

First edition
2002-06-15

Information technology — Multimedia content description interface —

Part 4: Audio

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

Partie 4: Audio

Reference number
ISO/IEC 15938-4:2002(E)



This is a preview of "ISO/IEC 15938-4:2002". [Click here to purchase the full version from the ANSI store.](#)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

© ISO/IEC 2002

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

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland

This is a preview of "ISO/IEC 15938-4:2002". [Click here to purchase the full version from the ANSI store.](#)

Contents

	Page
Foreword.....	v
Introduction.....	vi
1 Scope	1
1.1 Definition of Scope	1
1.2 Fields of application	1
2 Terms and definitions	2
3 Symbols and abbreviated terms	2
4 Conventions	3
4.1 Description Definition Language	3
4.2 Audio representation.....	3
5 Audio Framework	4
5.1 Introduction	4
5.2 Scalable Series	4
5.2.1 Introduction	4
5.2.2 ScalableSeriesType	5
5.2.3 SeriesOfScalarType	6
5.2.4 SeriesOfScalarBinaryType.....	9
5.2.5 SeriesOfVectorType	10
5.2.6 SeriesOfVectorBinaryType	13
5.3 Low level Audio Descriptors	13
5.3.1 Introduction	13
5.3.2 AudioLLDScalarType	14
5.3.3 AudioLLDVectorType	15
5.3.4 AudioWaveformType	16
5.3.5 AudioPowerType	17
5.3.6 Audio Spectrum Descriptors	17
5.3.7 AudioSpectrumEnvelopeType.....	18
5.3.8 AudioSpectrumCentroidType.....	21
5.3.9 AudioSpectrumSpreadType	23
5.3.10 AudioSpectrumFlatnessType	24
5.3.11 AudioSpectrumBasisType	26
5.3.12 AudioSpectrumProjectionType	29
5.3.13 AudioHarmonicityType	33
5.3.14 Timbre Descriptors	36
5.3.15 LogAttackTimeType	38
5.3.16 HarmonicSpectralCentroidType.....	39
5.3.17 HarmonicSpectralDeviationType	41
5.3.18 HarmonicSpectralSpreadType	42
5.3.19 HarmonicSpectralVariationType	44
5.3.20 SpectralCentroidType	45
5.3.21 TemporalCentroidType	46
5.4 Silence	46
5.4.1 Introduction	46
5.4.2 SilenceHeaderType.....	47
5.4.3 SilenceType	47
5.4.4 Usage, examples and extraction (informative).....	48
6 High Level Tools	49
6.1 Introduction	49
6.2 Audio Signature	49

This is a preview of "ISO/IEC 15938-4:2002". [Click here to purchase the full version from the ANSI store.](#)

6.2.1	Introduction	49
6.2.2	AudioSignatureType	50
6.2.3	Instantiation requirements	50
6.2.4	Usage and examples (informative)	50
6.3	Timbre	51
6.3.1	Introduction	51
6.3.2	InstrumentTimbreType	52
6.3.3	HarmonicInstrumentTimbreType	53
6.3.4	PercussiveInstrumentTimbreType	54
6.3.5	Usage, extraction and examples (informative)	55
6.4	General Sound Recognition and Indexing	56
6.4.1	Introduction	56
6.4.2	SoundModelType	57
6.4.3	SoundClassificationModelType	59
6.4.4	SoundModelStatePathType	61
6.4.5	SoundModelStateHistogramType	62
6.4.6	General Sound Classification and Indexing Applications (informative)	64
6.5	Spoken Content	66
6.5.1	Introduction	66
6.5.2	SpokenContentHeaderType	67
6.5.3	SpeakerInfoType	68
6.5.4	SpokenContentIndexEntryType	71
6.5.5	ConfusionCountType	71
6.5.6	WordType, PhoneType, WordLexiconIndexType and PhoneLexiconIndexType	73
6.5.7	LexiconType	74
6.5.8	WordLexiconType	74
6.5.9	phoneticAlphabetType	75
6.5.10	PhoneLexiconType	75
6.5.11	SpokenContentLatticeType	76
6.5.12	SpokenContentLinkType	78
6.5.13	Usage, extraction and examples (informative)	79
6.6	Melody	84
6.6.1	Introduction	84
6.6.2	MelodyType	84
6.6.3	Meter	85
6.6.4	scaleType	86
6.6.5	MelodyKey	86
6.6.6	MelodyContourType	88
6.6.7	contourType	88
6.6.8	beatType	89
6.6.9	MelodySequence	90
6.6.10	Usage of MelodyContour (informative)	92
6.6.11	Usage of MelodySequence (informative)	94
6.6.12	Examples (informative)	94
Annex A (informative) Usage, extraction and examples of Scalable Series		96
Annex B (informative) Patent statements		105

This is a preview of "ISO/IEC 15938-4:2002". [Click here to purchase the full version from the ANSI store.](#)

Foreword

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

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

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

ISO/IEC 15938-4 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee 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*

Annexes A and B of this part of ISO/IEC 15938 are for information only.

Introduction

This standard, also known as "Multimedia Content Description Interface," provides a standardized set of technologies for describing multimedia content. The standard 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 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
- **Description Definition Language (DDL)** defines Description Schemes, Descriptors, and Datatypes by specifying their syntax, and allows their extension.
- **Systems tools** support delivery of descriptions, multiplexing of descriptions with multimedia content, synchronization, file format, and so forth.

This standard is subdivided into eight 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 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 standard.

Part 7 – Conformance testing: specifies the guidelines and procedures for testing conformance of implementations of the standard.

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