

---

---

**Information technology — Coding of  
audio-visual objects —**

**Part 1:  
Systems**

*Technologies de l'information — Codage des objets audiovisuels —  
Partie 1: Systèmes*

**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.



**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2010

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](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

<b>Foreword</b> .....	<b>iv</b>
<b>0 Introduction</b> .....	<b>vi</b>
<b>1 Scope</b> .....	<b>1</b>
<b>2 Normative references</b> .....	<b>1</b>
<b>3 Additional references</b> .....	<b>2</b>
<b>4 Terms and definitions</b> .....	<b>2</b>
<b>5 Abbreviated terms</b> .....	<b>10</b>
<b>6 Conventions</b> .....	<b>11</b>
<b>7 Streaming Framework</b> .....	<b>11</b>
<b>8 Syntactic Description Language</b> .....	<b>99</b>
<b>9 Profiles</b> .....	<b>110</b>
<b>Annex A</b> (informative) <b>Time Base Reconstruction</b> .....	<b>112</b>
<b>Annex B</b> (informative) <b>Registration procedure</b> .....	<b>115</b>
<b>Annex C</b> (informative) <b>The QoS Management Model for ISO/IEC 14496 Content</b> .....	<b>119</b>
<b>Annex D</b> (informative) <b>Conversion Between Time and Date Conventions</b> .....	<b>120</b>
<b>Annex E</b> (informative) <b>Graphical Representation of Object Descriptor and Sync Layer Syntax</b> .....	<b>122</b>
<b>Annex F</b> (informative) <b>Elementary Stream Interface</b> .....	<b>130</b>
<b>Annex G</b> (informative) <b>Upstream Walkthrough</b> .....	<b>132</b>
<b>Annex H</b> (informative) <b>Scene and Object Description Carousel</b> .....	<b>137</b>
<b>Annex I</b> (normative) <b>Usage of ITU-T Recommendation H.264   ISO/IEC 14496-10 AVC</b> .....	<b>138</b>
<b>Annex J</b> (informative) <b>Patent statements</b> .....	<b>141</b>
<b>Bibliography</b> .....	<b>144</b>

## 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 14496-1 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This fourth edition cancels and replaces the third edition (ISO/IEC 14496-1:2004), which has been technically revised. It also incorporates the Amendments ISO/IEC 14496-1:2004/Amd.1:2005, ISO/IEC 14496-1:2004/Amd.2:2007, ISO/IEC 14496-1:2004/Amd.3:2007 and Technical Corrigenda ISO/IEC 14496-1:2004/Cor.1:2006 and ISO/IEC 14496-1:2004/Cor.2:2007.

ISO/IEC 14496 consists of the following parts, under the general title *Information technology — Coding of audio-visual objects*:

- *Part 1: Systems*
- *Part 2: Visual*
- *Part 3: Audio*
- *Part 4: Conformance testing*
- *Part 5: Reference software*
- *Part 6: Delivery Multimedia Integration Framework (DMIF)*
- *Part 7: Optimized reference software for coding of audio-visual objects*
- *Part 8: Carriage of ISO/IEC 14496 contents over IP networks*
- *Part 9: Reference hardware description*
- *Part 10: Advanced Video Coding*
- *Part 11: Scene description and application engine*
- *Part 12: ISO base media file format*
- *Part 13: Intellectual Property Management and Protection (IPMP) extensions*

- *Part 14: MP4 file format*
- *Part 15: Advanced Video Coding (AVC) file format*
- *Part 16: Animation Framework eXtension (AFX)*
- *Part 17: Streaming text format*
- *Part 18: Font compression and streaming*
- *Part 19: Synthesized texture stream*
- *Part 20: Lightweight Application Scene Representation (LSeR) and Simple Aggregation Format (SAF)*
- *Part 21: MPEG-J Graphics Framework eXtensions (GFX)*
- *Part 22: Open Font Format*
- *Part 23: Symbolic Music Representation*
- *Part 24: Audio and systems interaction*
- *Part 25: 3D Graphics Compression Model*
- *Part 26: Audio conformance*
- *Part 27: 3D Graphics conformance*

## 0 Introduction

### 0.1 Overview

ISO/IEC 14496 specifies a system for the communication of interactive audio-visual scenes. This specification includes the following elements.

- a) The coded representation of natural or synthetic, two-dimensional (2D) or three-dimensional (3D) objects that can be manifested audibly and/or visually (audio-visual objects) (specified in Parts 2, 3, 10, 11, 16, 19, 20, 23 and 25 of ISO/IEC 14496).
- b) The coded representation of the spatio-temporal positioning of audio-visual objects as well as their behavior in response to interaction (scene description, specified in Parts 11 and 20 of ISO/IEC 14496).
- c) The coded representation of information related to the management of data streams (synchronization, identification, description and association of stream content, specified in this Part and in Part 24 of ISO/IEC 14496).
- d) A generic interface to the data stream delivery layer functionality (specified in Part 6 of ISO/IEC 14496).
- e) An application engine for programmatic control of the player: format, delivery of downloadable Java byte code as well as its execution lifecycle and behavior through APIs (specified in Parts 11 and 21 of ISO/IEC 14496).
- f) A file format to contain the media information of an ISO/IEC 14496 presentation in a flexible, extensible format to facilitate interchange, management, editing, and presentation of the media specified in Part 12 (ISO File Format), Part 14 (MP4 File Format) and Part 15 (AVC File Format) of ISO/IEC 14496.
- g) The coded representation of font data and of information related to the management of text streams and font data streams (specified in Parts 17, 18 and 22 of ISO/IEC 14496).

