
**Information technology — Multimedia
application format (MPEG-A) —**

**Part 4:
Musical slide show application format**

*Technologies de l'information — Format pour application multimédia
(MPEG-A) —*

*Partie 4: Format pour application de présentation musicale de
diapositives*

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 2009

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 and definitions	2
4 Overview of MPEG Standards for Musical slide show application format.....	2
4.1 MPEG-1 Layer 3	2
4.2 ISO Base Media File Format:	3
4.3 The ISO Base Media and MPEG-4 File Formats	3
4.4 MPEG-4 “MPEG-1/2 Audio in MPEG-4”	3
4.5 MPEG-4 LAsER	4
4.6 MPEG-7 Multimedia Description Scheme	4
5 File structure of Musical slide show application format	4
5.1 General	4
5.2 Components of Musical slide show application format	4
5.3 File format	5
5.4 Playback	8
6 Technical features of Musical slide show application format	9
6.1 General	9
6.2 Synchronization of Media	10
6.3 Animation	11
6.4 Timed text.....	12
6.5 Metadata	13
7 Scope of Protected Musical slide show application format.....	30
7.1 Overview.....	30
7.2 Creating Protected Musical slide show application format	31
7.3 Metadata for Protection	32
7.4 Playback	36
8 Overview of Basic Standards for Protection	36
8.1 MPEG-21 IPMP Components Base Profile	36
8.2 MPEG-21 REL MAM Profile.....	37
9 Usage of File Format Brands.....	37
10 Conformance and List of Technologies	37
Annex A (informative) Use cases of Musical slide show application format.....	38
A.1 General	38
A.2 Personal slide show application	38
A.3 Photo-Music album application	38
A.4 Foreign language exercise materials	39
A.5 Storytelling application.....	39
A.6 Karaoke application	40
A.7 Slide show + Karaoke application	40
A.8 Online Musical slide show application format store	40
A.9 Storytelling content provider	41
A.10 Protected foreign language exercise	42
Annex B (informative) Examples of LAsER description element usage	43
B.1 General	43

B.2	Examples of basic transition effects	43
Annex C	(informative) Examples for MPEG-21 Metadata	47
C.1	Protecting all resources	47
C.2	Protecting MP3 audio	49
C.3	Protecting several JPEG images.....	50
C.4	Protecting specific segment in MP3 audio	51
C.5	Protecting specific region in selected JPEG images.....	53
C.6	Protecting slide show animation.....	54
Annex D	(informative) Implementation of Protected Musical slide show application format	56
D.1	General.....	56
D.2	Playing unprotected resource	56
D.3	Playing protected resource	57
D.4	Playing protected resource when exercise limit license already expired	58
D.5	Playing protected resource when validity condition license already expired	58
D.6	Playing protected resource with different protection tool	59
Bibliography		60

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 23000-4 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 which has been technically revised.

ISO/IEC 23000 consists of the following parts, under the general title *Information technology — Multimedia application format (MPEG-A)*:

- *Part 1: Purpose for multimedia application formats*
- *Part 2: MPEG music player application format*
- *Part 3: MPEG photo player application format*
- *Part 4: Musical slide show application format*
- *Part 5: Media streaming application format*
- *Part 6: Professional archival application format*
- *Part 7: Open access application format*
- *Part 8: Portable video application format*
- *Part 9: Digital Multimedia Broadcasting application format*
- *Part 10: Video surveillance application format*
- *Part 11: Stereoscopic video application format*

Introduction

ISO/IEC 23000 (also known as “MPEG-A”) is an MPEG standard that supports a fast track to standardization by selecting readily tested and verified tools taken from the MPEG body of standards and combining them to form an AF (Application Format). If a needed piece of technology is not provided within the MPEG, then additional technologies originating from other organizations can be included by reference in order to facilitate the envisioned application format.

The existing music player application format (ISO/IEC 23000-2) was designed as a format for enhanced MP3 players. It contains MP3 audio data, MPEG-7 metadata and an optional JPEG still image for cover art. The photo player application format (ISO/IEC 23000-3) is a format for digital photo library applications. It contains JPEG still images and associated MPEG-7 metadata.

The musical slide show application format (ISO/IEC 23000-4) is a richer multimedia format that builds on top of the music player and the photo player application format. This format supports the use of MP3 audio data along with multiple JPEG images in the form of a slide show presentation, and it is designed to render timed text data for annotations or lyrics. The format also features animation effects for image transitions and synchronization of media data.

Information technology — Multimedia application format (MPEG-A) —

Part 4: Musical slide show application format

1 Scope

This part of ISO/IEC 23000 specifies a file format for multimedia applications that feature MP3 audio playback and image slide show presentation. It also defines other technical features such as timed text (e.g. song lyrics) and animation (image transition effect).

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 14496-3:2005, *Information technology — Coding of audio-visual objects — Part 3: Audio*

ISO/IEC 14496-12:2005, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 14496-14:2003, *Information technology — Coding of audio-visual objects — Part 14: MP4 file format*

ISO/IEC 14496-20:2006, *Information technology — Coding of audio-visual objects — Part 20: Lightweight Application Scene Representation (LASER) and Simple Aggregation Format (SAF)*

ISO/IEC 15938-2, *Information technology — Multimedia content description interface — Part 2: Description definition language*

ISO/IEC 15938-5:2003, *Information technology — Multimedia content description interface — Part 5: Multimedia description schemes*

ISO/IEC 15938-10, *Information technology — Multimedia content description interface — Part 10: Schema definition*

ISO/IEC 21000-2, *Information technology — Multimedia framework (MPEG-21) — Part 2: Digital Item Declaration*

ISO/IEC 21000-4, *Information technology — Multimedia framework (MPEG-21) — Part 4: Intellectual Property Management and Protection Components*

ISO/IEC 21000-5, *Information technology — Multimedia framework (MPEG-21) — Part 5: Rights Expression Language*

ISO/IEC 21000-17, *Information technology — Multimedia framework (MPEG-21) — Part 17: Fragment Identification of MPEG Resources*

3GPP TS 26.245, *Transparent end-to-end Packet switched Streaming Service (PSS); Timed text format*, V7.0.0, 2007-06-21

3 Terms and definitions

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

3.1

slide show track

video track consisting of timed JPEG images

3.2

animation effect

simple image filtering effects applied on image transitions

4 Overview of MPEG Standards for Musical slide show application format

4.1 MPEG-1 Layer 3

ISO/IEC 11172-3:1993 specifies MPEG-1 Audio [1]. From that specification, MPEG-1 Layer 3 (or MP3) is one of the most widely deployed MPEG audio standards ever. Its wide appeal is due to both its good compression performance and its simplicity of implementation. The vast majority of compressed music archives use MP3 encoding.

