colours captured and known colour values of the chart patches. CIE Delta E 2000 [13] is a standard method for measuring this distance. A capture system's performance can be improved by minimizing the system's average (i.e., measured across all chart patches) and maximum (i.e., measured for any given chart patch) Delta E 2000.

An ideal system would have an average deltaE2000 of 1 and a maximum deltaE2000 of 5.

Annex D

(informative)

MTF test method according to ISO 12233:2014

Spatial resolution is a measure of the smallest discernible detail in an image. The image resolution measurements described herein are designed for the calibration of photo studios and office imaging systems. The methods generally involve photographing or scanning a standard target and analyzing the resulting images on a computer using standardized algorithms to compute a value.

Image fidelity factors are affected by the imaging sensor and lens. Resolution is a single frequency parameter that indicates whether the output signal contains a minimum threshold of detail information for visual detection (i.e., the highest spatial frequency that a camera or other similar imaging device can usefully capture).

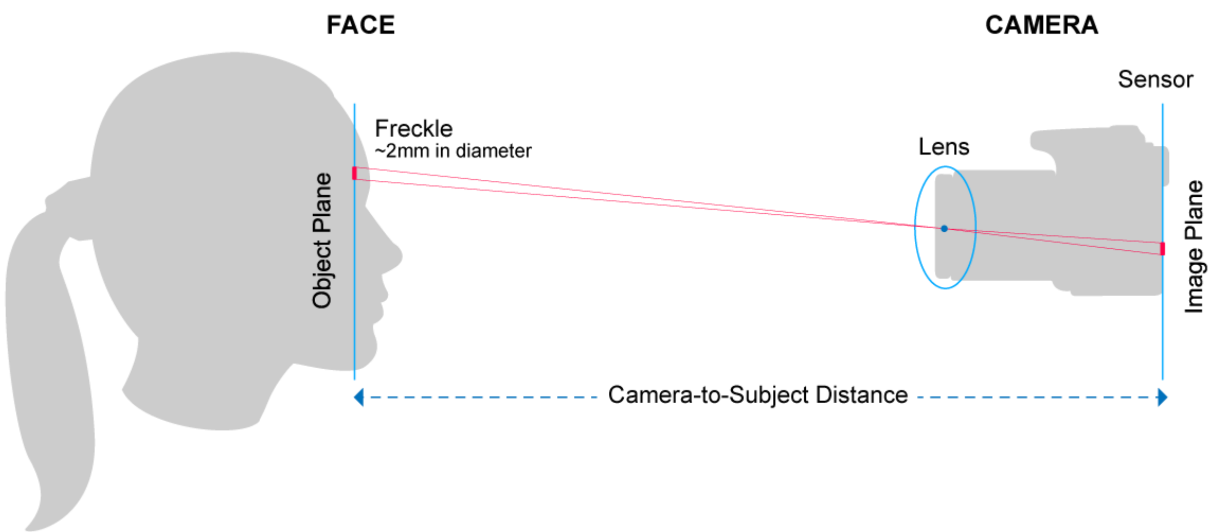


Figure D.1 — Relation between the physical size of an object at the face and its counterpart at the camera sensor.

Spatial frequency response is a multi-valued metric that measures contrast loss as a function of spatial frequency. Generally, contrast decreases as a function of spatial frequency to a level where detail can no longer be visually resolved. See Figure D.1. This limiting frequency value is the resolution of the camera, which is determined by the performance of the camera lens, the number of addressable photo elements in the optical imaging device, and the electrical circuits in the camera, which optionally perform image compression and gamma correction.

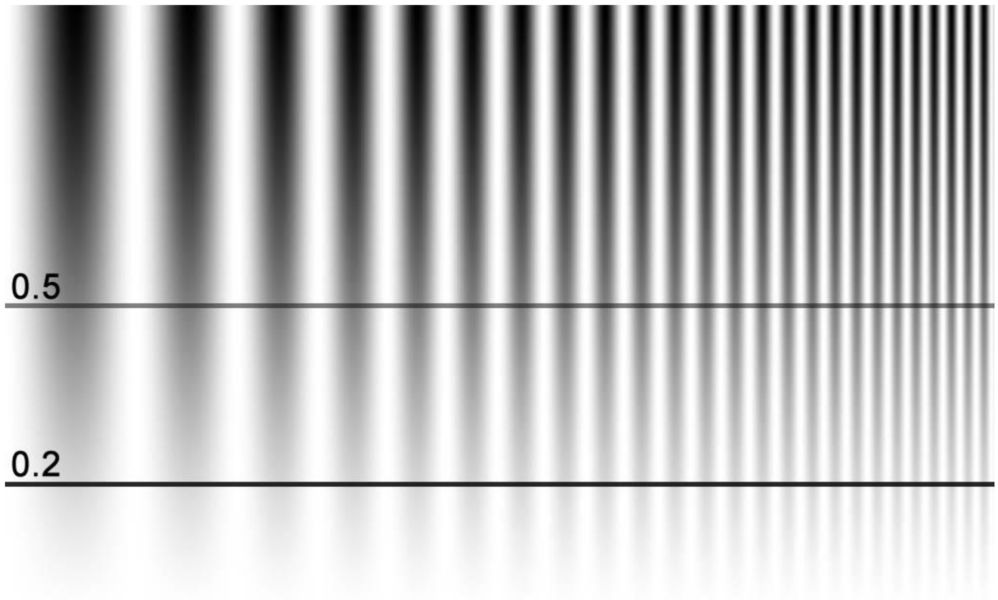


Figure D.2 — Sine wave test chart overlaid with a white contrast mask to show decreasing contrast effects from 100% value at the top to a 0% value at the bottom.

