
**Information technology — Rich media
user interfaces —**

**Part 2:
Advanced user interaction (AUI)
interfaces**

*Technologies de l'information — Interfaces d'utilisateur au support
riche —*

Partie 2: Interfaces d'interaction d'utilisateur avancé (AUI)



Reference number
ISO/IEC 23007-2:2012(E)

AVA Documents
Document Sharing Hub
© ISO/IEC 2012



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2012

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.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword	v
Introduction.....	vi
1 Scope	1
2 Normative references	1
3 Terms, definitions, and abbreviated terms	1
3.1 Terms and definitions	1
3.2 Abbreviated terms	2
4 Overview.....	2
4.1 Introduction.....	2
4.2 Advanced user interaction devices	3
5 Interactivity patterns	4
5.1 Common types.....	4
5.1.1 Schema wrapper conventions	5
5.1.2 Basic data types	5
5.1.3 Reference coordinate system	7
5.2 Geometric pattern.....	7
5.2.1 Introduction.....	7
5.2.2 Syntax.....	8
5.2.3 Semantics.....	9
5.2.4 Example	11
5.3 Symbolic pattern	12
5.3.1 Introduction.....	12
5.3.2 Syntax	12
5.3.3 Semantics.....	13
5.3.4 SymbolTypeCS	13
5.3.5 Example	14
5.4 Touch pattern.....	14
5.4.1 Introduction.....	14
5.4.2 Syntax	14
5.4.3 Semantics.....	15
5.4.4 TouchTypeCS	16
5.4.5 Example	17
5.5 Hand posture pattern	17
5.5.1 Introduction.....	17
5.5.2 Syntax	17
5.5.3 Semantics.....	18
5.5.4 HandPostureTypeCS.....	18
5.5.5 Example	19
5.6 Hand gesture pattern	19
5.6.1 Introduction.....	19
5.6.2 Syntax	20
5.6.3 Semantics.....	20
5.6.4 HandGestureCS	20
5.6.5 Examples	22
Annex A (informative) Relationship between MPEG-U and MPEG-V.....	23
Annex B (informative) Predefined Message Interfaces of AUI for Widget Manager	25
Annex C (informative) IDL interface definition of AUI	28

Annex D (informative) Patent statements 32

Bibliography 33

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.

ISO/IEC 23007-2 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 23007 consists of the following parts, under the general title *Information technology — Rich media user interfaces*:

- *Part 1: Widgets*
- *Part 2: Advanced user interaction (AUI) interfaces*
- *Part 3: Conformance and reference software*

Introduction

R&D activities regarding advanced user interaction devices and mechanisms (often referred as HCI) such as motion and voice recognitions were used to reside within academia and lab environments. However, user interaction devices of industry have evolved dramatically in recent years with the maturing technologies. Consequently various advanced interaction devices, such as multi-touch pad, g-sensor, etc., are already incorporated into consumer electronics and offered in the market.

Although multimedia technologies have been studied to provide mature applicable technologies; however, current user interaction standards mostly focus on basic interaction devices such as pointing and keying devices. This lack of support results in unavailability to utilize such industry ready advanced interaction devices in a standard way.

This part of ISO/IEC 23007 presents the list of new data formats for scene specific advanced user interaction devices. It also includes the high-level view of the relationship between semantically recognized information and physical sensed information. Moreover it provides other interface formats of advanced user interaction devices in order to show the applicability of this International Standard with the existing standard.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this part of ISO/IEC 23007 may involve the use of patents.

The ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured the ISO and IEC that he is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with the ISO and IEC. Information may be obtained from the companies listed in Annex D.

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 D. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO (www.iso.org/patents) and IEC (<http://patents.iec.ch>) maintain on-line databases of patents relevant to their standards. Users are encouraged to consult the databases for the most up to date information concerning patents.

Information technology — Rich media user interfaces —

Part 2: Advanced user interaction (AUI) interfaces

1 Scope

This part of ISO/IEC 23007 specifies advanced user interaction interfaces (AUI) to support various advanced user interaction devices. The AUI interface is a part of the bridge between scene descriptions and system resources. A scene description is a self-contained living entity composed of video, audio, 2D graphics objects, and animations. Through the AUI interfaces or other existing interfaces such as DOM events, a scene description accesses interesting system resources to interact with users. In general, a scene composition is conducted by a third party and remotely deployed.

