

DRAFT INTERNATIONAL STANDARD

ISO/IEC DIS 30116

ISO/IEC JTC 1/SC 31

Secretariat: ANSI

Voting begins on:
2015-09-30

Voting terminates on:
2015-12-30

Information technology — Automatic identification and data capture techniques — Optical Character Recognition (OCR) quality testing

Titre manque

ICS: 35.040

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/70c66299-7608-4027-870b-44d326537ece/iso-iec-30116-2016>

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.



Reference number
ISO/IEC DIS 30116:2015(E)

© ISO/IEC 2015

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/70c66299-7608-4027-870b-44d326537ece/iso-iec-30116-2016>



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2015, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

Contents

Page

Foreword.....	v
Introduction.....	vi
1 Scope.....	1
2 Normative references.....	1
3 Terms and definitions.....	1
3.1 binarized image.....	1
3.2 document reference edge.....	1
3.3 inspection area.....	2
3.4 COL.....	2
3.5 pixel.....	2
3.6 raw image.....	2
3.7 reference grey-scale image.....	2
3.8 scan grade.....	2
3.9 stroke width.....	2
3.10 SWT.....	2
3.11 symbol.....	3
3.12 X-Tolerance.....	3
3.13 Y-Tolerance.....	3
4 Abbreviated terms.....	3
5 Quality grading.....	3
6 Measurement methodology for OCR-B.....	3
6.1 Overview of methodology.....	3
6.2 Obtaining the test image.....	3
6.2.1 Measurement conditions.....	3
6.2.2 Raw image.....	4
6.2.3 Reference grey-scale image.....	4
6.2.4 Binarized image.....	4
6.3 Reference reflectivity measurements.....	4
6.3.1 General requirements.....	4
6.3.2 Light sources.....	4
6.3.3 Effective resolution.....	4
6.3.4 Optical geometry.....	4
6.3.5 Inspection area.....	7
6.4 Basis of symbol grading.....	8
6.5 Capture the raw image.....	8
6.6 Image assessment parameters and grading.....	8
6.6.1 Determining the document horizontal axis.....	8
6.6.2 Character best-fit algorithm.....	8
6.6.3 Position of a character.....	10
6.6.4 Character Evaluation Value (CEV) in the best-fit location.....	10
6.6.5 Background noise.....	11
6.6.6 Contrast PCS of the characters.....	11
7 Reporting the Grade.....	11
Annex A OCR-B Character Centreline Coordinates (Normative).....	13
Annex B Threshold Determination Method (Normative).....	21
B.1 Algorithm description.....	21
B.2 Example.....	21
Annex C OCR Reference Decode Algorithm (Normative).....	25
Annex D Example calculation of Character Evaluation Value (CEV)	

Bibliography

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/70c66299-7608-4027-870b-44d326537ece/iso-iec-30116-2016>

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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

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

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

ISO/IEC 30116 was prepared by Technical Committee ISO/TC 30, Subcommittee SC 31, *Automatic identification and data capture techniques*.

This second/third/... edition cancels and replaces the first/second/... edition(s), [clause(s) / subclause(s) / table(s) / figure(s) / annex(es)] of which [has / have] been technically revised.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/70c00299-7608-4027-870b-44d326537ece/iso-iec-30116-2016>

Introduction

For the inspection of ID documents, i.e. MRTDs (Machine Readable Travel Documents) according to ISO/IEC 7501 and driving licences according to ISO/IEC 18013, a reliable and ergonomic document inspection technology is essential. Considering RFID interoperability, strong improvement has been reached introducing mechanisms for interoperability evaluation and testing of MRTDs and reader devices. Similar standards for optical reading would improve the reliability of OCR. This is especially important because OCR of the document's MRZ (Machine Readable Zone) is essential for accessing BAC (Basic Access Control) and/or SAC (Supplementary Access Control) protected passports.

Thus, reliable OCR makes the performance of Automated Border Control systems as well as of many other applications more predictable. Furthermore, the evaluation of document reader products can be done much easier. This standardization project defines test methods to evaluate OCR document quality. Furthermore, it defines requirements ensuring the compliance to the applicable OCR standards. The project applies experiences from other domains such as bar code reading and possibly other test methods for OCR. Where conflicts in the specification work between MRTDs and driving licenses may arise, satisfying the definitions for MRTDs shall be given preference.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/70c66299-7608-4027-870b-44d326537ece/iso-iec-30116-2016>

Information technology — Automatic identification and data capture techniques — Optical Character Recognition (OCR) quality testing

1 Scope

This International Standard

- specifies the methodology for the measurement of specific attributes of OCR-B character strings;
- defines a method for evaluating these measurements and deriving an overall assessment of character string quality;
- defines a reference decode algorithm for OCR-B;
- gives information on possible causes of deviation from optimum grades to assist users in taking appropriate corrective action.

