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

Part 20:
**Lightweight Application Scene
Representation (LAsEeR) and Simple
Aggregation Format (SAF)**

Technologies de l'information — Codage des objets audiovisuels —

*Partie 20: Représentation de scène d'application allégée (LAsEeR) et
format d'agrégation simple (SAF)*

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.

This document is a preview generated by EVS



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
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword	v
Introduction.....	vii
1 Scope.....	1
2 Normative references.....	2
3 Terms, definitions and abbreviations.....	3
3.1 Terms and definitions	3
3.2 Abbreviations.....	4
4 Document conventions.....	4
5 Architecture	4
5.1 Overview.....	4
5.2 LAsER systems decoder model.....	5
6 Scene representation	8
6.1 Overview.....	8
6.2 Relationship with SVG	8
6.3 Timing model	10
6.4 Execution model.....	12
6.5 Events	12
6.6 Encoder configuration	16
6.7 LAsER Scene Commands.....	18
6.8 Scene description elements.....	31
6.9 Summary of possible attributes per element.....	47
6.10 Additions to the uDOM API.....	52
7 Simple Aggregation Format (SAF).....	54
7.1 Overview.....	54
7.2 Time and terminal model specification.....	54
7.3 SAF Packet.....	54
7.4 SAF Packet Header.....	57
7.5 SAF Access Unit.....	57
7.6 SimpleDecoderConfigDescriptor.....	59
7.7 SimpleDecoderSpecificInfo	59
7.8 RemoteStreamHeader	60
7.9 Cache unit	60
7.10 EndOfStream.....	61
7.11 EndOfSAFSession.....	61
7.12 GroupingDescriptor	62
7.13 SAF configuration	62
7.14 Stop cache	63
7.15 SAF Extended Access Unit.....	64
7.16 SAF first fragment unit.....	64
7.17 SAF fragment unit	65
8 Profiles.....	65
8.1 Overview.....	65
8.2 LAsER mini	65
8.3 LAsER full.....	68
8.4 LAsER Core.....	73
9 Compatibility of SAF Packet.....	76
10 Carriage of LAsER and SAF	77

10.1	Storage of LAsE _R in MP4 files	77
10.2	Carriage of SAF Streams over HTTP	80
10.3	Carriage of SAF Streams over RTP.....	80
10.4	Carriage of SAF Streams over MPEG-2 Systems	80
11	Electronic attachments	80
12	Binary syntax for the LAsE _R encoding	81
12.1	Decoding process	81
12.2	Binary syntax	86
13	Usage of ISO/IEC 23001-1	150
13.1	Introduction	150
13.2	Electronic attachments	151
13.3	Type codecs	151
13.4	Type codecs for use with ISO/IEC 23001-1 decoders	152
13.5	DecoderInit	155
13.6	Decoding process	156
Annex A (informative) Patent statements		160
Annex B (normative) Media Type Registrations		161
Bibliography		167

This document is a preview generated by EVS

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 14496-20 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 14496-20:2006) which has been technically revised. It also incorporates the Amendment ISO/IEC 14496-20:2006/Amd.1:2008 and the Technical Corrigenda ISO/IEC 14496-20:2006/Cor.1:2007 and ISO/IEC 14496-20:2006/Cor.2:2008.

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* [Technical Report]
- *Part 8: Carriage of ISO/IEC 14496 contents over IP networks*
- *Part 9: Reference hardware description* [Technical Report]
- *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 [Technical Report]*
- *Part 25: 3D Graphics Compression model*

This document is a preview generated by EVS

Introduction

ISO/IEC 14496-20 specifies syntax and semantics for the following:

- The Lightweight Application Scene Representation (LAsER), specified in Clause 6, which is a binary format for encoding 2D scenes and updates of scenes. The binary format and the scene representation (based on SVC Tiny) are both designed to be suitable for lightweight embedded devices such as mobile phones.
- A Simple Aggregation Format (SAF), specified in Clause 7, to efficiently and easily transport LAsER data together with audio and/or video content over various delivery channels. This multiplexing scheme is designed to be simple to implement and to allow efficient demultiplexing on low-end devices.

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

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

The holder of this patent right has assured 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 ISO and IEC. Information may be obtained from the companies listed in Annex A.

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

This document is a preview generated by EVS

Information technology — Coding of audio-visual objects —

Part 20:

Lightweight Application Scene Representation (LAsER) and Simple Aggregation Format (SAF)

1 Scope

This part of ISO/IEC 14496 defines a scene description format (LAsER) and an aggregation format (SAF) respectively suitable for representing and delivering rich-media services to resource-constrained devices such as mobile phones.