2 Normative references

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

W3C DOM Events, "Document Object Model Events", available at <http://www.w3.org/TR/DOM-Level-2-Events/events.html>

W3C XML, *Extensible Markup Language 1.0 (Second Edition)*, W3C Recommendation, 6 October 2000

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

3 Terms, definitions, and abbreviated terms

3.1 Terms and definitions

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

3.1.1

advanced user interaction

interactions that utilize more advanced interaction devices

EXAMPLE Multi-pointing user interaction, Motion sensing, Gaze tracking

3.1.2

sensor

combination of hardware and software or just an instance of software that detects features from real world

3.1.3

presentation engine

entity processing the Scene Description to provide its animated and interactive behaviour through composition and rendering

3.1.4

scene description

description defining an audiovisual and graphics presentation in terms of spatiotemporal layout, and interactions by using text, graphics, animations, images, videos, sounds, etc

3.1.5

widget

self-contained entity, with extensive communication capabilities, within a Rich Media User Interface; composed of a Manifest and associated Resources, including Scene Descriptions for the Full and Simplified Representations and Context Information

3.2 Abbreviated terms

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

BIFS	Binary Format for Scene
LASeR	Lightweight Application Scene Representation.
SVG	Scalable Vector Graphics
XML	eXtensible Markup Language
AUI	Advanced User Interaction
MPEG-U	ISO/IEC 23007
MPEG-V	ISO/IEC 23005

4 Overview

4.1 Introduction

This Subclause introduces the AUI patterns which this part of ISO/IEC 23007 provides as below.

1. The geometric patterns are a set of geometric shapes which are recognized with sensed geometric information as 2D or 3D Cartesian positions. Currently, the standard is trying to define the following geometric interactivity patterns: Point, Line, Rect., Arc and Circle patterns.
2. Instead of speaking or writing a word, simple well-known gestures help to communicate with others. For instance, V sign and Rock sign, which are well-known semantics, are already used in various situations. Therefore, the standard provides a container format for symbolic patterns and a classification scheme to enlist well-known symbolic shapes.
3. Many applications adopt well-known touch patterns for users to control themselves. In this part of ISO/IEC 23007, a container format for well-known touch patterns and a classification scheme to enlist basic touch patterns are provided.
4. This part of ISO/IEC 23007 describes the hand posture and gesture patterns to support the intuitive hand based interaction for scene description. For example, if a user wants to control an object in a scene description, the user in the real life makes a hand gesture such as grabbing, fist, and open palm which would be a good candidate gesture to support such an interaction modality. In addition, when the user wants to interact with people using hand gestures, "waving a hand" gesture is common semantic to all the human beings, even though the contextual meaning of the gesture may vary.

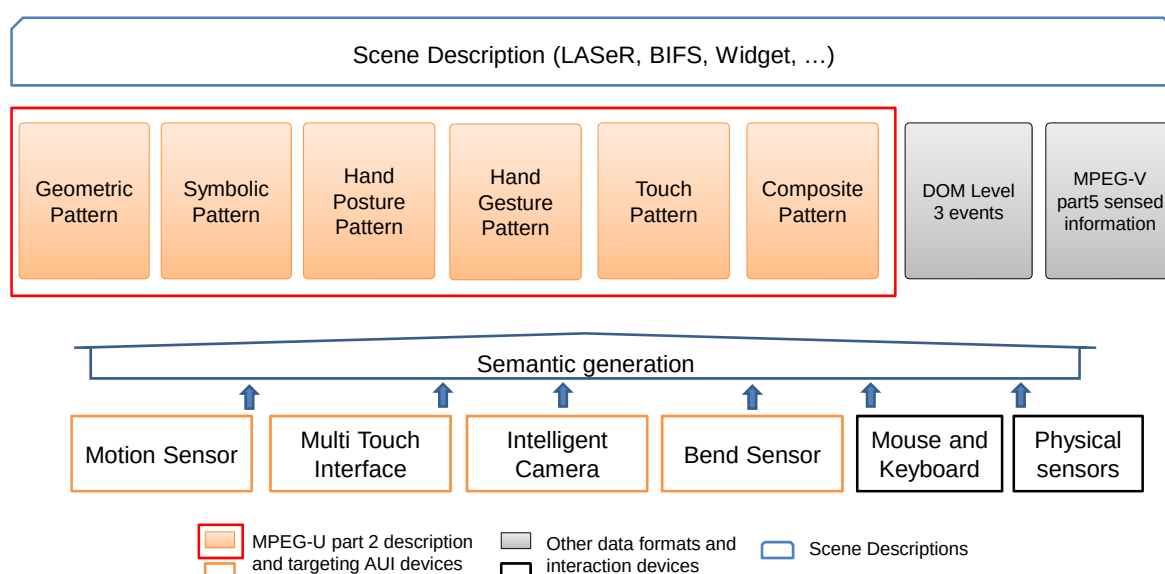


Figure 1 — Architecture of MPEG-U part 2

Figure 1 presents the overall architecture of this part of ISO/IEC 23007. User interaction devices such as motion sensors and multi touch interfaces generate the physical sensed information from user's environment. According to the use cases which are mentioned in the context and objectives document for this part of ISO/IEC 23007, a set of physical information can generate a semantic meaning by a recognition process. For instance, some feature points drawn by user's finger can be understood as a circle which is specified with a circle center position and a radius value. Therefore, this part of ISO/IEC 23007 provides a set of data formats which describes geometric patterns, symbolic patterns, touch patterns, hand posture patterns and their composite pattern.

At the right side of the above figure, other standard activities are also providing similar data structures for new user interaction interfaces such as physical sensors and even new features of mouse and keyboard. Therefore, scene description can use not only the data formats of this part of ISO/IEC 23007 and existing data sets from others.