Modulation Transform Functions (MTFs) are normally measured by optics experts using purpose-printed sine wave test charts, such as in Figure D.2, and standardized procedures. Alternatively, MTF can be determined from the magnitude of the Fourier transform of a system's point or line spread function. The Fourier transform decomposes the spread function into the frequencies that comprise it. By looking at the amplitudes of the frequencies, it is possible to define the resolution characteristics of the measured imaging system.

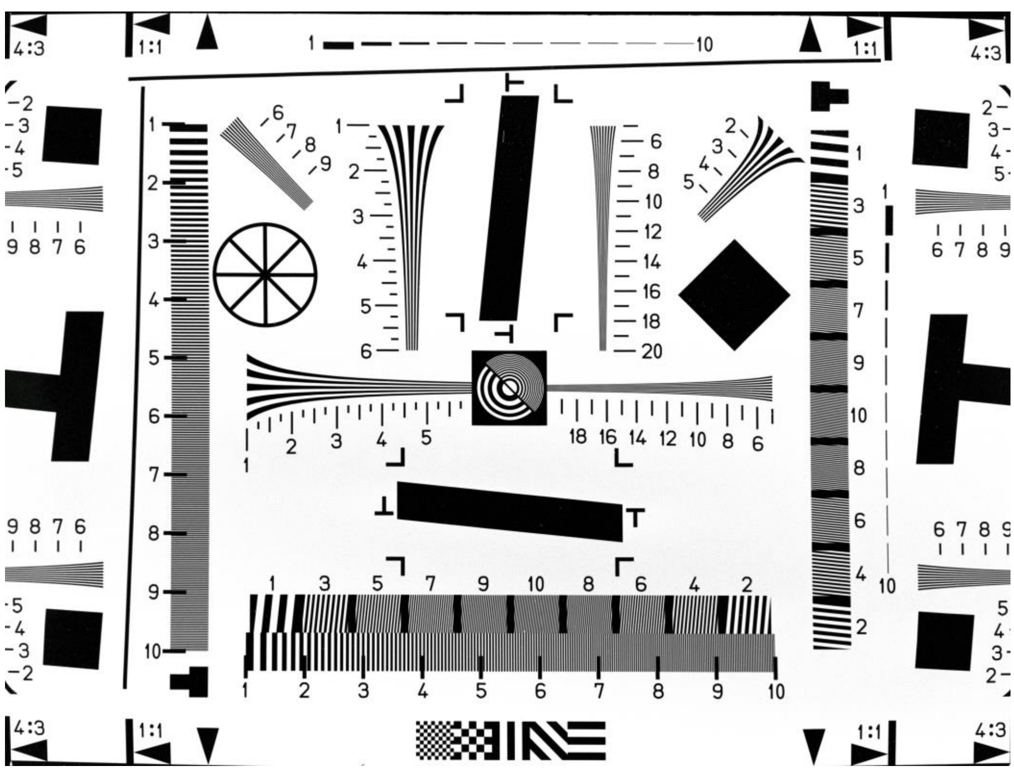


Figure D.3 — Test pattern: ISO 12233:2014.

ISO 12233:2014 specifies methods for measuring the Spatial Frequency Response (SFR) of electronic still picture cameras and similar imaging devices such as video cameras and flatbed scanners. The SFR measurement closely approximates the mathematically-defined system MTF of the camera. The MTF of the camera can only be approximated through the SFR, because most electronic still-picture cameras provide spatial colour sampling and nonlinear processing.

SFR is measured by capturing an image of a bi-tonal, rotated square test pattern, such as those shown in Figure D.3, which can subsequently be analyzed using readily-available edge analysis image processing software. It is important that the test pattern is captured in the same photographic environment that is prescribed for the face by this standard, including camera-to-subject distance, image dimensions, etc.

To perform slanted edge measurements using a software program, the user selects a region containing the edge to be measured. The software processes digital image values near slanted vertical and horizontal black to white edges to derive super sampled edge spread data, which is then filtered and converted to the frequency domain to get the SFR values. Horizontal edge is used for vertical SFR measurement.

Slanted edge measurements are less sensitive to noise than sine patterns. Gamma influences the MTF measurement accuracy, and for this reason, gamma value should be measured using a grey scale chart. Gamma correction, or often simply gamma, is the name of a nonlinear operation used to encode and decode luminance or colour related tristimulus values in imaging systems. An incorrect gamma setting for the MTF calculation causes an error situation where the MTF at 2 * Nyquist is not equal to 0, as it should be. In practice, a measurement of exactly 0 is not required to achieve an acceptable measurement.

Figure D.4 — Test pattern: ISO 16067-1. Note that the physical distance between the fiducials on the standardized test pattern is 66.8 mm.