One aspect of the simplicity of Layer 3 is that it specifies a self-synchronizing transport, making it amenable to both storage in a computer file and transmission over a channel without byte framing. In the context of transmission channels, Layer 3 can operate over a constant-rate isochronous link, and has constant-rate headers (as does Layer 1 and 2). However Layer 3 is an instantaneously-variable-rate coder, which adapts to the constant-rate channel by using a “bit buffer” and “back pointers.” Each of the headers signals the start of another block of audio signal, however due to the Layer 3 syntax, the data associated with that next block of audio signal may be in a prior segment of the bit stream, pointed to by the back pointer (see Figure 1, specifically the curved arrow pointing to `main_data_begin`). We note that this is in contrast to the MPEG-4 view of data stream segmentation, in which one access unit contains all information necessary to decode one segment of audio.

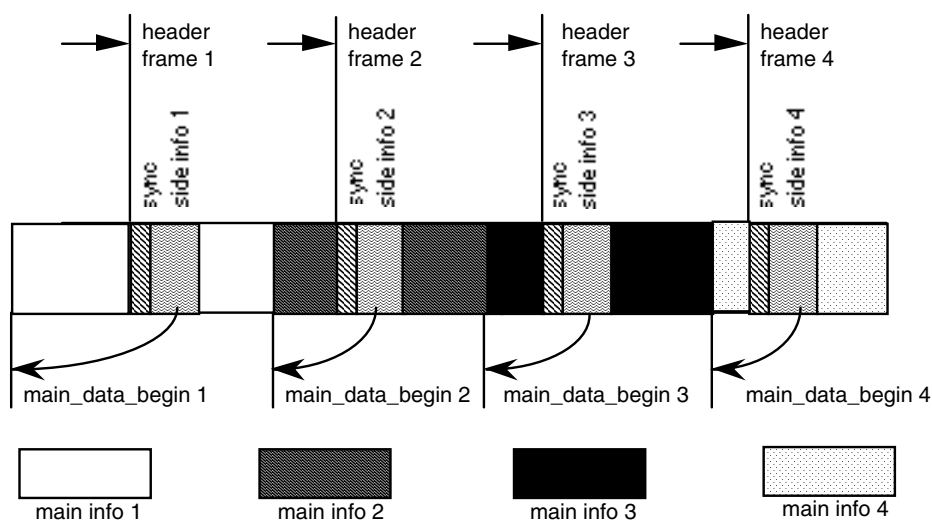


Figure 1 — Layer 3 bit stream organization

4.2 ISO Base Media File Format:

The ISO Base Media File Format is designed to contain timed media information for a presentation in a flexible, extensible format that facilitates interchange, management, editing, and presentation of the media. The ISO Base Media File Format is a base format for media file formats. In particular, the MPEG-4 file format derives from this base file format.

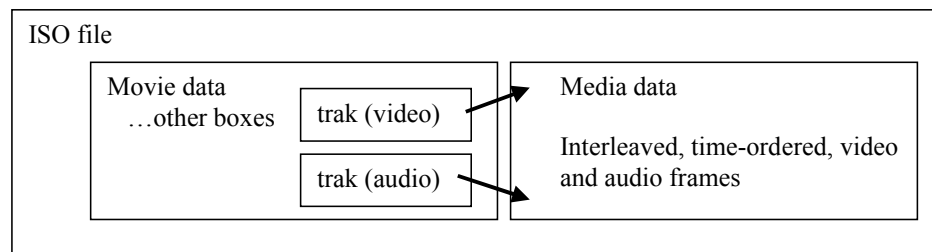


Figure 2 — Example of a simple ISO file used for interchange, containing two streams

The file structure is object-oriented as shown in Figure 2 which means that a file can be decomposed into constituent objects very simply, and the structure of the objects inferred directly from their type. The file format is designed to be independent of any particular network protocol while enabling efficient support for them in general.

4.3 The ISO Base Media and MPEG-4 File Formats

ISO/IEC 14496-12:2005, and ISO/IEC 14496-14:2003 together specify the MPEG-4 File Format. This supports storage of compressed audio data in tracks. It also provides support for metadata in the form of 'meta' boxes at the file-, movie- and track-level. This allows support for static (un-timed) metadata. Figure 3 schematically illustrates the location of these un-timed MPEG-7 metadata boxes.