4.2 Advanced user interaction devices

Recently various new types of user interaction devices are available on the market such as multi-touch pad, motion sensing remote controller, etc. Figure 2 shows a few available advanced user interaction devices and Figure 3 shows multi-touch devices currently available on the market.



Figure 2 — Motion sensing remote controllers



Figure 3 — Multi-touch devices

Among various possible Advanced User Interaction (AUI) Interfaces, multi-touch interface is the first recognized interaction interface to be ready both in terms of device availability on the market and in terms of applicability to various MPEG scene representation standards.

Almost unlimited number of actions can be mapped with the combination of 1) the number of fingers touched, 2) the moving direction of each touched finger, 3) the amount of displacement of each finger, etc. Therefore, multi-touch interface should be defined by carefully extracting necessary core attributes and make the structure expandable.



Figure 4 — Hand based 3D interaction

Another example among possible Advanced User Interaction (AUI) Interfaces is related to hand based interface shown in Figure 4. If a user wants to select an object and move it to a certain place in the scene, the user in the real life moves his/her hand toward the object, and then move the object while holding and grabbing the object. The user finally releases the object by opening the hand. That is, hand postures such as grabbing, fist, open palm, pointing, thumb-up, and thumb-down would be a good candidate posture to support such an interaction modality. The semantics of the postures shall be selected as the descriptions of the shapes, which are independent upon the cultural difference.

5 Interactivity patterns

5.1 Common types

This Clause describes common types including the schema wrapper conventions, basic data types. The types defined in this Clause are defined to be used in this part of ISO/IEC 23007, and are not intended to be instantiated again.

5.1.1 Schema wrapper conventions

The syntax defined in this subclause assumes the following Schema Wrapper to form a valid XML schema document.

```
<schema xmlns="http://www.w3.org/2001/XMLSchema"
  xmlns:aui="urn:mpeg:mpegu:schema:aui:2012"
  xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
  targetNamespace="urn:mpeg:mpegu:schema:aui:2012"
  elementFormDefault="qualified"
  attributeFormDefault="unqualified" version="ISO/IEC 23007-2"
  id="MPEG-U-AUI.xsd">
  <import namespace="urn:mpeg:mpeg7:schema:2004" schemaLocation="mp74U.xsd"/>
```

Additionally, the following line should be appended to the resulting schema document in order to obtain a well-formed XML document.

```
</schema>
```

5.1.2 Basic data types

This Subclause describes structure of the basic data types which are commonly used in this part of ISO/IEC 23007 as a basic building block of the tools.

5.1.2.1 Syntax

```
<!-- ##### -->
<!-- Vector Datatype -->
<!-- ##### -->
<complexType name="VectorType">
  <sequence>
    <element name="X" type="float"/>
    <element name="Y" type="float"/>
    <element name="Z" type="float" minOccurs="0"/>
  </sequence>
</complexType>
<!-- ##### -->
<!-- Chirality Datatype -->
<!-- ##### -->
<simpleType name="ChiralityType">
  <restriction base="string">
    <enumeration value="Right"/>
    <enumeration value="Left"/>
  </restriction>
</simpleType>
<!-- ##### -->
<!-- AUI Base Datatype -->
<!-- ##### -->
<complexType name="AUIBaseType">
  <attribute name="capturedTimeStamp" type="unsignedLong" default="0"
  use="optional"/>
  <attribute name="userId" type="ID" use="optional"/>
</complexType>
<!-- ##### -->
<!-- Composite Pattern -->
<!-- ##### -->
```

```

<element name="CompositePattern" type="aui:CompositePatternType"/>
<complexType name="CompositePatternType">
  <complexContent>
    <extension base="aui:AUIBaseType">
      <sequence>
        <element name="AUIPattern" type="aui:AUIBaseType"
maxOccurs="unbounded"/>
      </sequence>
      <attribute name="sameObject" type="boolean" default="false"
use="optional"/>
    </extension>
  </complexContent>
</complexType>

```

5.1.2.2 Semantics

Semantics of the basic datatypes :

Name	Definition
VectorType	This type describes the vector type composed of two float values and one optional value to represent a set of values.
X	A Value that describes a float value(can be force, torque, position) for x-axis
Y	A Value that describes a float value(can be force, torque, position) for y-axis
Z	A Value that describes a float value(can be force, torque, position) for z-axis
ChiralityType	This type describes whether the hand of interest is a left hand or a right hand. The value of "Right" describes that the hand is a right hand and the value of "Left" describes that the hand is a left hand.
AUIBaseType	This type provides the topmost type of the base type hierarchy which each individual AUI data formats can inherit.
capturedTimeStamp	This attribute specifies the time (in milliseconds relative to the epoch) at which a user interaction was captured. When the value of capturedTimeStamp is not available, a value of 0 will be returned. Examples of epoch time are the time of the system start or 0:0:0 UTC 1st January 1970. The capturedTimeStamp shall be used as defined in W3C DOM Events.
userId	This attribute describes an index referencing the user who is generating AUI patterns.
CompositePatternType	This type provides the container type as the placeholder for the set of AUI Patterns. Since users may generate more than one AUI patterns simultaneously, this type helps to transfer that set of patterns.