DRAFT INTERNATIONAL STANDARD

ISO/IEC DIS 23005-3

ISO/IEC JTC **1**/SC **29** Secretariat: **JISC**

Voting begins on:
2015-05-11

Voting terminates on:
2015-08-11

Information technology — Media context and control —

Part 3:
Sensory information

Technologies de l'information — Contrôle et contexte de supports —

Partie 3: Information sensorielle

ICS: 35.040

Reference number
ISO/IEC DIS 23005-3:2015(E)

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.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword v
Introduction vi
1 **Scope** 1
2 **Normative references** 2
3 **Terms, definitions, and abbreviated terms** 2
3.1 **Terms and definitions** 2
3.2 **Abbreviated terms** 3
3.3 **Schema documents** 3
3.4 **Use of prefixes** 3
4 **Sensory effects description language** 4
4.1 **Introduction** 4
4.2 **Validation** 4
4.3 **Processing** 4
4.4 **Basic building blocks** 5
5 **Sensory effect vocabulary** 37
5.1 **Introduction** 37
5.2 **Validation** 38
5.3 **Schema wrapper** 38
5.4 **Light effect** 38
5.5 **Flash effect** 41
5.6 **Temperature effect** 42
5.7 **Wind effect** 43
5.8 **Vibration effect** 45
5.9 **Spraying effect** 47
5.10 **Scent effect** 49
5.11 **Fog effect** 51
5.12 **Color correction effect** 52
5.13 **Rigid body motion effect** 56
5.14 **Passive kinesthetic motion effect** 77
5.15 **Passive kinesthetic force effect** 79
5.16 **Active kinesthetic effect** 81
5.17 **Tactile effect** 83
5.18 **Parameterized Tactile effect** 87
5.19 **Bubble Effect Type** 95
Annex A (informative) **Intended Usage of Sensory Information** 97

Annex B (informative) **Schema documents** **98**
Annex C (informative) **Patent statements** **99**
Bibliography **100**

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 2.

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.

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.

ISO/IEC 23005-3 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This third edition cancels and replaces the second edition (ISO/IEC 23005-3:2013) which has been technically revised.

ISO/IEC 23005 consists of the following parts, under the general title *Information technology — Media context and control*:

— *Part 1: Architecture*

— *Part 2: Control information*

— *Part 3: Sensory information*

— *Part 4: Virtual world object characteristics*

— *Part 5: Data formats for interaction devices*

— *Part 6: Common types and tools*

— *Part 7: Conformance and reference software*

Introduction

ISO/IEC 23005 (MPEG-V) provides an architecture and specifies associated information representations to enable interoperability between virtual worlds, e.g. digital content provider of a virtual world, (serious) gaming, simulation, DVD, and with the real world, e.g. sensors, actuators, vision and rendering, robotics (e.g. for revalidation), (support for) independent living, social and welfare systems, banking, insurance, travel, real estate, rights management and many others.

Virtual worlds (often referred to as 3D3C for 3D visualization & navigation and the 3C's of Community, Creation and Commerce) integrate existing and emerging (media) technologies (e.g. instant messaging, video, 3D, VR, AI, chat, voice, etc.) that allow for the support of existing, and the development of new kinds of, social networks. The emergence of virtual worlds as platforms for social networking is recognized by businesses as an important issue for at least two reasons:

— It offers the power to reshape the way companies interact with their environments (markets, customers, suppliers, creators, stakeholders, etc.) in a fashion comparable to the Internet.

— It allows for the development of new (breakthrough) business models, services, applications and devices.

Each virtual world however has a different culture and audience making use of these specific worlds for a variety of reasons. These differences in existing metaverses permit users to have unique experiences. Resistance to real-world commercial encroachment still exists in many virtual worlds, where users primarily seek an escape from real life. Hence, marketers should get to know a virtual world beforehand and the rules that govern each individual universe.

Although realistic experiences have been achieved via devices such as 3-D audio/visual devices, it is hard to realize sensory effects only with the presentation of audiovisual contents. The addition of sensory effects leads to even more realistic experiences in the consumption of audiovisual contents. This will lead to the application of new media for enhanced experiences of users in a more realistic sense.

Such new media will benefit from the standardization of control and sensory information which can include sensory effect metadata, sensory device capabilities/commands, user sensory preferences, and various delivery formats. The MPEG-V architecture can be applicable for various business models for which audiovisual contents can be associated with sensory effects that need to be rendered on appropriate sensory devices.

This part of ISO/IEC 23005 contains the sensory information which can stimulate other senses than vision or audition, e.g. olfaction, mechanoreception, equilibrioception, or thermoception. That is, in addition to the audio-visual content of, e.g., a movie, also other sense shall be stimulated giving her/him the sensation of being part of the particular media which shall result in a worthwhile, informative user experience.