An example SFR measured for an imaging system is shown in Figure D.4.The contrast value is depicted either in percent or on a scale from 0 to 1, where 1 corresponds to 100% on the vertical axis, and frequency values start from 0 and increase to the right on the horizontal axis. The frequency unit of an SFR measurement can be represented in cy/px or cy/mm. For the purposes of this standard, where the system specification references the object being photographed (i.e., the size of features on a face), the frequency unit should be in cycles/mm in the object plane. The size of a freckle/mole that should be detectable in face photos is 2 to 3 mm. The unit cy/mm is preferred over cy/px as sensor dimensions, camera distances, and head sizes vary. The physical distance of the object is the universal unit for all issuers. This measurement methodology requires that the user compute the sampling rate in px/mm, which is determined by measuring the number of pixels across the known physical dimensions of the test pattern.

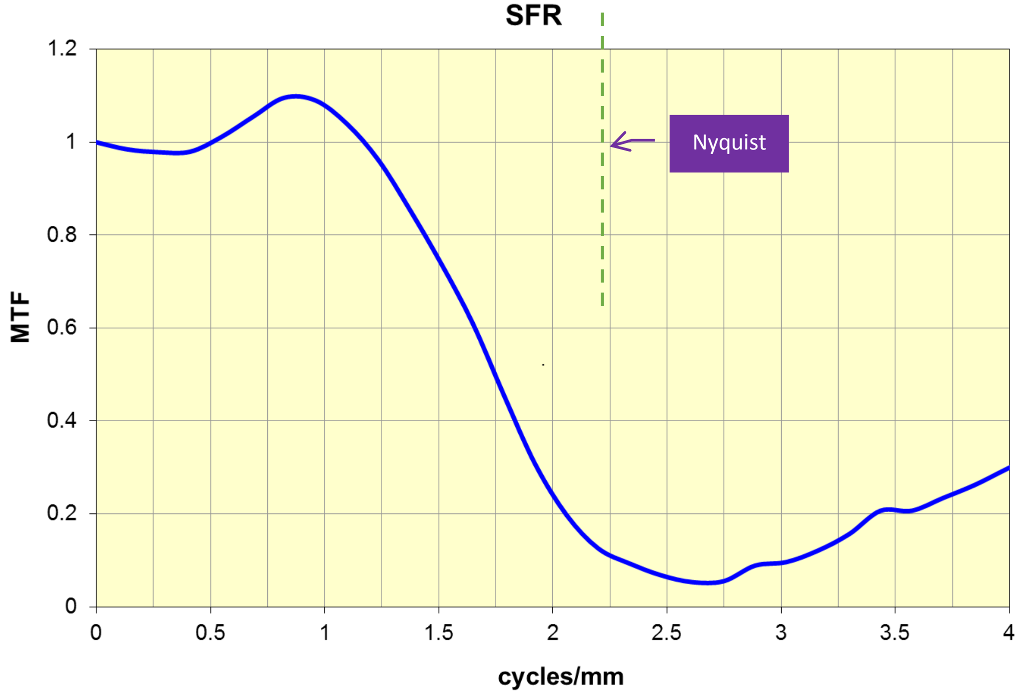


Figure D.5 — Example SFR graph.

The Nyquist frequency is half of the sampling rate of a discrete (opposite of continuous) signal processing system. The Nyquist frequency is also called the Nyquist limit. It is the highest frequency that can be encoded at a given sampling rate while still fully reconstructing the image. The Nyquist frequency for an imaging system is 0,5 cy/px as two samples (i.e., pixels) is the minimum required to represent a complete cycle. However, cy/px does not clarify the size of the object being resolved and must therefore be converted to cy/mm.

The SFR is deemed acceptable if the MTF remains sufficiently high up to a specified frequency. The Nyquist limit for the specification is determined by the minimum sampling rate specified in this standard multiplied by the Nyquist limit in cy/px. A different specification is needed for scanned photographs versus live capture, as the object plane in each case is different. The minimum sampling rates in this standard are 90 pixels per inter eye distance (approximately 60 mm) for cameras, and 300 PPI for scanners. The Nyquist limits are therefore: 0,75 cy/mm for cameras and 5,9 cy/mm for scanners.

MTF20, which is an indicator of SFR at higher spatial frequencies, should occur at approximately 80% of the Nyquist frequency, or 0,6 cy/mm for cameras and 4,7 cy/mm for scanners.

Annex E

(informative)

Focus and depth-of-field considerations

Proper focus and depth-of-field will be assured by pre-focusing the lens at the distance of the subject's eyes and by selecting an appropriate aperture (F-stop) to ensure a depth-of-field containing a subject's nose and ears. The depth-of-field of a lens is dependent upon its focal length, its effective aperture, and the focus distance. Point sources which are closer or farther than the distance at which a lens is well focused will be blurred, with the extent of the blur described by a "circle of confusion." If the maximum diameter of the circle of confusion is limited by, for example, the spacing between adjacent pixels in a CCD image sensor, the front and rear distances from the plane of optimum focus that produce acceptably focused images can be determined. The sum of these front and rear distances is the depth-of-field (D_{DoF}).

$$D_{DoF} = D_{front} + D_{rear}$$

$$D_{front} = \frac{cFs(s-f)}{f^2 + cF(s-f)}$$

$$D_{rear} = \frac{cFs(s-f)}{f^2 - cF(s-f)}$$

where:

D_{front} = the front focal distance, the distance from the plane of focus to the plane closest to the lens that is still in acceptable focus,

D_{rear} = the rear focal distance, the distance from the plane of focus to the plane farthest from the lens that is still in acceptable focus,

c = the diameter of the circle of confusion,

s = the distance from the lens to the object plane d

$F = f/a$ is the F-stop, the lens focal length f divided by the effective lens aperture a

Figure E.1 illustrates these dimensions.