The overall operation of a system communicating audio-visual scenes can be paraphrased as follows:

At the sending terminal, the audio-visual scene information is compressed, supplemented with synchronization information and passed to a delivery layer that multiplexes it into one or more coded binary streams that are transmitted or stored. At the receiving terminal, these streams are demultiplexed and decompressed. The audio-visual objects are composed according to the scene description and synchronization information and presented to the end user. The end user may have the option to interact with this presentation. Interaction information can be processed locally or transmitted back to the sending terminal. ISO/IEC 14496 defines the syntax and semantics of the bitstreams that convey such scene information, as well as the details of their decoding processes.

This part of ISO/IEC 14496 specifies the following tools.

- A terminal model for time and buffer management.
- A coded representation of metadata for the identification, description and logical dependencies of the elementary streams (object descriptors and other descriptors).
- A coded representation of descriptive audio-visual content information [object content information (OCI)].
- An interface to intellectual property management and protection (IPMP) systems.
- A coded representation of synchronization information (sync layer – SL).
- A multiplexed representation of individual elementary streams in a single stream (M4Mux).

These various elements are described functionally in this clause and specified in the normative clauses that follow.

## 0.2 Architecture

The information representation specified in ISO/IEC 14496 describes the means to create an interactive audio-visual scene in terms of coded audio-visual information and associated scene description information. The entity that composes and sends, or receives and presents such a coded representation of an interactive audio-visual scene is generically referred to as an “audio-visual terminal” or just “terminal”. This terminal may correspond to a stand-alone application or be part of an application system.

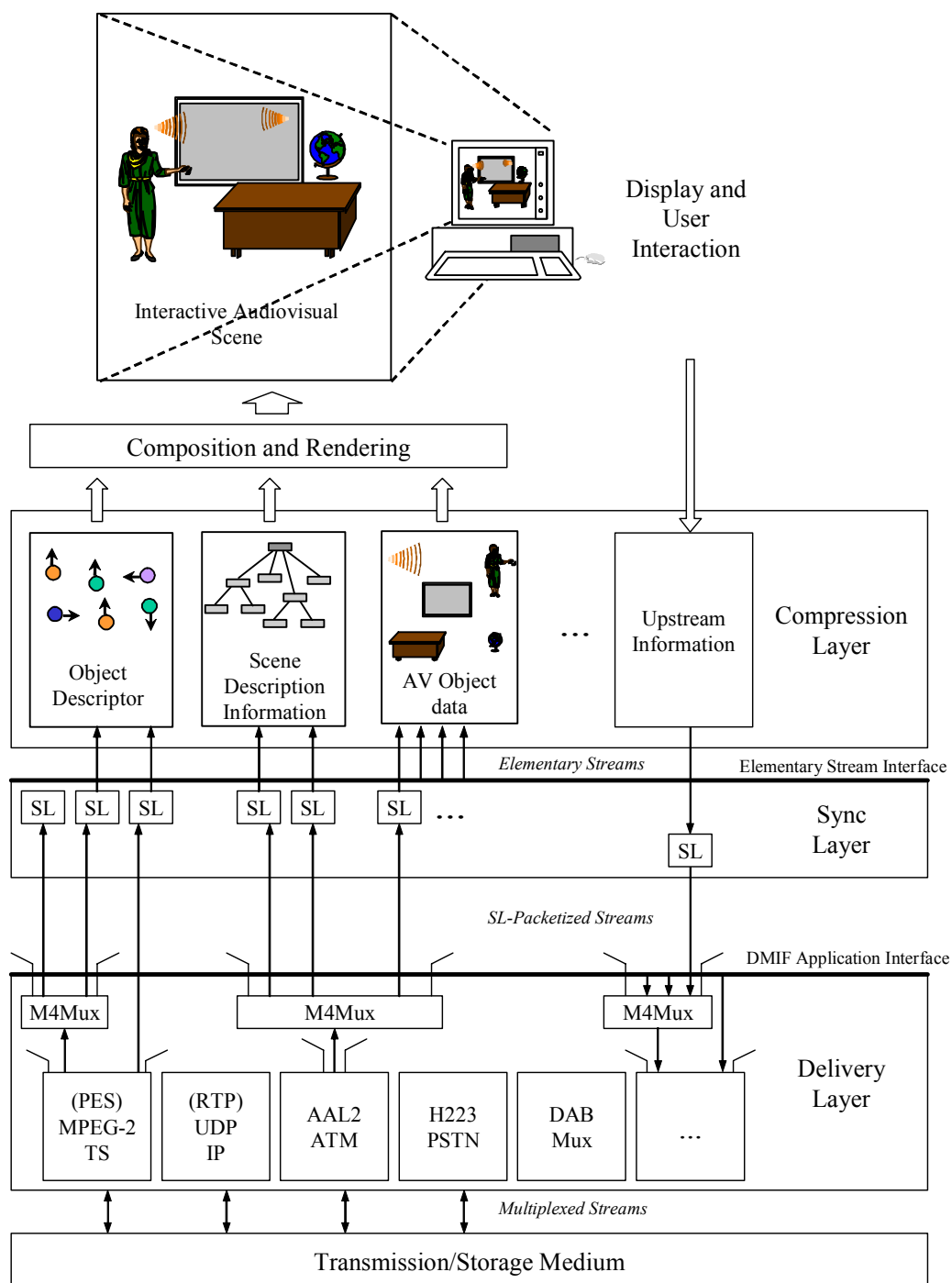


Figure 1 — The ISO/IEC 14496 Terminal Architecture