

First edition  
2008-01-15

---

---

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

Part 24:  
**Audio and systems interaction**

*Technologies de l'information — Codage d'objets audiovisuels —  
Partie 24: Codage audio et interaction de systèmes*

---

---

Reference number  
ISO/IEC TR 14496-24:2008(E)



© ISO/IEC 2008

**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 2008

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

Foreword.....	iv
<b>1</b> <b>Scope</b> .....	<b>1</b>
<b>2</b> <b>Motivating audio composition time stamp handling</b> .....	<b>1</b>
<b>3</b> <b>AAC Encoder/Decoder Behavior</b> .....	<b>2</b>
<b>3.1</b> <b>Example 1: AAC</b> .....	<b>2</b>
<b>3.1.1</b> <b>Overview</b> .....	<b>2</b>
<b>3.1.2</b> <b>Pre-roll</b> .....	<b>4</b>
<b>3.1.3</b> <b>Edit-list</b> .....	<b>4</b>
<b>3.1.4</b> <b>Compressed Information and Decoder behavior</b> .....	<b>4</b>
<b>3.2</b> <b>Example 2: HE-AAC</b> .....	<b>5</b>
<b>3.2.1</b> <b>Overview</b> .....	<b>5</b>
<b>4</b> <b>Streaming Considerations</b> .....	<b>6</b>
<b>Annex A</b> (informative) <b>Relevant ISO Base Media File Format Syntax</b> .....	<b>7</b>
<b>A.1</b> <b>Pre-roll syntax</b> .....	<b>7</b>
<b>A.2</b> <b>Edit-list syntax</b> .....	<b>8</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.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

- type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;
- type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;
- type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard (“state of the art”, for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

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 TR 14496-24, which is a Technical Report of type 3, 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 TR 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 (LAsER) 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*

# Information technology — Coding of audio-visual objects —

## Part 24:

## Audio and systems interaction

### 1 Scope

This part of ISO/IEC TR 14496 describes the desired joint behavior of MPEG-4 Systems (MPEG-4 File Format) and MPEG-4 Audio codecs. It is desired that MPEG-4 Audio encoders and decoders permit finite length signals to be encoded to a file (particularly MPEG-4 files) and decoded again to obtain the identical signal, subject to codec distortions. This will allow the use of audio in systems implementations (particularly MPEG-4 Systems), perhaps with other media such as video, in a deterministic fashion. Most importantly, the decoded signal will have nothing “extra” at the beginning or “missing” at the end.

This permits:

- a) an exact ‘round trip’ from raw audio to encoded file back to raw audio (excepting encoding artifacts);
- b) predictable synchronization between audio and other media such as video;
- c) correct behavior when performing random access as well as when starting at the beginning of a stream;
- d) identical behavior when edits are applied in the raw domain and the encoded domain (again, excepting encoding artifacts).

It is also required that there be predictable interoperability between encoders (as represented by files) and decoders. There are two kinds of audio ‘offsets’ (or ‘delay’ in the context of transmission): those that result from the encoding process, and those that result from the decoding process. This document is primarily concerned with the latter.

These issues are resolved by the following:

- The handling of composition time stamps for audio composition units is specified. Special care is taken in the case of compressed data, like HE-AAC coded audio, that can be decoded in a backward compatible fashion as well as in an enhanced fashion.
- Examples are given that show how finite length signals can be encoded to an MPEG-4 file and decoded again to obtain the identical signal, excepting codec distortions. Most importantly, the decoded signal has nothing “extra” at the beginning or “missing” at the end.