This International Standard applies to OCR-B as defined in ISO 1073-2, but its methodology can be applied partially or wholly to other OCR fonts.

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.

ISO/IEC 19762 *Information technology -- Automatic identification and data capture (AIDC) techniques -- Harmonized vocabulary*

3 Terms and definitions

3.1

binarized image

binary (black/white) image created by applying the Global Threshold to the pixel values in the reference grey-scale image

3.2

document reference edge

physical (i.e. mechanical) end of the surface with the MRZ whose position is determined by putting a black background under the surface with the MRZ and sliding the document up against a physical stop

3.3

inspection area

rectangular area which contains the entire symbol to be tested inclusive of its quiet zones

3.4

character outline limits

outlines of an ideal printed image of a character

Note: this is a qualitative evaluation utilized in ISO 1836 that is replaced in this standard with SWT

3.5

pixel

individual light-sensitive element in a light-sensitive array (e.g. CCD (charge coupled device) or CMOS (complementary metal oxide semiconductor) device)

3.6

raw image

matrix of the reflectance values in x and y coordinates across a two-dimensional image, derived from the discrete reflectance values of each pixel of the light-sensitive array

3.7

reference grey-scale image

raw image convolved with a synthesised circular aperture

3.8

scan grade

result of the assessment of a single scan of an OCR symbol, derived by taking the lowest grade achieved for any measured parameter of the reference grey-scale and binarized images

3.9

stroke width

nominal dimension perpendicular to the direction of the line making up an OCR character

3.10

stroke width template

the inner and outer character boundaries defined by circles whose centres follow the line created by the character centreline coordinates defined in Annex A

3.11**symbol**

group of OCR characters comprising the entire machine-readable entity (e.g. Machine Readable Zone (MRZ) as specified in ICAO Doc 9303, sizes ID-1, ID-2, and ID-3) including quiet zones and the document reference edge

3.12**X-Tolerance**

0,08 mm for Size I with a nominal stroke width of 0,35 mm

Note: 0,08 mm for Size I with a nominal stroke width of 0,35 mm was originally defined in ISO 1831 Table 2

3.13**Y-Tolerance**

0,15 mm for Size I with a nominal stroke width of 0,35 mm

Note: 0,15 mm for Size I with a nominal stroke width of 0,35 mm was originally defined in ISO 1831 Table 2

4 Abbreviated terms

COL = character outline limits

CEV = character evaluation value

MRZ = Machine Readable Zone

SWT = Stroke width template

5 Quality grading

Quality grades for Best-Fit, PCS, Position and Background Noise are determined as one of three levels, Recommended, Needs attention and Not recommended. The parameter with the lowest grade is the grade of the symbol.

6 Measurement methodology for OCR-B**6.1 Overview of methodology**

The basis of the measurement methodology is the evaluation of reflectance from the symbol. This methodology is also intended to correlate with conditions encountered in OCR scanning systems. The method starts by obtaining the raw image, which is a high-resolution grey-scale image of the symbol captured under controlled illumination and viewing conditions.

6.2 Obtaining the test image**6.2.1 Measurement conditions**

A test image of the symbol shall be obtained in a configuration that mimics the typical scanning situation for that symbol, but with substantially higher resolution (see 6.3.3), uniform illumination, and at best focus. The reference optical arrangement is defined in 6.3.4. Alternative optical arrangements may be used provided that the measurements obtained with them can be correlated with the use of the reference optical arrangement.

Ambient light levels shall be controlled in order not to influence the measurement results. Whenever possible, measurements shall be made on the symbol in its final configuration, i.e. the configuration in which it is intended to be scanned. For MRTD evaluation, optically personalized samples shall be used. This includes that all layers available at a document including laminations, security features and protective layers shall be present.

Two principles govern the design of the optical set-up. First, the test image's grey-scale shall be nominally linear and not be enhanced in any way. Second, the image resolution shall be adequate to produce consistent readings, which generally requires that the character stroke-widths span at least ten image pixels.

6.2.2 Raw image

The raw image is a matrix of the actual reflectance values for each pixel of the light-sensitive array, from which are derived the reference grey-scale image and the binarized image which are evaluated for the assessment of symbol quality.