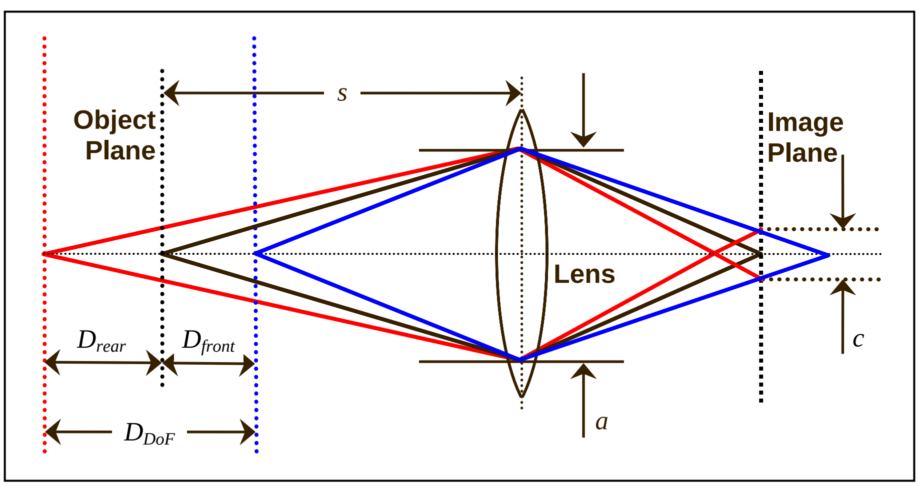


Figure E.1 — Dimensions for depth-of-field calculations

Annex F

(informative)

Report about the study of the effect of the camera subject distance of reference face images on face verification performances

F.1 Origin of this study

During the 51st meeting of ISO/IEC JTC1/SC17/WG3 in Vienna from 28 to 30 September 2015 a question was raised: May distortion of the facial image by taking photographs too close up, lead to decreased accuracy in facial recognition? ISO/IEC JTC1/SC17/WG3 composed a Special Group to work out to which degree the algorithms used for face comparison are impacted. This annex summarizes the findings of this distortion study group.

F.2 What is magnification distortion?

Taking photographs from a short camera-subject distance causes magnification distortion of face images. Figure F.1 illustrates the magnification distortion. Let the distance between the tip of the nose of a capture subject and the lens of a camera be *D*, the distance between the eyes plane of the capture subject and the lens of the camera, i.e. the camera-subject distance, be $D + \Delta D$, and the height of a structure at nose level be *h*. Then, the structure at nose level appears to have the same size as a structure of the height $h + \Delta h$ at eye level. The magnification distortion is defined as $K_{magnification} = \Delta D/(D+\Delta D) \times 100\%$.

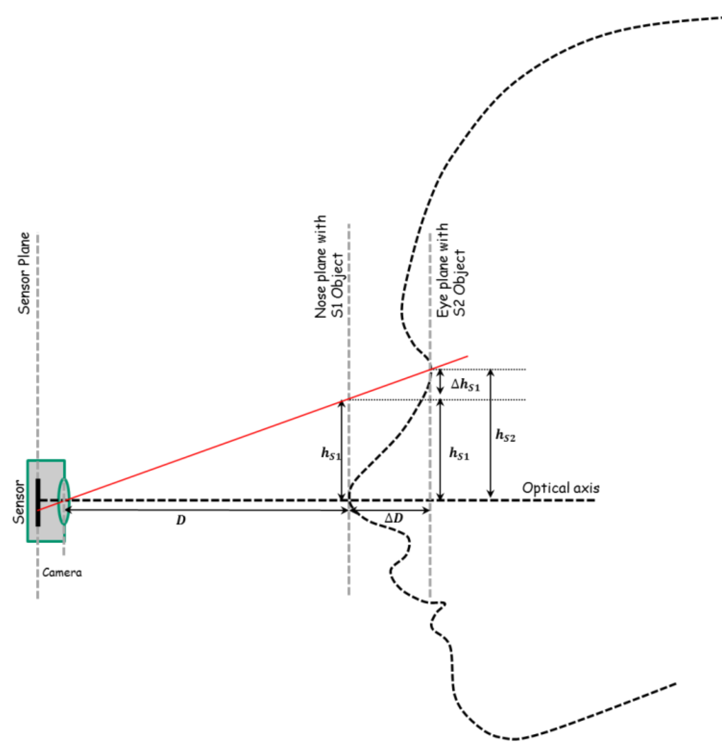


Figure F.1 — Illustration of magnification distortion

If the distance between the nose plane and the eyes plane ∆D is 5 cm, the magnification distortions are as in Table F.1.