The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of patents.

ISO and the IEC take no position concerning the evidence, validity and scope of these patent rights.

The holders of these patent rights have assured ISO and the IEC that they are willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statements of the holders of these patent rights are registered with ISO and the IEC. Information may be obtained from the companies listed in Annex C.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those identified in Annex C. ISO and the IEC shall not be held responsible for identifying any or all such patent rights.

Information technology — Media context and control — Part 3: Sensory information

1 Scope

This Part of ISO/IEC 23005 specifies syntax and semantics of description schemes and descriptors that represent sensory information. This international standard is applicable to enhance the experience of users while consuming media resources by stimulating human multi-sensor such as tactile, orfactory, light sense, temperature sense, etc.

The system architecture is depicted in Figure 1 and the scope of this Part of ISO/IEC 23005 is highlighted. That is, only the information representation that acts as an input to the possible Adaptation VR – as defined in ISO/IEC 23005-1 – is specified in this Part of ISO/IEC 23005.

The adaptation engine for Sensory Information is to adapt Sensory information to Device Command which is the actual signal to control multi-sensory devices as defined in ISO/IEC 23005-5. This adaptation process is not mandatory incase the sensory information may directly control the actual devices.

NOTE 1 The actual Adaptation VR is deliberately informative and left open for industry competition.

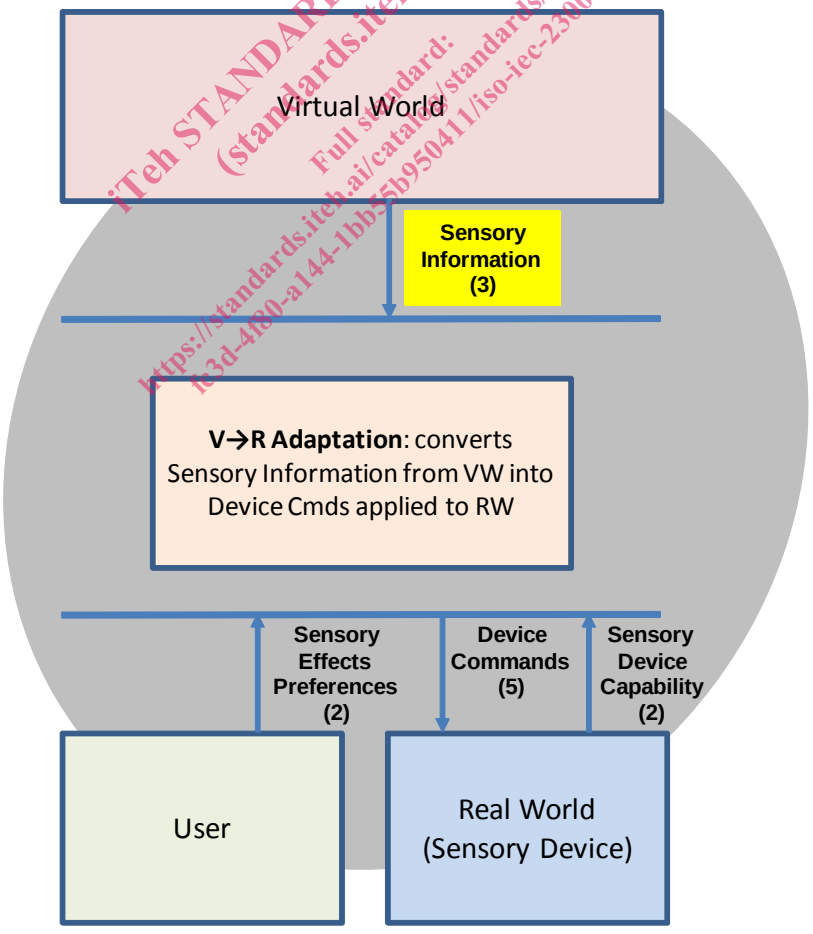

Figure 1 — System Architecture

NOTE 2 Additional informative information can be found in Annex A.

The usage scenarios are described in detail in MPEG-V Architecture (ISO/IEC 23005-1).

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 21000-7, *Information technology — Multimedia framework (MPEG-21) — Part 7: Digital Item Adaptation*

ISO/IEC 23005 (all parts), *Information technology — Media context and control*

W3C XML, *Extensible Markup Language (XML) 1.1, Second Edition*, W3C Recommendation 16 August 2006, edited in place 29 September 2006

W3C XMLSCHEMA, *XML Schema Part 1: Structures and XML Schema Part 2: Datatypes*, Second Edition W3C Recommendation, 28 October 2004

3 Terms, definitions, and abbreviated terms

3.1 Terms and definitions

For the purpose of this document, the terms and definitions in ISO/IEC 23005-6 and the following apply.