6.2.3 Reference grey-scale image

The reference grey-scale image is obtained from the raw image by processing the individual pixel reflectance values through a synthetic circular aperture equal to 0,2 mm.

6.2.4 Binarized image

The binarized image is obtained from the reference grey-scale image by applying the algorithm defined in Annex B.

6.3 Reference reflectivity measurements

6.3.1 General requirements

Equipment for assessing the quality of symbols in accordance with this clause shall comprise a means of measuring and analysing the variations in the reflectivity of a symbol on its substrate over an inspection area which shall cover the full height and width of the symbol.

The measured reflectance values shall be expressed in percentage terms by means of calibration and reference to recognised national standards laboratories, where 100 per cent should correspond to the reflectance of a barium sulphate or magnesium oxide reference sample.

It should be ensured that all materials visible to the camera or close to the optical path are reflection free, at least in IR illumination. In particular, the background the symbol is attached to shall be IR absorbing. The environment temperature shall be between 20 and 25°C.

6.3.2 Light sources

Measurements shall be made using LED light sources at 890 and 940 nm wavelengths.

All illumination elements shall have a diameter of 25 mm or less and may be shaped as circles, squares or similar.

6.3.3 Effective resolution

The effective resolution of an instrument that implements this international standard shall be sufficient to ensure that the parameter grading results are consistent irrespective of the rotation of the symbol. The effective resolution is the product of the resolution of the light-sensitive array and of the magnification of the associated optical system and effected by distortions introduced by the optical system. The reference optical arrangement requires an effective resolution of not less than ten pixels per stroke width.

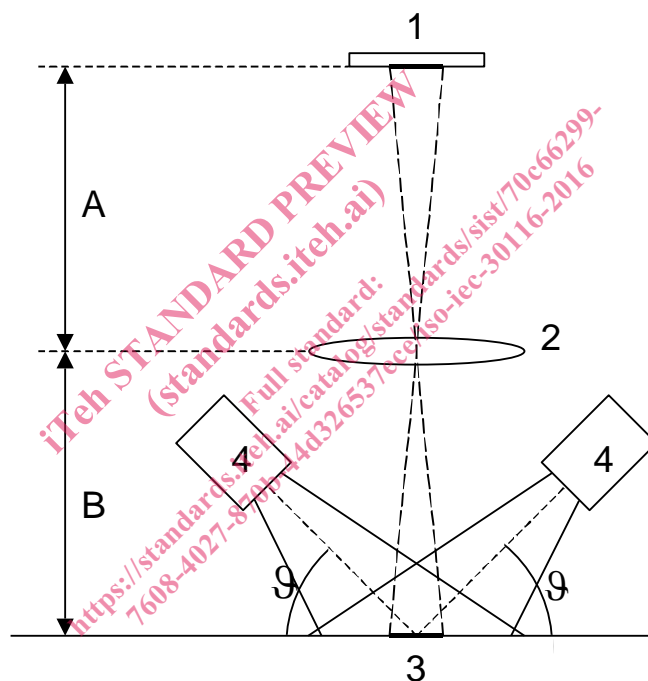
6.3.4 Optical geometry

A reference optical geometry is defined for reflectivity measurements and consists of:

- flood incident illumination, uniform across the inspection area, from a set of four light sources arranged at 90 degree intervals around a circle concentric with the inspection area and in a plane parallel to that of the inspection area, at a height which will allow incident light to fall on the centre of the inspection area at an angle of 45° to its plane, and
- a light collection device, the optical axis of which is perpendicular to the inspection area and passes through its centre, and which focuses an image of the test symbol on a light-sensitive array.

The light reflected from the inspection area shall be collected and focussed on the light-sensitive array.

Implementations may use alternative optical geometries and components, provided that their performance can be correlated with that of the reference optical arrangement defined in this section. Figures 1 and 2 illustrate the principle of the optical arrangement, but are not intended to represent actual devices; in particular the magnification of the device is likely to differ from 1:1. For example, it is possible to use a 10MP industrial camera without IR cut filter with a sensor size of 1/2". The image could be captured from a distance of approx. 350 mm and the lens chosen appropriately. The resulting magnification then would be 1:21.



- 1 – Light sensing element
- 2 – Lens providing 1:1 magnification (measurement A = measurement B)
- 3 – Inspection area
- 4 – Light sources
- ϑ - Angle of incidence of light relative to plane of symbol = 45°

Figure 1 — Reference optical arrangement – side view