Table F.1 — Magnification distortion as a function of camera-subject distance

Camera-subject distance	Magnification distortion $\Delta D/(D+\Delta D)$
0,5 m	10,0%
0,7 m	7,1%
1,0 m	5,0%
1,2 m	4,2%
1,5 m	3,3%
2,5 m	2,0%
3,0 m	1,7%

F.3 Methodology from enrolment to score calculation

A bench has been created capable of rapidly capturing pictures of the same subject under strictly controlled conditions. That bench (see Figure F.2) ensured that for all face images, the capture conditions were the same and in line with the provisions specified in this document, except for the camera subject distance, which is variable and ranges from 0,5 m to 3 m. The ten distances are 0,5 m, 0,6 m, 0,7 m, 0,8 m, 0,9 m, 1,0 m, 1,5 m, 2,0 m, 2,5 m, and 3,0 m.

The bench contains a Canon EOS 6D digital camera:

- 36 mm × 24 mm CMOS photo sensor with around 20,2 megapixels (5472 × 3648),
- Manual focus by SDK,
- ISO speed set to 100,
- Aperture = 22,
- Flash synchronisation: 1/160,
- White balance set to flash.

The lens used is a Canon EF 50 mm f/1,8 STM with radial distortion of less than 0,8%. The bench uses two front flashes PROFILITE 250 and one background flash PROFILITE 250 (each 250 W). The two front flashes are set to 5,0. The background flash is set to 1,0. The ambient light of the flashes is set to 50%.

After seating a test subject in front of the camera of the bench, a capture session is started. The camera automatically captures 40 face images in two passes. In each pass, the camera moves to ten positions with different camera-subject distances from 0,5 m to 3 m. In the first pass, the camera moves away from the test subject. In the second pass, the camera moves back towards the test subject. At each stop, two images are taken in order to mitigate the risk of closed-eye effect due to flash. After capturing 40 images, the capture session is completed.

Time between acquisitions is 12 s, the total duration of a capture session is 4 min. The precision of the movement is below 2 mm (0,08% of the full movement). Given the difficulty to locate the optical centre of the camera lens and given the morphological and behavioural differences of the test subjects, the actual camera-subject distances may be up to 30 mm smaller than the recorded camera-subject distances.

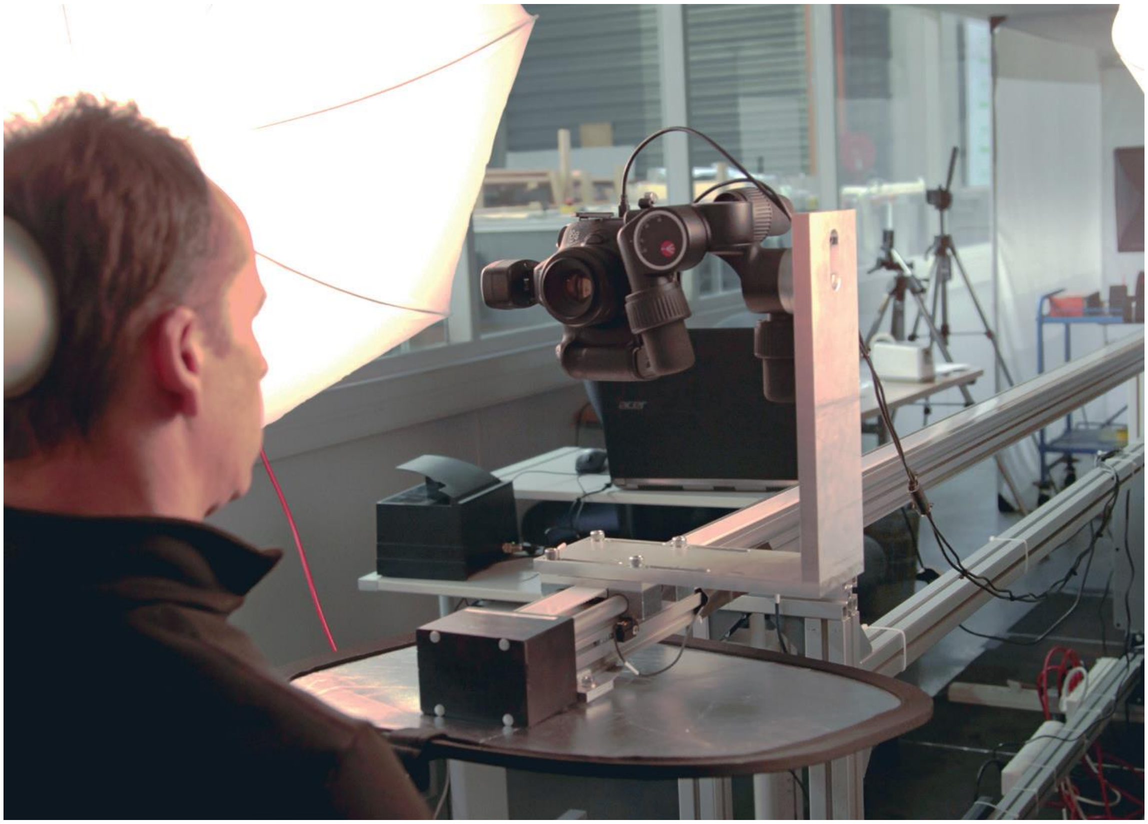


Figure F.2 – Bench created to rapidly capture a number of pictures at different distances

Using the bench, local databases of face images with different camera-subject distances were collected at the premises of members of the study group from as many volunteer test subjects as possible.