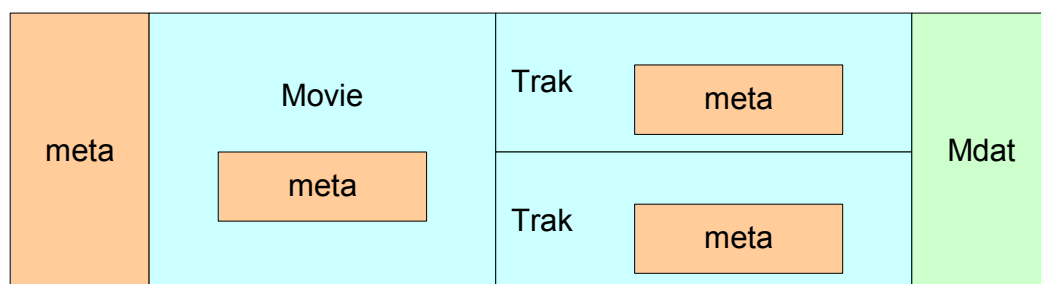
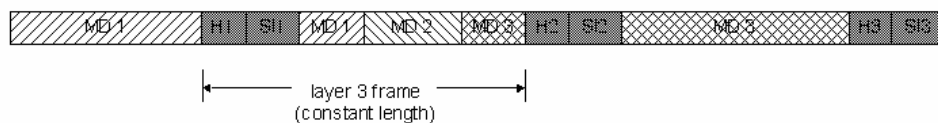
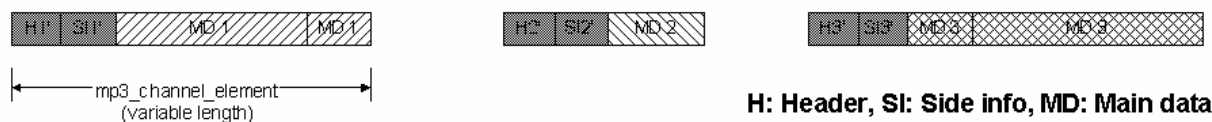


Figure 3 — Support of static un-timed metadata in ISO/MP4 files

4.4 MPEG-4 “MPEG-1/2 Audio in MPEG-4”

ISO/IEC 14496-3:2005 MPEG-4 Audio, Section 9 “MPEG-1/2 Audio in MPEG-4” specifies a method for segmenting and formatting Layer III bitstreams into MPEG-4 Access Units, and therefore is often referred to as “MP3onMP4”. This consists primarily of re-arranging the compressed data associated with a given header such that it follows the header. This typically results in new segments that are no longer of constant length, but that is perfectly in accordance with the definition of MPEG-4 Access Units. See example in Figure 4.

Layer 3 bitstream**Access units consisting of mp3_channel_elements****Figure 4 — Converting an MPEG-1/2 Layer 3 bitstream into mp3_channel_elements****4.5 MPEG-4 LAsER**

ISO/IEC 14496-20:2006, the Lightweight Application Scene Representation (LAsER), is a scene description format that specifies various aspects of 2D scene representation and updates of scenes as a part of rich media content. A scene description is composed of graphics, animation, text, and spatial and temporal layout.

A scene description specifies the following areas of a presentation:

- Spatial layout of the visual elements
- Temporal organization of the media elements (e.g. synchronization)
- Interactivity (e.g. mouse clicks, key inputs)
- Change of scenes (e.g. animation effects)

LAsER is designed to be suitable for lightweight embedded devices such as mobile phones.

4.6 MPEG-7 Multimedia Description Scheme

ISO/IEC 15938-5:2003 Multimedia description scheme (MDS) specifies all non-Visual and non-Audio specific metadata (e.g. Artist, Title, Date) in the MPEG-7 standard. As such it is able to represent all of the information found in the popular ID3V1 [2] metadata specification system.

5 File structure of Musical slide show application format**5.1 General**

This clause provides necessary information for creating and playing a Musical slide show application format file.

5.2 Components of Musical slide show application format

Creating a Musical slide show application format file involves formatting different types of media data, and storing them into an MPEG-4 file format.

A Musical slide show application format file consists of:

- MP3 audio
- JPEG images
- Timed text (e.g. lyrics, if any)
- LAsER scene description for animation effects (if any)

Figure 5 shows an example of Musical slide show application format creator system architecture. MP3 audio, JPEG images, and text data are formatted as individual MP4 media tracks. Descriptions for the animation effects are stored as LAsER scene description in XML format.

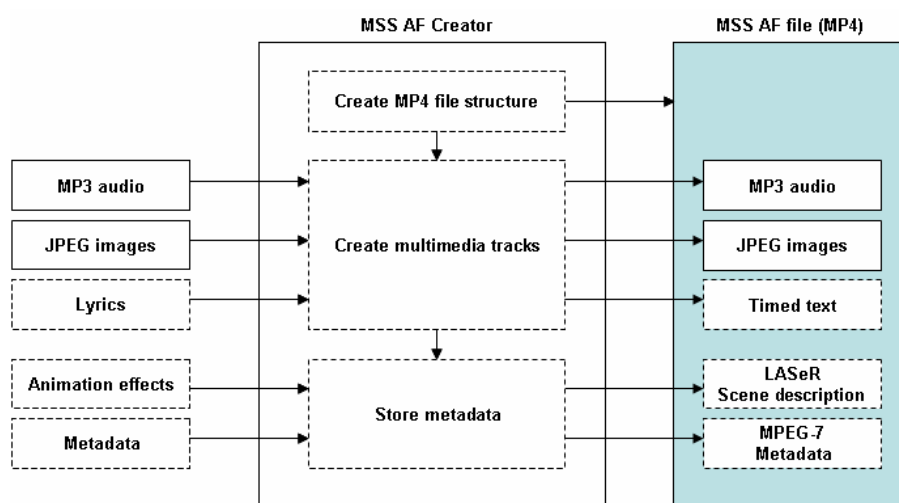


Figure 5 — Example of Musical slide show application format creator system architecture

5.3 File format

The file structure for the Musical slide show application format is based on the MPEG-4 file format. An example of a file structure containing a single MP3 audio track, a timed text track, and a slide show track for N-number of images is shown in Figure 6.

