

Australian/New Zealand Standard™

**Information technology—Multimedia  
content description interface**

**Part 4: Audio**

### **AS/NZS 15938.4:2003**

This Joint Australian/New Zealand Standard was prepared by Joint Technical Committee IT-001, Information Systems—Interconnection. It was approved on behalf of the Council of Standards Australia on 22 April 2003 and on behalf of the Council of Standards New Zealand on 22 April 2003. It was published on 2 June 2003.

---

The following are represented on Committee IT-001:

Australian Bureau of Statistics  
Australian Communications Industry Forum  
Australian Information Industry Association  
Australian Telecommunications Users Group  
Australian Vice-Chancellors Committee  
Electrical Compliance Testing Association  
Information Technology Association of New Zealand

---

#### **Keeping Standards up-to-date**

Standards are living documents which reflect progress in science, technology and systems. To maintain their currency, all Standards are periodically reviewed, and new editions are published. Between editions, amendments may be issued. Standards may also be withdrawn. It is important that readers assure themselves they are using a current Standard, which should include any amendments which may have been published since the Standard was purchased.

Detailed information about joint Australian/New Zealand Standards can be found by visiting the Standards Australia web site at [www.standards.com.au](http://www.standards.com.au) or Standards New Zealand web site at [www.standards.co.nz](http://www.standards.co.nz) and looking up the relevant Standard in the on-line catalogue.

Alternatively, both organizations publish an annual printed Catalogue with full details of all current Standards. For more frequent listings or notification of revisions, amendments and withdrawals, Standards Australia and Standards New Zealand offer a number of update options. For information about these services, users should contact their respective national Standards organization.

We also welcome suggestions for improvement in our Standards, and especially encourage readers to notify us immediately of any apparent inaccuracies or ambiguities. Please address your comments to the Chief Executive of either Standards Australia International or Standards New Zealand at the address shown on the back cover.

---

Australian/New Zealand Standard™

**Information technology—Multimedia  
content description interface**

**Part 4: Audio**

First published as AS/NZS 15938.4:2003.

**COPYRIGHT**

© Standards Australia/Standards New Zealand

All rights are reserved. No part of this work may be reproduced or copied in any form or by any means, electronic or mechanical, including photocopying, without the written permission of the publisher.

Jointly published by Standards Australia International Ltd, GPO Box 5420, Sydney, NSW 2001 and Standards New Zealand, Private Bag 2439, Wellington 6020

ISBN 0 7337 5252 7

## PREFACE

This Standard was prepared by the Joint Standards Australia/Standards New Zealand Committee IT-001, Information Systems—Interconnection. This Standard is identical with and has been reproduced from ISO/IEC 15938-4:2002, *Information technology—Multimedia content description interface*, Part 4: *Audio*.

The objective of this Standard is to provide 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.

This Standard is Part 4 of AS/NZS 15938, *Information technology—Multimedia content description interface*, which, when complete, will consist of the following:

Part 1: Systems

Part 2: Description definition language

Part 3: Visual

Part 4: Audio (this Standard)

Part 5: Multimedia description schemes

Part 6: Reference software

Part 7: Conformance testing

Part 8: Extraction and use of MPEG-7 descriptions

The term ‘informative’ has been used in this Standard to define the application of the annex to which it applies. An ‘informative’ annex is only for information and guidance.

As this Standard is reproduced from an International Standard, the following applies:

- (a) Its number does not appear on each page of text and its identity is shown only on the cover and title page.
- (b) In the source text ‘this part of ISO/IEC 15938’ should read ‘this part of AS/NZS 15938’.
- (c) A full point should be substituted for a comma when referring to a decimal marker.

