INTERNATIONAL STANDARD

ISO/IEC 15938-13

First edition
2015-09-01

Information technology — Multimedia content description interface —

Part 13: Compact descriptors for visual search

Technologies de l'information — Interface de description du contenu multimédia —

Partie 13: Descripteurs compacts pour recherche visuelle

Reference number
ISO/IEC 15938-13:2015(E)

© ISO/IEC 2015



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 Terms and definitions **1**

3 Symbols and abbreviated terms **2**
3.1 General 2
3.2 Abbreviations 2
3.3 Arithmetic operations 3
3.4 Logical operators 3
3.5 Relational operators 3
3.6 Bitwise operators 4
3.7 Assignment 4
3.8 Mnemonics 4
3.9 Constants 4
3.10 Functions 4

4 CDVS syntax **5**
4.1 Binary representation syntax 5
4.2 Descriptor component semantics 6

5 CDVS encoding **9**
5.1 General 9
5.2 Original image preprocessing 9
5.3 Interest point detection 9
5.3.1 Introduction 9
5.3.2 Scale space construction 9
5.3.3 Detection of scale-space extrema 10
5.3.4 Coordinate refinement to subpixel precision. 14
5.3.5 Transformation of coordinates and scale to the converted image resolution 17
5.3.6 Elimination of duplicates 17
5.3.7 Orientation Assignment 17
5.3.8 Interest point characteristics 19
5.4 Local feature selection 19
5.4.1 Operation 19
5.4.2 Descriptor components 20
5.5 Local feature description 21
5.6 Local feature descriptor aggregation 23
5.6.1 Operation 23
5.6.2 Descriptor components 26
5.7 Local feature descriptor compression 27
5.7.1 Operation 27
5.7.2 Descriptor components 30
5.8 Local feature location compression 31
5.8.1 Operation 31
5.8.2 Descriptor components 36
5.9 Encoding order of compressed local feature descriptors and relevance bits 37
5.10 Computation of the number of compressed local feature descriptors at different image descriptor lengths 37

Annex A (informative) **CDVS encoder organization** **38**

Annex B (normative) **Coefficients for coordinate refinement** **39**

Annex C (normative) **Probability values for the feature selection** **41**

Annex D (normative) **PCA projection matrix for local feature descriptor aggregation** **44**

© ISO/IEC 2015 – All rights reserved

Annex E (normative) **GMM parameters for local feature descriptor aggregation** **55**
Annex F (normative) **Gaussian function selection parameters for local feature descriptor aggregation** **135**
Annex G (normative) **Bit selection masks for local feature descriptor aggregation** **136**
Annex H (normative) **Scalar quantization thresholds for local feature descriptor compression** **138**
Annex I (normative) **Histogram count arithmetic coding model probabilities** **142**
Annex J (normative) **Histogram map arithmetic coding model probabilities** **144**
Annex K (informative) **CDVS decoding** **145**

 © ISO/IEC 2015 – All rights reserved

Foreword

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

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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

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

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

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 15938 consists of the following parts, under the general title *Information technology — Multimedia content description interface*:

— *Part 1: Systems*

— *Part 2: Description definition language*

— *Part 3: Visual*

— *Part 4: Audio*

— *Part 5: Multimedia description schemes*

— *Part 6: Reference software*

— *Part 7: Conformance testing*

— *Part 8: Extraction and use of MPEG-7 descriptions*

— *Part 9: Profiles and levels*

— *Part 10: Schema definition*

— *Part 11: MPEG-7 profile schemas*

— *Part 12: Query format*

— *Part 13: Compact descriptors for visual search*

© ISO/IEC 2015 – All rights reserved

Introduction

This International Standard, also known as "Multimedia Content Description Interface," provides a standardized set of technologies for describing multimedia content. It addresses a broad spectrum of multimedia applications and requirements by providing a metadata system for describing the features of multimedia content.

The following are specified in this International Standard:

— **Description schemes (DS)** describe entities or relationships pertaining to multimedia content. Description schemes specify the structure and semantics of their components, which may be Description Schemes, descriptors, or datatypes.

— **Descriptors (D)** describe features, attributes, or groups of attributes of multimedia content.

— **Datatypes** are the basic reusable datatypes employed by description schemes and descriptors.

— **Systems tools** support delivery of descriptions, multiplexing of descriptions with multimedia content, synchronization, file format, and so forth.

This International Standard is subdivided into 13 parts:

— **Part 1 — Systems**: specifies the tools for preparing descriptions for efficient transport and storage, compressing descriptions, and allowing synchronization between content and descriptions.

— **Part 2 — Description definition language**: specifies the language for defining the International Standard set of description tools (DSs, Ds, and datatypes) and for defining new description tools.

— **Part 3 — Visual**: specifies the description tools pertaining to visual content.

— **Part 4 — Audio**: specifies the description tools pertaining to audio content.

— **Part 5 — Multimedia description schemes**: specifies the generic description tools pertaining to multimedia including audio and visual content.

— **Part 6 — Reference software**: provides a software implementation of the International Standard.

— **Part 7 — Conformance testing**: specifies the guidelines and procedures for testing conformance of implementations of the International Standard.

— **Part 8 — Extraction and use of MPEG-7 descriptions**: provides guidelines and examples of the extraction and use of descriptions.

— **Part 9 — Profiles and levels**: provides guidelines and standard profiles.

— **Part 10 — Schema definition**: specifies the schema using description definition language.

— **Part 11 — Profile Schemas**: listing of profile schemas using description definition language.

— **Part 12 — Query format**: contains the tools of the MPEG Query Format (MPQF).

— **Part 13 — Compact descriptors for visual search**: specifies an image description tool for visual search applications.

© ISO/IEC 2015 – All rights reserved

Information technology — Multimedia content description interface —

Part 13: Compact descriptors for visual search

1 Scope

The structure of this part of ISO/IEC 15938 is as follows. Clauses 2 and 3 specify the terms, abbreviations, symbols, and conventions used in the International Standard. Clause 4 specifies the binary representation syntax and descriptor component semantics for a CDVS image descriptor. Clause 5 specifies the extraction and encoding process for a CDVS image descriptor. Annexes A-J specify information relevant to the encoding process of Clause 5. Annex K contains an informative description of the decoding process of a CDVS image descriptor.

This part of the MPEG-7 standard specifies an image description tool designed to enable efficient and interoperable visual search applications, allowing visual content matching in images. Visual content matching includes matching of views of objects, landmarks, and printed documents, while being robust to partial occlusions as well as changes in viewpoint, camera parameters, and lighting conditions.

2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

2.1
image descriptor
descriptor extracted from one image

2.2
image descriptor length
size of an image descriptor in bytes

Note 1 to entry: This International Standard specifies six average (i.e. over a large number of images) image descriptor lengths, i.e. 512 bytes, 1024 bytes, 2048 bytes, 4096 bytes, 8192 bytes, and 16384 bytes, and the encoding process for each image descriptor length.

2.3
original image
input image to the image descriptor encoder

2.4
converted image
image which is a spatially resampled version of the original image and from which the image descriptor is extracted

2.5
pixel
indexable element of the original image or the converted image, comprising spatial coordinates and a luminance value

2.6
interest point
point in an image showing detection stability under local and global perturbations in the image domain, including perspective transformations, changes in image scale, and illumination variations

© ISO/IEC 2015 – All rights reserved