Local databases were collected at the premises of (in temporal order)

- KIS SAS in France,
- Oberthur Technologies in France,
- Photo-Me International in the UK,
- Gemalto in the Czech Republic,
- Fotofix Schnellphotoautomaten in Germany, and
- Nippon Auto-Photo in Japan.

Each test subject participated in only one capture session. The captured face images were cropped and resized in conformity with this document (i.e. "ICAO cropped"). The format of the ICAO cropped images was the JPEG file interchange format.

The local databases were encrypted and then sent to the Biometrics Evaluation Laboratory at the Fraunhofer Institute for Computer Graphics Research IGD, in order to be processed using various state-of-the-art face recognition algorithms.

The local databases were merged into one consolidated database containing 20 ICAO cropped face images from 435 test subjects, i.e. in total 8700 images. The filenames of all face images were pseudonymised such that it was not apparent from the filenames which face images matched and from which camera-subject distance they were captured. The consolidated face image database was divided into a directory of reference face images and a directory of probe face images. For each test subject, the directory of reference images contained ten ICAO cropped face images from the first pass (one per camera-subject distance), and the directory of probe images contained ten ICAO cropped face images from the second pass (one per camera-subject distance). The consolidated face image database was sequestered for the official run of face comparisons.

Providers of commercial off-the-shelf, state-of-the-art algorithms for one-to-one face comparison were invited to participate in this study. The following algorithm providers (in alphabetic order) submitted face comparison software:

- Dermalog,
- id3 Technologies,
- Innovatrics,
- NEC, and
- OT-Morpho.

The executable face comparison software was submitted by the algorithm providers to Fraunhofer IGD to be executed there on the sequestered face image database.

Each participating face comparison algorithm has been interfaced with two software interfaces. The participating algorithm providers were asked to build two executables in the form of Windows console applications:

- extract.exe, to extract comparable features from each of a list of face images, and
- compare.exe, to compare features and produce a comparison score for each pair of faces given in a comparison list.

Each participating face comparison algorithm compared the features from each reference face image with the features from each probe face image. This means $4350 \times 4350 = 18\,922\,500$ comparisons. For each comparison the file name of the reference image, the file name of the probe image, and the comparison score has been recorded in a CSV file.

F.4 Data analysis

F.4.1 Methodology

The research hypothesis is that

- the camera-subject distance of a face image or
- different magnification distortions of a face image and of probe face images compared with that face image

have an effect on the usefulness of that face image as a reference image. The usefulness of biometric sample for telling mated and non-mated samples apart is referred to as “utility” [17]. If for several of the participating state-of-the-art face comparison algorithms, the camera-subject distance of reference images has negligible effect on their utility, the research hypothesis would be refuted.

F.4.2 False Non-Match Rate at fixed False Match Rate

For four out of five of the participating algorithms, the highest non-mated similarity score is lower than the lowest mated similarity score, i.e. the distribution of mated scores was clearly separated from that of non-mated scores, allowing perfect classification by setting the decision threshold between the two distributions of scores. No matter what is the allowed value for FMR > 0 %, no false non-match error was observed in 43500 mated comparisons. Thus, according to the “Rule of 3” [18], with 95 % confidence FNMR ≤ 0,0069 % for most of the participating algorithms.

F.4.3 Distance-related differences in score distributions

Distance-related differences in score distributions can be significant in real scenarios where mated similarity scores are lower because of other factors affecting recognition performance (such as aging, pose variation, and illumination).

A measure of how well the distributions of mated and non-mated comparison scores are separated is d' (pronounced “d-prime”), defined as $d' = \frac{|\mu_m - \mu_n|}{\sqrt{{\sigma_m}^2 + {\sigma_n}^2}}$, where

- μ_m is the arithmetic mean of the mated comparison scores,
- μ_n is the arithmetic mean of the non-mated comparison scores,
- σ_m is the standard deviation of the mated comparison scores, and
- σ_n is the standard deviation of the non-mated comparison scores [19].

Figure F.3 shows the average d' values over the three best participating commercial face comparison algorithms as a function of the camera-subject distance of the reference image and the camera-subject distance of the probe image. The individual values are represented as colours. The lowest value is mapped to dark blue and the highest value to dark red.