3.1.1
digital content provider
entity that acts as the source of digital information of various nature

NOTE The digital content may be provided in real-time or non real-time.

EXAMPLE Digital content from an on-line virtual world, simulation environment, multi user game, a broadcasted multimedia production, a peer-to-peer multimedia production, or packaged content like a DVD or game.

3.1.2
sensory information
standardized representation format of ISO/IEC 23005 in the standardization area B as defined in ISO/IEC 23005-1

EXAMPLE Sensory effect metadata, haptic (kinesthetic/tactile) information, emotion information, avatar information.

3.1.3
sensory effect metadata
defines the description schemes and descriptors to represent **sensory effects**

3.1.4
sensory effect
effect to augment perception by stimulating human senses in a particular scene of a multimedia application

EXAMPLE Scent, wind, light, haptic(kinesthetic-force, stiffness, weight, friction, texture, widget (button, slider, joystick), tactile: air-jet, suction pressure, thermal, current, vibration, note: combinations of tactile display may provide also directional, shape information).

3.1.5
adaptation VR
entity that can process the **sensory information** in order to be consumed within the real world's context

NOTE This may include the adaptation or transformation of the sensory information according to the capabilities of real world devices or the preferences of the user. A specification of these capabilities and preferences can be found in ISO/IEC 23005-2.

3.2 Abbreviated terms

For the purpose of this document, the abbreviated terms given in the following apply:

DIA	digital item adaptation (ISO/IEC 21000-7)
MPEG-21	multimedia framework (ISO/EC 21000)
MPEG-7	multimedia content description interface (ISO/IEC 15938)
SEDL	sensory effects description language
SEM	sensory effect metadata
SEV	sensory effects vocabulary
UMA	universal multimedia access
UME	universal multimedia experience
XML	extensible mark-up language
XSI	XML streaming instructions

3.3 Schema documents

In the main text of this specification, the syntax of description schemes and descriptors is provided whenever possible as a single schema document.

In some cases though, and in particular for Clause 5, the syntax of description schemes and descriptors is provided as a collection of schema snippets imbricated with other text. In order to form a valid schema document, these schema components should be gathered in a same document with the schema wrapper provided at the head of the clause. For better readability, the relevant schema documents are provided in Annex B, but as non-normative information.

In all cases, each schema document has a `version` attribute, the value of which is "ISO/IEC 23005-3". Furthermore, an informative identifier is given as the value of the `id` attribute of the `schema` component. This identifier is non-normative and used as a convention in this specification to reference another schema document. In particular, it is used for the `schemaLocation` attribute of the `include` and `import` schema components.

3.4 Use of prefixes

For clarity, throughout this Part of ISO/IEC 23005, consistent namespace prefixes are used.

"`xsi:`" prefix is not normative. It is a naming convention in this document to refer to an element of the `http://www.w3.org/2001/XMLSchema-instance` namespace.

"`xml:`" and "`xmlns:`" are normative prefixes defined in [1]. The prefix "`xml:`" is by definition bound to "`http://www.w3.org/XML/1998/namespace`". The prefix "`xmlns:`" is used only for namespace bindings and is not itself bound to any namespace name.

All other prefixes used in either the text or examples of this specification are not normative, e.g., "`sedl:`", "`sev:`", "`dia:`", "`si:`", "`mpeg7:`".

In particular, most of the informative examples in this specification are provided as XML fragments without the normally required XML document declaration and, thus, miss a correct namespace binding context declaration. In these descriptions fragments the different prefixes are bound to the namespaces as given in the following table.

Table 1 — Mapping of prefixes to namespaces in examples and text

Prefix	Corresponding namespace
ct	urn:mpeg:mpeg-v:2014:01-CT-NS
sedl	urn:mpeg:mpeg-v:2014:01-SEDL-NS
sev	urn:mpeg:mpeg-v:2014:01-SEV-NS
dia	urn:mpeg:mpeg21:2003:01-DIA-NS
si	urn:mpeg:mpeg21:2003:01-DIA-XSI-NS
mpeg7	urn:mpeg:mpeg7:schema:2004
xsi	http://www.w3.org/2001/XMLSchema-instance
xsd	http://www.w3.org/2001/XMLSchema

Unlike the informative descriptions examples, the normative specification of the syntax of tools in XML Schema follows the namespace binding context defined in the relevant schema declaration such as the one defined in 5.3.

4 Sensory effects description language

4.1 Introduction

This Clause specifies the syntax and semantics of the sensory effects description language (SEDL) which provides basic building blocks for the authoring of sensory effect metadata.

4.2 Validation

Validating a document against the SEDL schema (as specified in W3C XMLSCHEMA) is necessary, but not sufficient, to determine its validity with respect to SEDL. After a document is validated against the SEDL schema, it shall also be subjected to additional validation rules. These additional rules are given below in the descriptions of the elements to which they pertain.