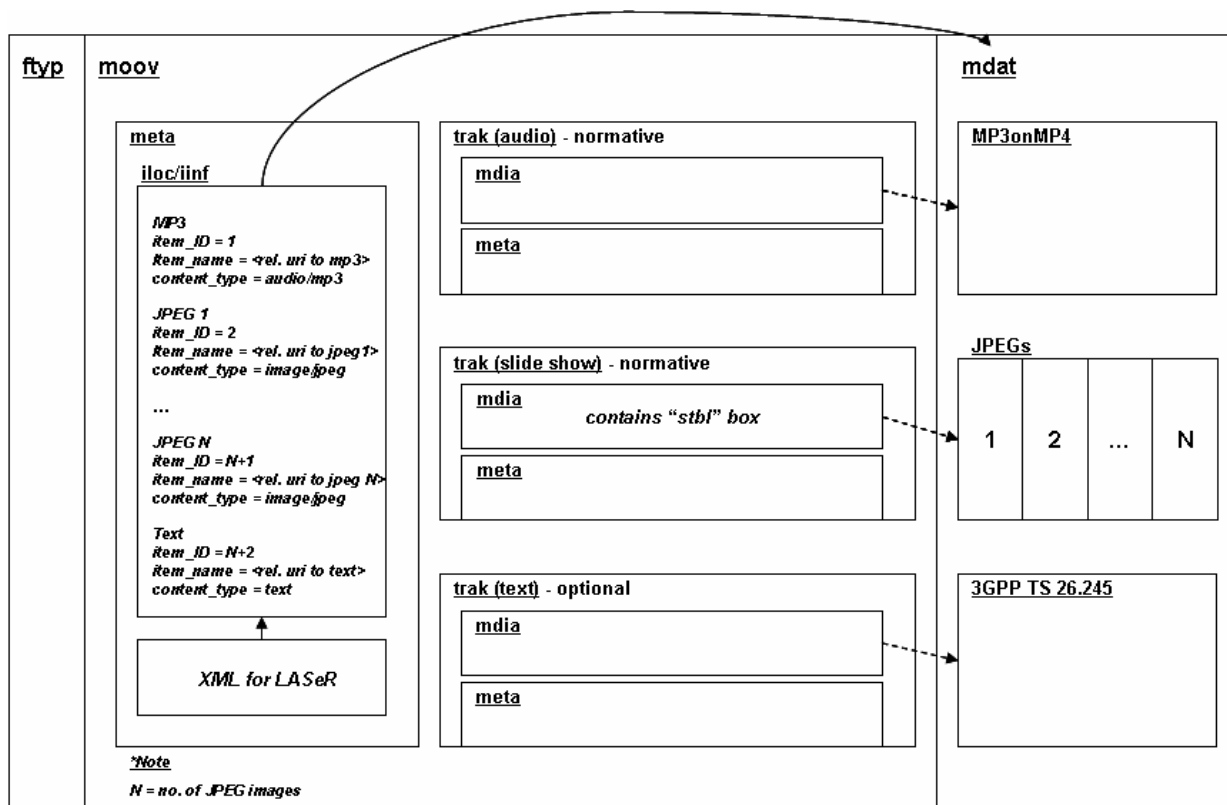


Figure 6 — Example of Musical slide show application format file structure with animation

The normative file structure consists of three boxes at the file level:

- File Type Box
- Movie Box
- Media Data Box

File Type Box ('ftyp'):

The 'ftyp' box defines the type of the file format that the file structure complies to. For the file type box ('ftyp'), the major-brand is 'mp42.'

The brand that identifies the Musical slide show application format file format is:

- 'mss1'

The brand 'mss1' is used as a compatible brand for the Musical slide show application format file format. A detailed description of the two modes is provided in subclause 5.4.

Movie Box ('moov'):

The 'moov' box contains three types of tracks (slide show, audio, text) and a metadata box:

- Normative Slide show Track Box (A 'trak' box for timed JPEG images)
- Normative Audio Track Box (MP3 audio)
- Optional Text Track Box (Timed text)

- Optional Metadata Box (Media resource information and LAsER scene description)

The 'trak' boxes contain temporal and spatial information of the media data (JPEG images, MP3 audio, timed text). For the Musical slide show application format, all the images that are used in the slide show presentation are arranged in a single track.

A Musical slide show player shall support application formats with the following number of tracks:

- Single slide show track (normative)
- Single audio track (normative)
- Single text track (optional)

The track handler types for the above tracks are:

- 'vide' for the slide show track
- 'soun' for the audio track
- 'text' for the text track

A detailed description of the "Slide show" track is provided in subclause 6.2.

The movie level metadata box ('meta') contains the item information box ('iinf') and the item location ('iloc') box. For each media data, an item ID is assigned, and the physical location and size of the media data are contained in the item location box. The item name and the content type information are contained in the item information box.

The 'xml' box that is located in the movie level 'meta' box contains the LAsER scripts responsible for the animation effects, and since it exists as a single "file," the meta handler-type is 'lsr1' for the 'meta' box.

Media Data Box ('mdat'):

The "mdat" box contains the actual media data bytes.

In case of the resources of Musical slide show application format are protected with a flexible protection scheme, a profile of MPEG-21 IPMP Components may be used to signal such protection in the movie level metadata box. For this necessity, the LAsER script can be wrapped or carried under the *DIDL/Container/Descriptor/Statement* element as shown in Figure 7.