## CONTENTS

	<i>Page</i>
<b>1</b>	<b>Scope .....</b> 1
1.1	Definition of Scope .....
1.2	Fields of application .....
<b>2</b>	<b>Terms and definitions .....</b> 2
<b>3</b>	<b>Symbols and abbreviated terms .....</b> 2
<b>4</b>	<b>Conventions .....</b> 3
4.1	Description Definition Language .....
4.2	Audio representation.....
<b>5</b>	<b>Audio Framework .....</b> 4
5.1	Introduction .....
5.2	Scalable Series .....
5.2.1	Introduction .....
5.2.2	ScalableSeriesType .....
5.2.3	SeriesOfScalarType .....
5.2.4	SeriesOfScalarBinaryType.....
5.2.5	SeriesOfVectorType .....
5.2.6	SeriesOfVectorBinaryType .....
5.3	Low level Audio Descriptors .....
5.3.1	Introduction .....
5.3.2	AudioLLDScalarType .....
5.3.3	AudioLLDVectorType .....
5.3.4	AudioWaveformType .....
5.3.5	AudioPowerType .....
5.3.6	Audio Spectrum Descriptors .....
5.3.7	AudioSpectrumEnvelopeType.....
5.3.8	AudioSpectrumCentroidType.....
5.3.9	AudioSpectrumSpreadType .....
5.3.10	AudioSpectrumFlatnessType .....
5.3.11	AudioSpectrumBasisType .....
5.3.12	AudioSpectrumProjectionType .....
5.3.13	AudioHarmonicityType .....
5.3.14	Timbre Descriptors .....
5.3.15	LogAttackTimeType .....
5.3.16	HarmonicSpectralCentroidType.....
5.3.17	HarmonicSpectralDeviationType .....
5.3.18	HarmonicSpectralSpreadType .....
5.3.19	HarmonicSpectralVariationType .....
5.3.20	SpectralCentroidType .....
5.3.21	TemporalCentroidType .....
5.4	Silence .....
5.4.1	Introduction .....
5.4.2	SilenceHeaderType.....
5.4.3	SilenceType .....
5.4.4	Usage, examples and extraction (informative).....
<b>6</b>	<b>High Level Tools .....</b> 49
6.1	Introduction .....
6.2	Audio Signature .....

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
<b>Annex A (informative) Usage, extraction and examples of Scalable Series .....</b>		<b>96</b>
<b>Annex B (informative) Patent statements .....</b>		<b>105</b>

## AUSTRALIAN/NEW ZEALAND STANDARD

**Information technology — Multimedia content description interface —****Part 4:  
Audio****1 Scope****1.1 Definition of Scope**

This International Standard defines a Multimedia Content Description Interface, specifying a series of interfaces from system to application level to allow disparate systems to interchange information about multimedia content. It describes the architecture for systems, a language for extensions and specific applications, description tools in the audio and visual domains, as well as tools that are not specific to audio-visual domains. As a whole, this International Standard encompassing all of the aforementioned components is known as “MPEG-7.” MPEG-7 is divided into eight parts (as defined in the Foreword).

This part of the MPEG-7 Standard (Part 4: Audio) specifies description tools that pertain to multimedia in the audio domain. See below for further details of application.

This part of the MPEG-7 Standard is intended to be implemented in conjunction with other parts of the standard. In particular, MPEG-7 Part 4: Audio assumes knowledge of Part 2: Description Definition Language (DDL) in its normative syntactic definitions of Descriptors and Description Schemes. This part of the standard also has dependencies upon clauses in Part 5: Multimedia Description Schemes, namely many of the fundamental Description Schemes that extend the basic type capabilities of the DDL.

MPEG-7 is an extensible standard. The method to extend the standard beyond the Description Schemes provided in the standard is to define new ones in the DDL, and to make those DSs available with the instantiated descriptions. Further details are available in Part 2. To avoid duplicate functionality with other parts of the standard, the DDL is the only extension facility provided.

**1.2 Fields of application**

MPEG-7 Part 4: Audio is applicable to all forms of audio content. The encoding format or medium of the said audio is not limited in any way, and may include audio held in an analogue medium such as magnetic tape or optical film. The content of the audio is not limited within or without music, speech, sound effects, soundtracks, or any mixtures thereof.

The tools listed in this part of the International Standard are applicable to both audio in isolation and to audio associated with video.

The specific tools provided within the Audio portion of the standard are designed to work in conjunction with the Multimedia Description Schemes that apply to both audio and video. Because of the “toolbox” nature of the standard, the most appropriate tools from the different parts of the standard may be mixed, within the constraints of the DDL.

The MPEG-7 Audio tools are applicable to two general areas: low-level audio description, in the case of the Audio Framework (clause 5), and application-driven description, in the case of the High Level Tools (clause 6).