4.3 Processing

The processing model for the sensory effect metadata is defined as an XML processor (as specified by W3C XML) and the utilization of the elements and attributes as defined in the subsequent (Sub)clauses.

NOTE The processing of the sensory effect metadata may follow existing XML decoding/parsing models such as the Document Object Model (DOM) or the Simple API for XML (SAX).

The time information that may be associated to sensory effects may be used for the synchronization with respect to other media assets.

EXAMPLE These other media assets may be video and/or audio.

4.4 Basic building blocks

4.4.1 Introduction

This Subclause specifies the syntax and semantics of the basic building blocks for authoring sensory effect metadata.

This Part of ISO/IEC 23005 adopts the XML streaming instructions (XSI) as defined in ISO/IEC 21000-7 for the purpose of identifying process units and associating time information to them. In this context, a process unit is defined as a well-formed fragment of XML-based metadata that can be consumed as such and to which time information may be attached, indicating the point in time when it becomes available for consumption. A process unit is specified by one element named anchor element and by a process unit mode indicating how other connected elements are aggregated to this anchor to compose the process unit. Depending on the mode, the anchor element is not necessarily the root of the process unit. Anchor elements are ordered according to the navigation path of the XML document. Process units may overlap, i.e. some elements (including anchor elements) may belong to several process units. Additionally, the content provider may require that a given process unit be encoded as a random access point, i.e. that the resulting access unit does not require any other access units to be decoded. The syntax and semantics of the XML streaming instructions is fully specified in 8.6 of ISO/IEC 21000-7:2007.

In addition to the XML streaming instructions, this standard adopts the following basic time model for sensory effects metadata which is depicted in Figure 2.

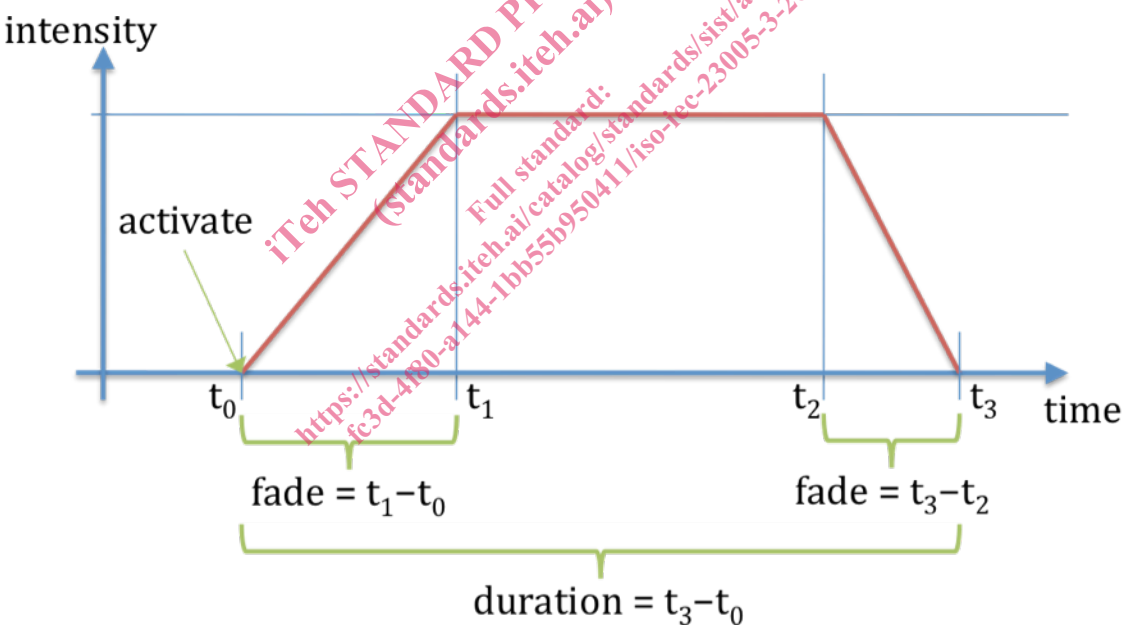

Figure 2 — Time model for sensory effect metadata

Each effect may be activated (i.e., t_0) and deactivated (i.e., t_3) at certain points in time. The deactivation of an effect may be explicitly defined (i.e., `activate="false"`) or indicated by means of a duration attribute during activation (i.e., $t_3 - t_0$). Furthermore, each effect may specify a fade-in (i.e., $t_1 - t_0$) or fade-out (i.e., $t_3 - t_2$) time within which the corresponding effect shall reach its specified intensity.

NOTE The actual implementation of some effects may require one or more elements as defined in the following. An example implementation of Figure 2 using the syntax as defined in the following is provided in 4.4.14.