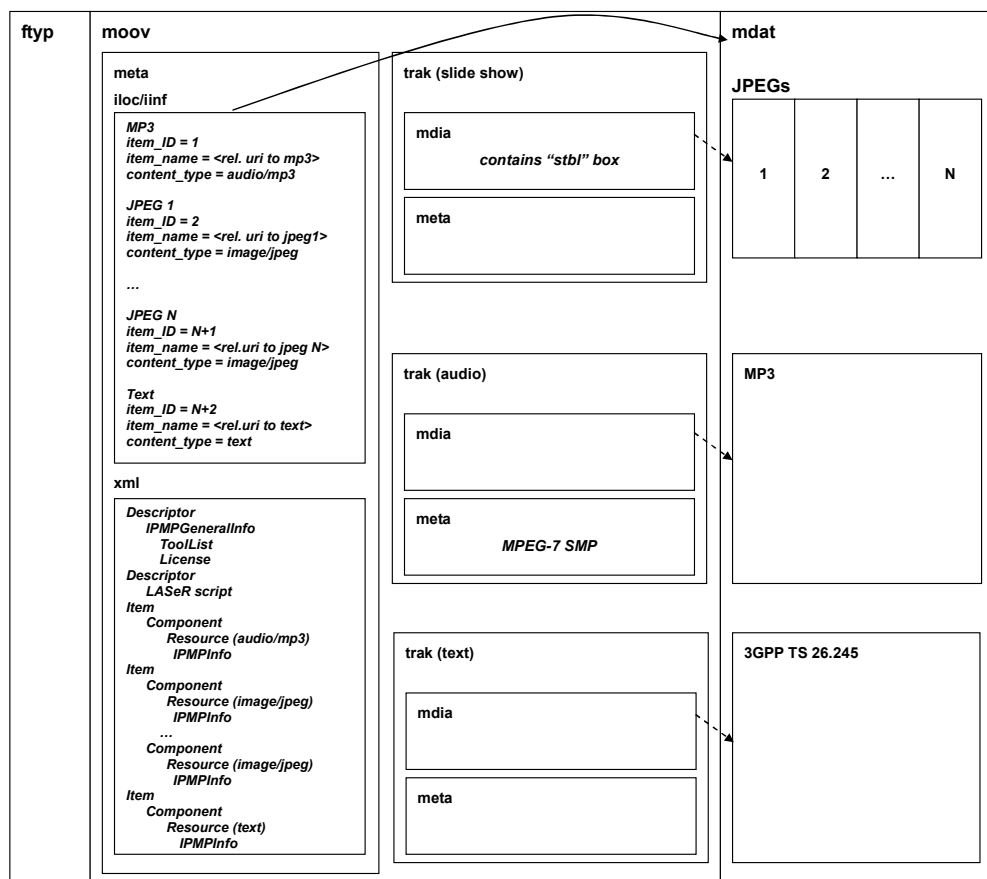


Figure 7 — The file structure of Musical slide show application format with IPMP DIDL

The more detailed description of protection scheme is described in clause 7 and clause 8.

5.4 Playback

For the Musical slide show application format, there are two possible rendering modes:

- Basic mode
- Enhanced mode

For the “Basic” mode, the timed text and slide show of JPEG images shall be rendered using the sample table (‘stbl’) box in the file format. In the “Enhanced” mode, LASer scene description shall be responsible for coordinating the overall presentation (slide show, animation effects, and timed text).

The operational flow for both Basic and Enhanced mode is shown in Figure 8.

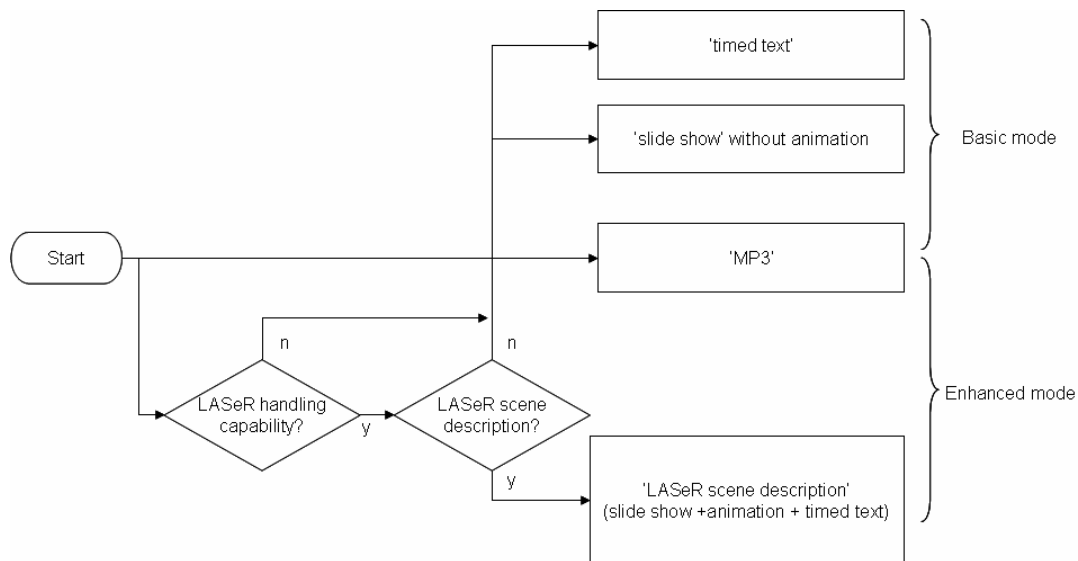


Figure 8 — Basic and Enhanced mode operational flow diagram

5.4.1 Basic mode

In the “Basic” mode, MP3, JPEG images and timed text (3GPP TS 26.245) are rendered concurrently by only using the information (timing, sample size, sample offset) obtained from the ‘stbl’ boxes. Therefore, when the file is loaded, the ‘moov’ box is parsed first, then, the tracks are read. For each track, the ‘stbl’ box is parsed in order to gain access to the spatial and temporal information regarding the sample data.

Players that are not capable of handling LASEr scene description may ignore the ‘meta’ box where LASEr scene description (‘xml’ box) is placed.

In this mode, the JPEG images are rendered based on the timing information in the ‘stbl’ box.

5.4.2 Enhanced mode

In the “Enhanced” mode, animation effects (described in subclause 6.3) shall be applied to the JPEG images in the slide show presentation.

The LASEr script is responsible for rendering the JPEG images and the timed text data (using the ‘text’ element). Therefore, the part that describes the timing information (i.e. subclause 6.2) regarding the JPEG images is ignored. For the “Enhanced” mode, the timeline of the slide show presentation is fully dependent on how the LASEr scene description is formed. The MP3 is played in the same way as in the “Basic” mode.

However, the sample table box of the slide show track shall contain timing information in case the player does not support LASEr decoding capability.

6 Technical features of Musical slide show application format

6.1 General

The following section describes the technical details that are featured in Musical slide show application format.