LAsER aims at fulfilling all the requirements of rich-media services at the scene description level. LAsER supports

- an optimized set of objects inherited from SVG to describe rich-media scenes,
- a small set of key compatible extensions over SVG,
- the ability to encode and transmit a LAsER stream and then reconstruct SVG content,
- dynamic updating of the scene to achieve a reactive, smooth and continuous service,
- simple yet efficient compression to improve delivery and parsing times, as well as storage size, one of the design goals being to allow both for a direct implementation of the SDL as documented, as well as for a decoder compliant with ISO/IEC 23001-1 to decode the LAsER bitstream,
- an efficient interface with audio and visual streams with frame-accurate synchronization,
- use of any font format, including the OpenType industry standard, and
- easy conversion from other popular rich-media formats in order to leverage existing content and developer communities.

Technology selection criteria for LAsER included compression efficiency, but also code and memory footprint and performance. Other aims included: scalability, adaptability to the user context, extensibility of the format, ability to define small profiles, feasibility of a J2ME implementation, error resilience and safety of implementations.

SAF aims at fulfilling all the requirements of rich-media services at the interface between media/scene description and existing transport protocols:

- simple aggregation of any type of stream,
- signalling of MPEG and non-MPEG streams,
- optimized packet headers for bandwidth-limited networks,
- easy mapping to popular streaming formats,

- cache management capability, and
- extensibility.

SAF has been designed to complement LAsER for simple, interactive services, bringing

- efficient and dynamic packaging to cope with high latency networks,
- media interleaving, and
- synchronization support with a very low overhead.

This part of ISO/IEC 14496 defines the usage of SAF for LAsER content; however, LAsER can be used independently from SAF.

2 Normative references

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

- ISO/IEC 9899, *Programming languages — C*
- ISO/IEC 13818-1, *Information technology — Generic coding of moving pictures and associated audio information: Systems*
- ISO/IEC 14496-1, *Information technology — Coding of audio-visual objects — Part 1: Systems*
- ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*
- ISO/IEC 14496-18, *Information technology — Coding of audio-visual objects — Part 18: Font compression and streaming*
- ISO/IEC 14882:2003, *Programming languages — C++*
- ISO/IEC 16262:2002, *Information technology — ECMAScript language specification*
- IETF BCP 13, *RFC 4288 on Media Type Specifications and Registration Procedures*, <http://www.ietf.org/rfc/rfc4288.txt>
- IETF RFC 3023, *XML Media Types*, M. Murata, S. St. Laurent, D. Kohn, January 2001, <http://www.ietf.org/rfc/rfc3023.txt>
- IETF RFC 3986, *Uniform Resource Identifiers (URI): Generic Syntax*, T. Berners-Lee, R. Fielding, L. Masinter, January 2005, <http://www.ietf.org/rfc/rfc3986.txt>
- IETF RFC 2045, *MIME formats and encodings*, <http://www.ietf.org/rfc/rfc2045.txt>
- IETF RFC 2326, *Real Time Streaming Protocol*, <http://www.ietf.org/rfc/rfc2326.txt>
- IETF RFC 2965, *HTTP State Management Mechanism*, Kristol and Montulli, <http://www.ietf.org/rfc/rfc2965.txt>
- W3C SVG11, *Scalable Vector Graphics (SVG) 1.1 Specification [Recommendation]*, <http://www.w3.org/TR/2003/REC-SVG11-20030114>

W3C SMIL2, *Synchronized Multimedia Integration Language (SMIL 2.0) [Second Edition]*, J. Ayars, D. Bulterman et. al., 7 January 2005, <http://www.w3.org/TR/2005/REC-SMIL2-20050107>

W3C CSS, *Cascading Style Sheets, level 2 [Recommendation]*,
<http://www.w3.org/TR/1998/REC-CSS2-19980512>

W3C DOM, *Document Object Model Level 2 Events Specification, Version 1.0, W3C Recommendation*, 13 November, 2000, <http://www.w3.org/TR/2000/REC-DOM-Level-2-Events-20001113>

W3C XML Events, *an Events Syntax for XML, W3C Recommendation*, 14 October 2003, <http://www.w3.org/TR/2003/REC-xml-events-20031014>

W3C xml:id, *Version 1.0, W3C Recommendation*, 9 September 2005, <http://www.w3.org/TR/2005/PR-xml-id-20050712>

W3C XLink, *XML Linking Language, W3C Recommendation*, 27 June 2001, <http://www.w3.org/TR/2001/REC-xlink-20010627>

3 Terms, definitions and abbreviations

3.1 Terms and definitions

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

3.1.1

access unit

individually accessible portion of data within a media stream

NOTE An access unit is the smallest data entity to which timing information can be attributed.

3.1.2

media time line

axis on which times are expressed within the transport or system carrying a LAsER or other stream

3.1.3

normal play time

indicates the stream absolute position relative to the beginning of the presentation

[IETF RFC 2326]

3.1.4

packet

smallest data entity managed by SAF consisting of a header and a payload

3.1.5

scene segment

set of access units of a LAsER stream, where only the first access unit contains a LAsERHeader

3.1.6

scene time line

axis on which times are expressed within the SVG/LAsER scene, e.g. begin and end

3.1.7

waiting tree

separate tree defined in addition to the scene tree

NOTE The compositor and renderer have no knowledge of the waiting tree, thus objects in the waiting tree are neither composited nor rendered.