INTERNATIONAL STANDARD

ISO/IEC 23000-4

Second edition
2009-01-15

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

Reference number
ISO/IEC 23000-4:2009(E)

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

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.

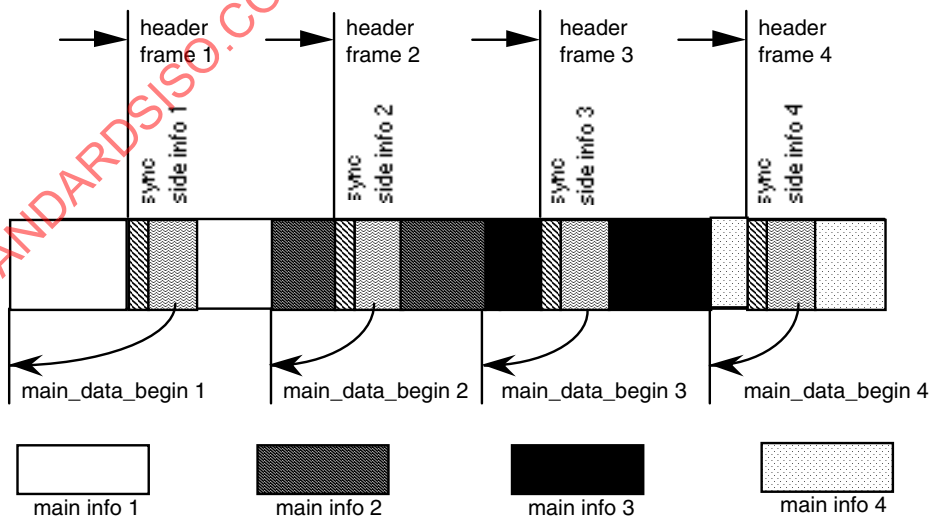


Figure 1 — Layer 3 bit stream organization

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.

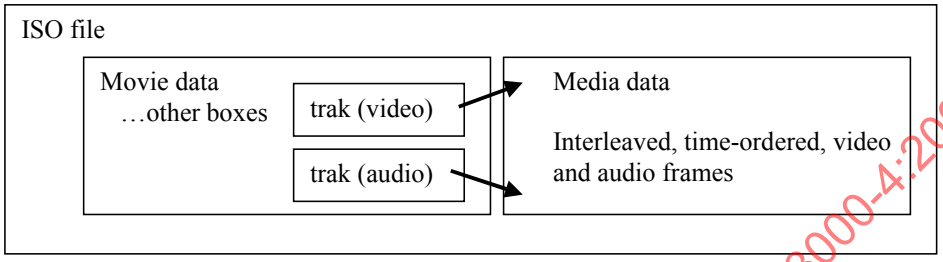


Figure 2 — Example of a simple ISO file used for interchange, containing two streams

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.

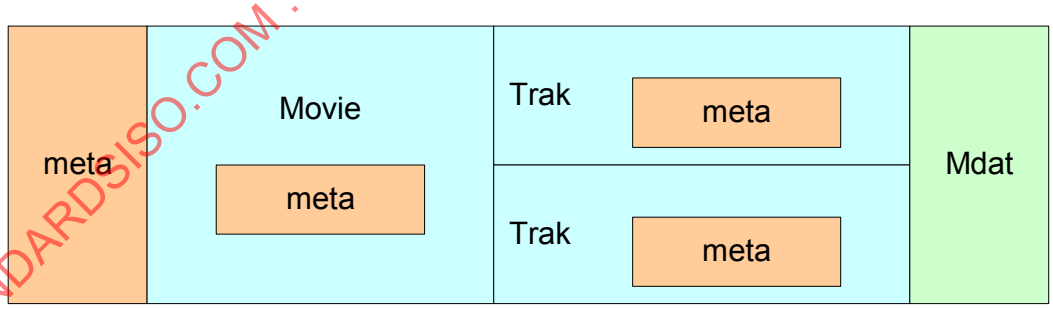


Figure 3 — Support of static un-timed metadata in ISO/MP4 files

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.

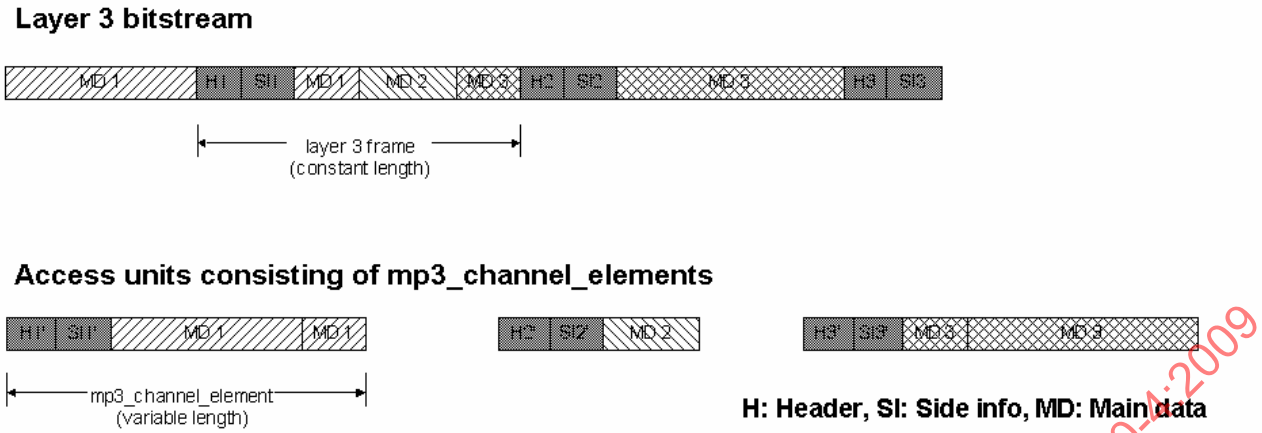


Figure 4 — Converting an MPEG-1/2 Layer 3 bitstream into mp3_channel_elements

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.