6.2 Synchronization of Media

Synchronization of the media data is primarily achieved with the use of the sample table box ('stbl'). The sample table contains all the time and data indices of the media samples in a track. For the slide show track, each JPEG image is considered to be a sample. Therefore, the timing information (slide show duration), and the physical sizes and locations of the images regarding the slide show presentation are stored inside the 'stbl' box. Specifically, the following sub-boxes are used:

- 'stts' (Decoding Time to Sample Box)
- 'stsz' (Sample Size Box)
- 'stco' (Chunk Offset Box)

The slide show duration is stored in the 'stts' box, and the image size and the image location are stored in the 'stsz' box and the 'stco' box respectively.

Figure 9 shows an example of allocating JPEG samples and referring them from the sample table box.

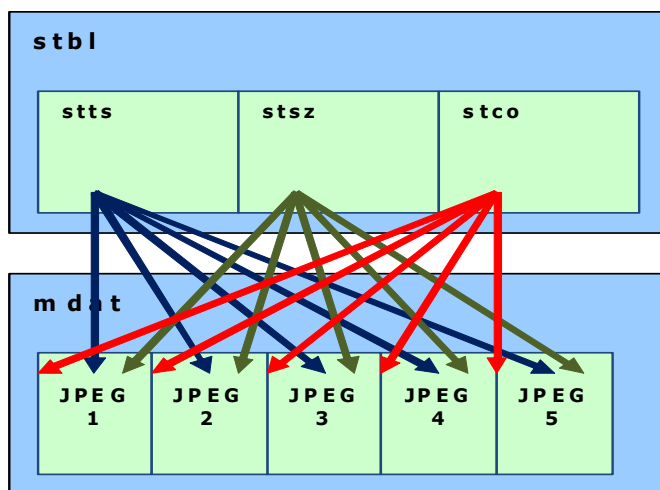


Figure 9 — Allocating several JPEGs as a collection of samples

This allocating method is also applied for the text track, where each line of text can be considered as a sample, consisting of a collection of strings.

In order for the timeline of the text track to be aligned with the slide show track, the timing information in the 'stts' box of the slide show track should act as the "clock." In other words, the text track should be fully dependent on the timeline of the slide show track.

Figure 10 shows an example of a synchronized Musical slide show application format presentation.

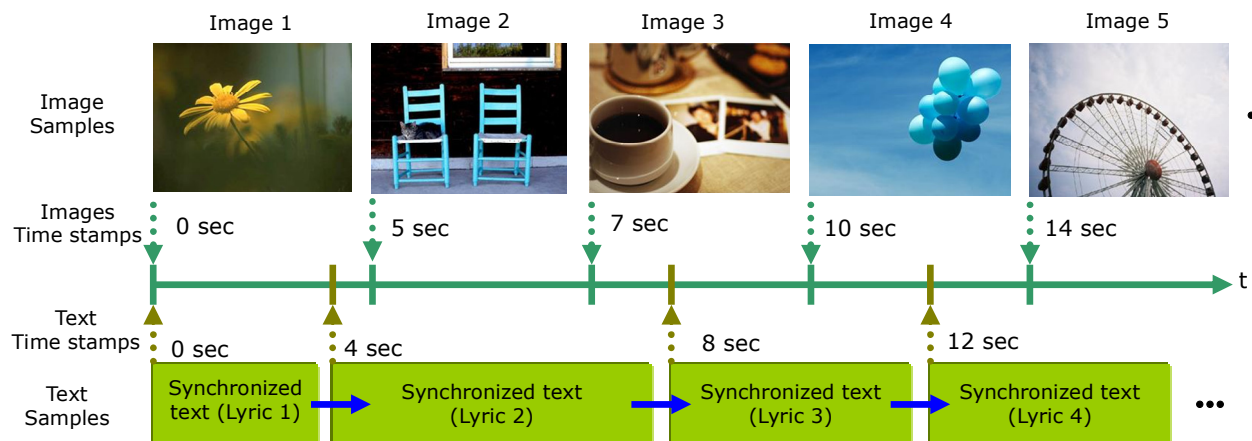


Figure 10 — Example of a synchronized Musical slide show application format presentation

6.3 Animation

The animation effects for the Musical slide show application format focus on creating a more entertaining image slide show presentation. Figure 11 shows an example of an animation effect that combines image transformation and opacity control.



Figure 11 — Example of animation

It is important to note that the effects featured in the Musical slide show application format shall only be used for image transitions during the slide show presentation, and they shall be comprised of simple image filtering effects. In addition, the timing information within LAsER can independently be defined regardless of the timing information in the sample table box.

6.3.1 Basic transition effects

Table 1 shows the functionalities and the description elements (LAsER) that shall be supported for the basic transition effects.

Table 1 — List of basic transition effects

Effects	Functionalities	Description elements	Semantics
Grouping	Effect grouping	<i>g</i>	Defined in subclause 6.8.15 of ISO/IEC 14496-20:2006
Image referencing	Image dimension	<i>image</i>	Defined in subclause 6.8.16 of ISO/IEC 14496-20:2006
	Referencing image		
Opacity control	Fade-in / Fade-out	<i>animate</i>	Defined in subclause 6.8.4 of ISO/IEC 14496-20:2006
Geometrical transformation	Translation	<i>animateTransform</i>	Defined in subclause 6.8.7 of ISO/IEC 14496-20:2006
	Scale		
	Rotation		
	Skew		

Object motion	Object motion on a predefined path	<i>animateMotion</i>	Defined in subclause 6.8.6 of ISO/IEC 14496-20:2006
Color change	Changes object color	<i>animateColor</i>	Defined in subclause 6.8.5 of ISO/IEC 14496-20:2006
Attribute control	Sets the value of an attribute	<i>set</i>	Defined in subclause 6.8.28 of ISO/IEC 14496-20:2006
Shapes & Motion path		<i>path</i>	Defined in subclause 6.8.22 of ISO/IEC 14496-20:2006
Basic shapes		<i>rect</i>	Defined in subclause 6.8.26 of ISO/IEC 14496-20:2006
		<i>circle</i>	Defined in subclause 6.8.9 of ISO/IEC 14496-20:2006
		<i>ellipse</i>	Defined in subclause 6.8.13 of ISO/IEC 14496-20:2006
		<i>line</i>	Defined in subclause 6.8.17 of ISO/IEC 14496-20:2006
		<i>polyline</i>	Defined in subclause 6.8.24 of ISO/IEC 14496-20:2006
		<i>polygon</i>	Defined in subclause 6.8.23 of ISO/IEC 14496-20:2006