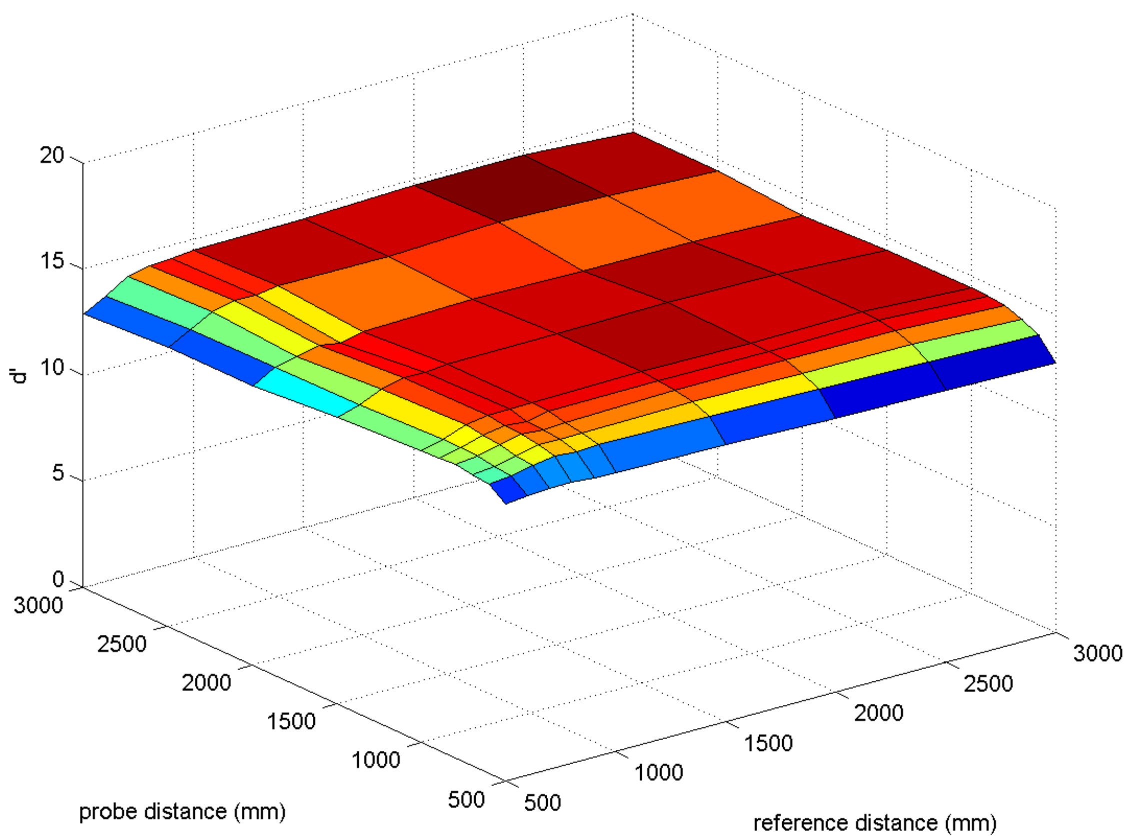


Figure F.3 – Average d' values over three commercial face comparison algorithms

F.5 Theoretical predictions

If a reference face image is compared with *N* face images, an algorithm returns *N* scores s_1 to s_N. We sort these *N* values in a decreasing order to get s'_1 to s'_N (s'_1 is the biggest score). Considering a *d'* value over 12 means that the probability to have s'_1 not correspond to the right person is below $2{,}15 \cdot 10^{-32}$ (see the demonstration below).

If N is the world population of $7{\cdot}10^9$ people, we estimate the probability to have one individual wrongly classified due to magnification distortion extrapolates to $77{\cdot}10^9 \cdot 2{,}15{\cdot}10^{-32} = 1{,}51{\cdot}10^{-22}$. This estimate assumes the matcher scores are normally distributed, that the mated pairs are captured in a single sitting, and that the matcher scores depend only on the probe and reference images (no per-search normalization). If our assumptions are valid, false rejects will hardly ever happen.

NOTE: Be aware that the performance of a facial recognition system also depends on other factors than distance, like e.g. illumination, pose, exposure, ageing, which may impact the mate score distribution.

Demonstration: Let U_n be a random variable representing a non-mated comparison score. Its mean is μ_n and its standard deviation is σ_n. We assume that the probability distribution of U_n is a normal distribution f_n. Let U_m be a random variable representing a mated comparison score. Its mean is μ_m and its standard deviation is σ_m. We assume that the probability distribution of U_m is a normal distribution f_m. Furthermore, we assume that the probability distribution of U_n and the probability distribution of U_m are independent.

Then, $X = U_m - U_n$ is also a random variable with a normal distribution f_X. Its mean is μ_X and its standard deviation is σ_X. We can determine the mean and the standard deviation of *X*:

$\mu_X = \mu_m - \mu_n$

$\sigma_X = \sqrt{\sigma_m^2 + \sigma_n^2}$

Figure F.4 shows examples of mated and non-mated score distributions and the distribution of differences of mated and non-mated scores.

In our study, $d' \in [12{,}7; 14{,}8]$. So, $d' = \frac{|\mu_m - \mu_n|}{\sqrt{{\sigma_m}^2 + {\sigma_n}^2}} = \frac{|\mu_X|}{\sigma_X} > 12$. So, $|\mu_X| > 12 \cdot \sigma_X$.

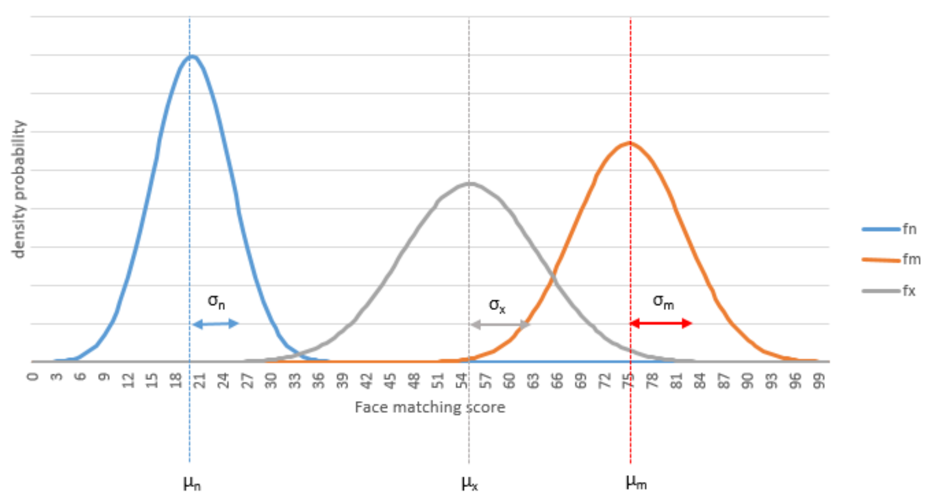


Figure F.4 – Distribution of differences of mated and non-mated scores

Let *X* be normally distributed with mean $|\mu_X| = 12 \cdot \sigma_X$ and standard deviation σ_X.