6.3.2 Using LAsER in a textual format (XML)

Using LAsER in a textual format provides an easy way to create and edit descriptions for the animation effects, since the input data can simply be typed in, and the data itself can be more intuitive in terms of understanding the functionalities. Therefore, textual format shall be the normative way of using LAsER. In the Musical slide show application format, a reduced set (Table 1) of scene description elements for animation is used in local playback settings. Therefore, the data size or the decoding speed may not be an issue in terms of parsing or decoding the data. Figure 12 shows a possible model for a LAsER renderer.

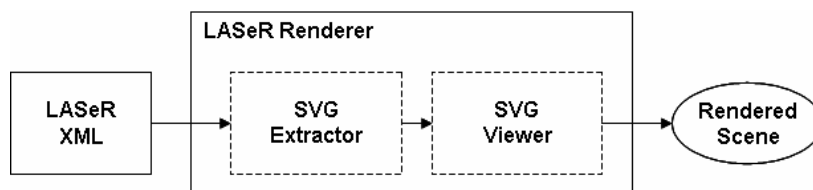


Figure 12 — LAsER renderer model for Musical slide show application format

6.4 Timed text

Timed text is intended to be used for applications (e.g. “karaoke” and language study materials) that require extensive use of textual presentation. In the Musical slide show application format, there are two possible ways to render timed text.

For players that are not capable of handling LAsER scene description (Basic mode) or contents that only require minimum use of textual presentation, 3GPP TS 26.245 timed text format is used.

3GPP TS 26.245 timed text data consists of:

- Text samples
- Sample descriptions

A text sample consists of one text string and text modifiers (optional). Sample descriptions and text modifiers are parameters that determine how the text string is to be displayed.

Sample descriptions provide global information such as font, position, background color about a text sample or samples, where as text modifiers provide information about a text string when it is displayed.

For the Musical slide show application format, there are four types of text modifiers (optional):

- 'styl' (for text style)
- 'hlit' (for highlighted text)
- 'krok' (for Karaoke, closed captioning and dynamic highlighting)
- 'blnk' (for blinking text)

For detailed sample description and text modifier syntax, refer to 3GPP TS 26.245 specification.

Figure 13 shows the structural interpretation of text sample.

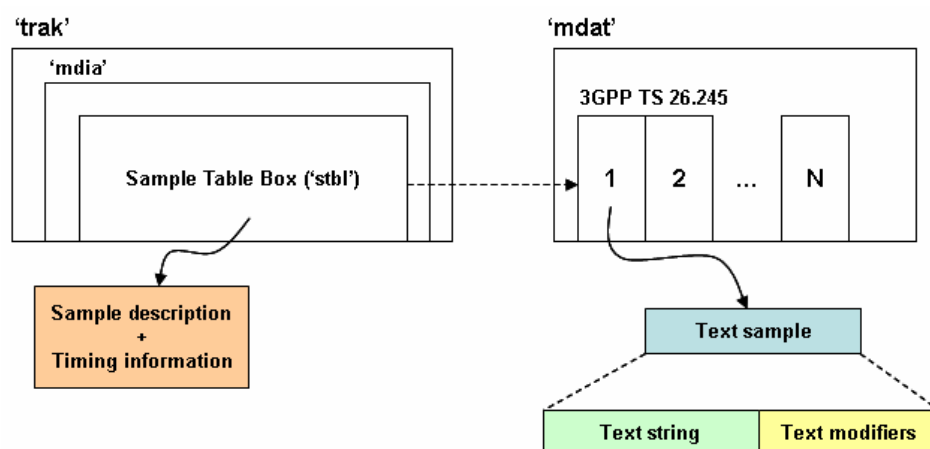


Figure 13 — Structural interpretation of text sample

The players that are capable of handling LAsER scene description (Enhanced mode), the 'text' element in LAsER is used for timed text functionality.

The supported functionalities (optional) of timed text are:

- Characters and glyphs support
- Font Support
- Color Support
- Text rendering position and composition
- Highlighting, closed captioning and "karaoke"

6.5 Metadata

For the Musical slide show application format, metadata provides simple background information, such as creation date, artist/creator information, and title of a photo series or a song.

The two types of normative metadata (textual XML) included in the Musical slide show application format are:

- Collection and item level metadata for the slide show track (Collection-level metadata are for allowing users to define groups/categories/sets of photos and to store metadata relating

to those groups, independently of the ordering of the slide-show; Item-level metadata are for enabling content-based search of slides)

■ Metadata for the audio track

The current Musical slide show application format file structure allows the metadata to be stored inside the media tracks. Figure 14 shows the locations of the metadata in the Musical slide show application format file structure.