Let $X' = \frac{X-12\cdot\sigma_X}{\sigma_X}$. Then, *X'* is normally distributed with mean 0 and standard deviation 1, and $P(X' \leq -12) = 2{,}15 \cdot 10^{-32}$.

$$P(X' \leq -12) = P\left(\frac{X-12\cdot\sigma_X}{\sigma_X} \leq -12\right) = P(X - 12\cdot\sigma_X \leq -12\cdot\sigma_X) = P(X \leq 0)$$

So, $P(X \leq 0) = 2{,}15 \cdot 10^{-32}$.

F.6 Conclusions

Based on the data collected for this study, camera subject distance does not have a great influence on face verification performance over the range of camera subject distances investigated. Over a database of excellent-quality face images of 435 test subjects, captured at single capture sessions per test subject from different camera subject distances, several face verification algorithms avoided verification errors altogether.

It is noticeable that the average d' value is not sensitive at all to a reference distance above 0,7 m, i.e. below 7,1 % of magnification distortion. Nearly symmetrically, the average d' value is not sensitive at all to a probe distance above 0,7 m, i.e. below 7,1 % magnification distortion. Even at 0,5 m, i.e. at 10 % magnification distortion, d' decreases only 15 % to 12,7 versus the maximum value of 14,8 obtained with magnification distortion below 7,1 %.

The enrolment and verification images should be captured from a similar distance whenever possible.

BIBLIOGRAPHY

[1] IEC 61966-8:2001 Multimedia systems and equipment - Colour measurement and management - Part 8: Multimedia colour scanners

[2] ANSI IT8.7/2-1993 (R2013) Graphic technology - Color reflection target for input scanner calibration

[3] ISO 12641-1:2016 Graphic technology -- Prepress digital data exchange -- Colour targets for input scanner calibration -- Part 1: Colour targets for input scanner calibration

[4] ISO 11664-3:2012 (CIE S014-3/E:2011) Colorimetry -- Part 3: CIE tristimulus values

[5] ISO 11664-4:2008 (CIE S 014-4/E:2007) Colorimetry -- Part 4: CIE 1976 L*a*b* Colour space

[6] IEC 61966-2-1:1999 Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB

[7] Rochester Institute of Technology, https://www.cs.rit.edu/%7Encs/color/t_convert.html#RGB

[8] Field Application Report: Materials Characterization: UV/Vis/NIR Spectroscopy, Jillian F. Dlugos, Jeffrey L. Taylor, Ph.D., PerkinElmer, Inc., 2012

[9] Digital imaging & image processing techniques for the comparison of human hair features, Carolyn Jane McLaren, University of Canberra, Faculty of Applied Science, 2012

[10] Grading of Iris Color with an Extended Photographic Reference Set, Luuk Franssen, Joris E. Coppens, Thomas J.T.P. van den Berg, Netherlands Institute for Neuroscience, an institute of the Royal, Netherlands Academy of Arts and Sciences. Amsterdam, The Netherlands, 2008

[11] IEC 62676-4:2014 Video surveillance systems for use in security applications – Part 4: Application guidelines

[12] Spectral Filter Optimization for the Recovery of Parameters Which Describe Human Skin, Stephen J. Preece and Ela Claridge, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 7, July 2004

[13] CIE Delta E 2000. Improvement to industrial color-difference evaluation. Vienna: CIE Publication No. 142-2001, Central Bureau of the CIE; 2001.

[14] ISO 16067-1:2003 Photography -- Spatial resolution measurements of electronic scanners for photographic images -- Part 1: Scanners for reflective media

[15] Characterising the variations in ethnic skin colours: a new calibrated data base for human skin. K. Xiao, J. M. Yates, F. Zardawi, S. Sueeprasan, N. Liao, L. Gill, C.Li and S. Wuerger. Skin Research and Technology 2017; 23: 21–29. Published by John Wiley & Sons Ltd, 2016.

[16] Measuring Human Skin Colour. Mengmeng Wang, Kaida Xiao, Sophie Wuerger, Vien Cheung, Ming Ronnier Luo. 23rd Color and Imaging Conference Final Program and Proceedings, Society for Imaging Science and Technology, 2015.

[17] ISO/IEC 29794-1 Information technology – Biometric sample quality – Part 1: Framework.

[18] ISO/IEC 19795-1 Information technology – Biometric performance testing and reporting – Part 1: Principles and framework.

[19] R.M. Bolle, S. Pankanti, and N.K. Ratha. Evaluation techniques for biometrics-based authentication systems (FRR). In Proceedings of the 15th International Conference on Pattern Recognition ICPR, volume 2, 2000.