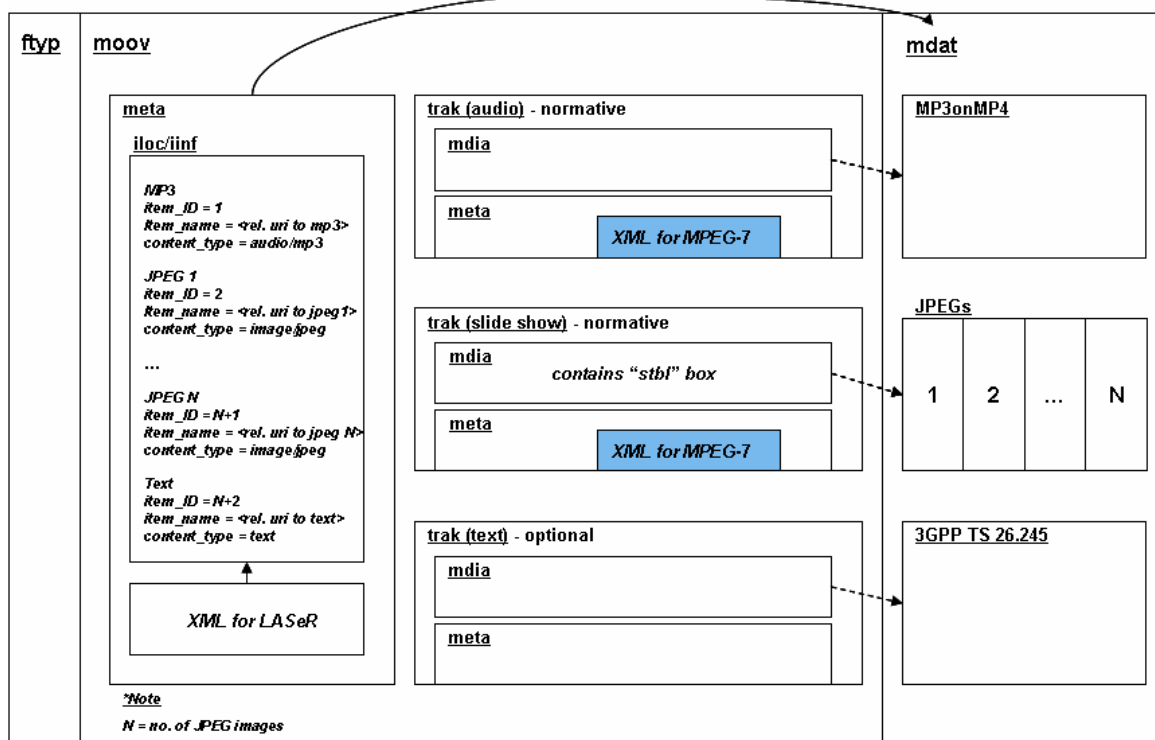


Figure 14 — Locations of metadata in the Musical slide show application format file structure

The metadata handler type is 'mp7t' for both slideshow and audio tracks.

6.5.1 Metadata for the slide show track

For the Musical slide show application format, images are structurally arranged in a single track, therefore, both collection and item level metadata are contained inside the slide show track as single XML data.

For the collection and item level (metadata for individual photos) descriptive metadata, MPEG-7 *ContentCollection* and *Image* DS are used, respectively (aligned with the Photo player application format). In order to combine the two metadata, the *Image* DS for the item-level metadata shall be contained under the *Content* element in the *ContentCollection* DS for the collection-level metadata. Every photo in the file shall have a corresponding *Content* element in the root collection. This means there shall be as many *Content* elements in the root collection as there are photos.

In the item-level metadata, *MediaLocator* DS is used for associating the metadata pertaining to the individual image with its resource data within the file, identified by its **item_ID**. In the collection level, images can be referenced using the *ContentRef* element. The *ContentRef* element shall only exist in sub-collections.

Tables 2 and 4 summarize the semantics of available tools. The normative specification of all semantics is given in ISO/IEC 15938-2 and ISO/IEC 15938-5:2003.

The slide show track schema is defined with respect to the MPEG-7 Version 2 schema as specified in ISO/IEC 15938-10. The namespace of the Version 2 schema providing a basis for the slide show track schema is "urn:mpeg:mpeg7:schema:2004". Tables 3 and 5 list the ISO/IEC 15938 description tools (global elements, global attributes, attribute groups, complexTypes and simpleTypes) selected to be included and any further constraints imposed on these description tools for the slide show track metadata schema. The interpretation of the Table is as follows:

- "element/attribute/attributeGroup excluded" – this MPEG-7 metadata shall not be instantiated in the slide show track metadata
- `xsi:type="[TypeName]"` – the element shall have this attribute, when instantiated in the slide show track metadata. Therefore, it shall only be instantiated with type [TypeName]
- `minOccurs="n"` – no fewer than *n* occurrences of the element shall be instantiated in the slide show track metadata
- `maxOccurs="m"` – no more than *m* occurrences of the element shall be instantiated in the slide show track metadata

Where no constraint is listed, the metadata elements/attributes shall be instantiated in accordance with ISO/IEC 15938-10.

Table 2 — Semantics for the slide show track metadata (Collection-level)

Tag Name		Semantics
DescriptionMetadata/ Creator	Optional	The author of the collection definition. The term "Author" (registered in RoleCS) shall be used as his/her role.
DescriptionMetadata/ CreationTime	Optional	The time stamp when the collection definition was created.
DescriptionMetadata/ LastUpdate	Mandatory	The time stamp of the most recent change to the collection definition.
ContentCollection/ name (attribute)	Optional	The name of (i.e., title for) the collection.
ContentCollection/ CreationInformation/ Creation/ Title	Mandatory	A title of (all) the photos in the collection. Note that this is not the title of the collection itself, which is expressed by ContentCollection/name above.
ContentCollection/ CreationInformation/ Creation/ TitleMedia/ TitleImage	Optional	A representative thumbnail picture for the collection. The image coding format for thumbnails shall be either JPEG or one of the uncompressed formats supported by Exif headers – that is, Baseline TIFF Rev. 6.0 RGB Full Color and TIFF Rev. 6.0 Extension YCbCr Image.
ContentCollection/ CreationInformation/ Creation/ TitleMedia/ TitleImage/ InlineMedia/ type (Attribute)	Mandatory (if InlineMedia is present)	Describing the format of the thumbnail image data. This shall be one of the MIME types "image/jpeg" or "image/tiff".
ContentCollection/ CreationInformation/ Creation/ Creator	Optional	Describing a person or organization who relates to the creation process of the images in the collection, such as photographer, publisher and so on. Their roles should be described using RoleCS. A variety of methods can be used to identify the Creators, including electronic address elements such as url or email. If the Role is set to "Actor", this field shall describe the identity of persons who appear in the images in the collection.