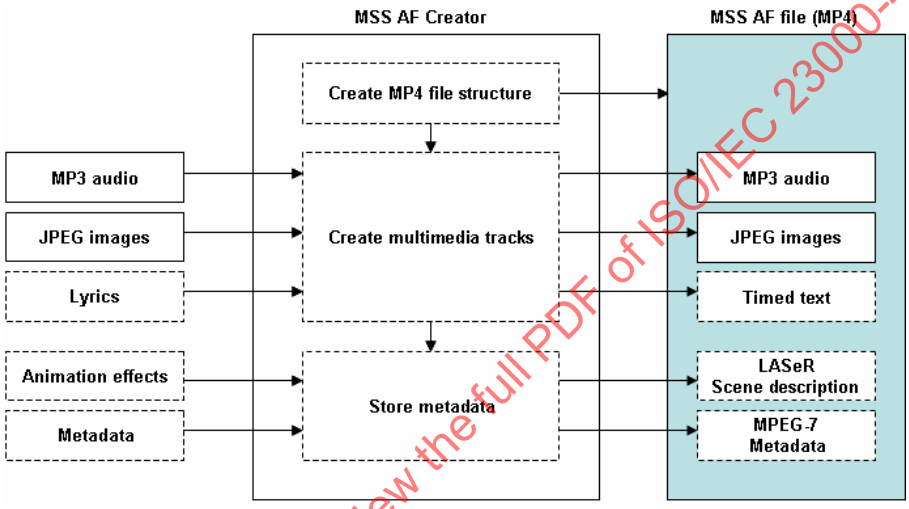


Figure 5 — Example of Musical slide show application format creator system architecture

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.

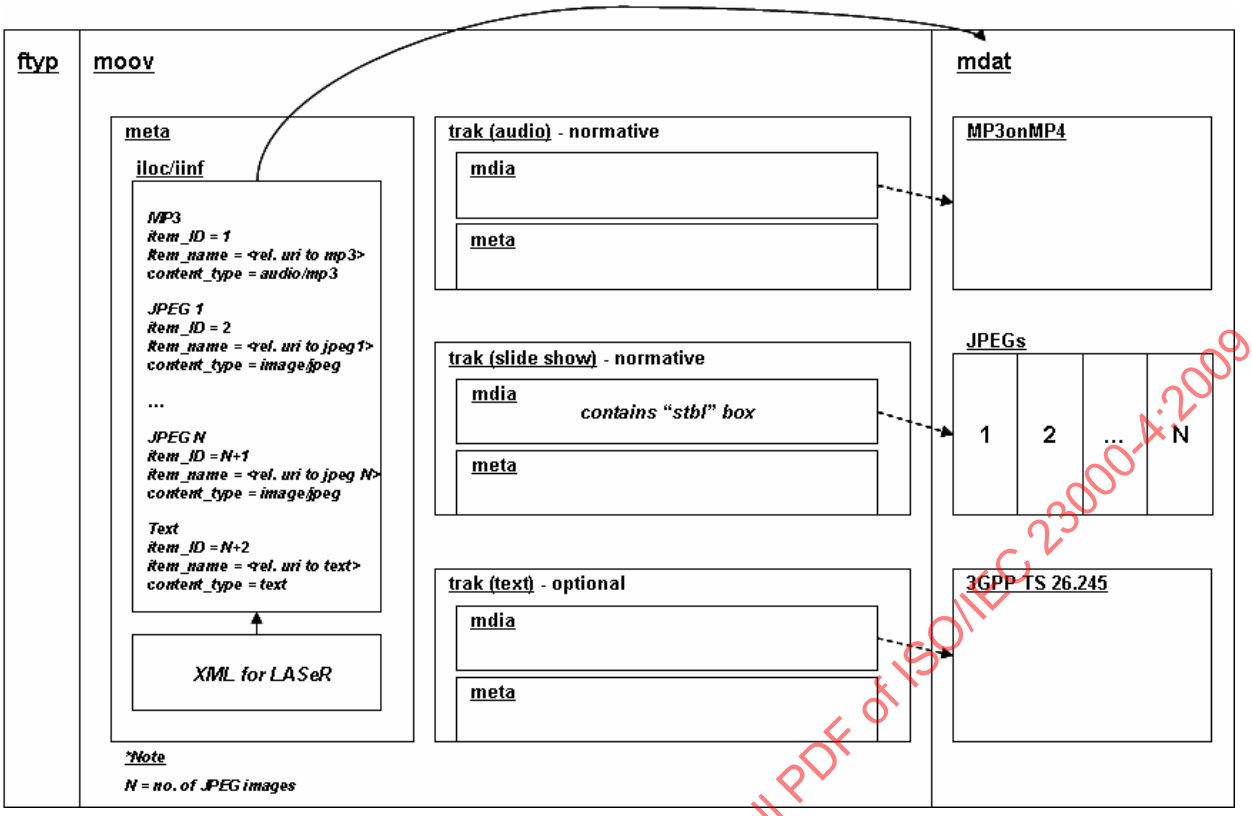


Figure 6 — Example of Musical slide show application format file structure with animation

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.

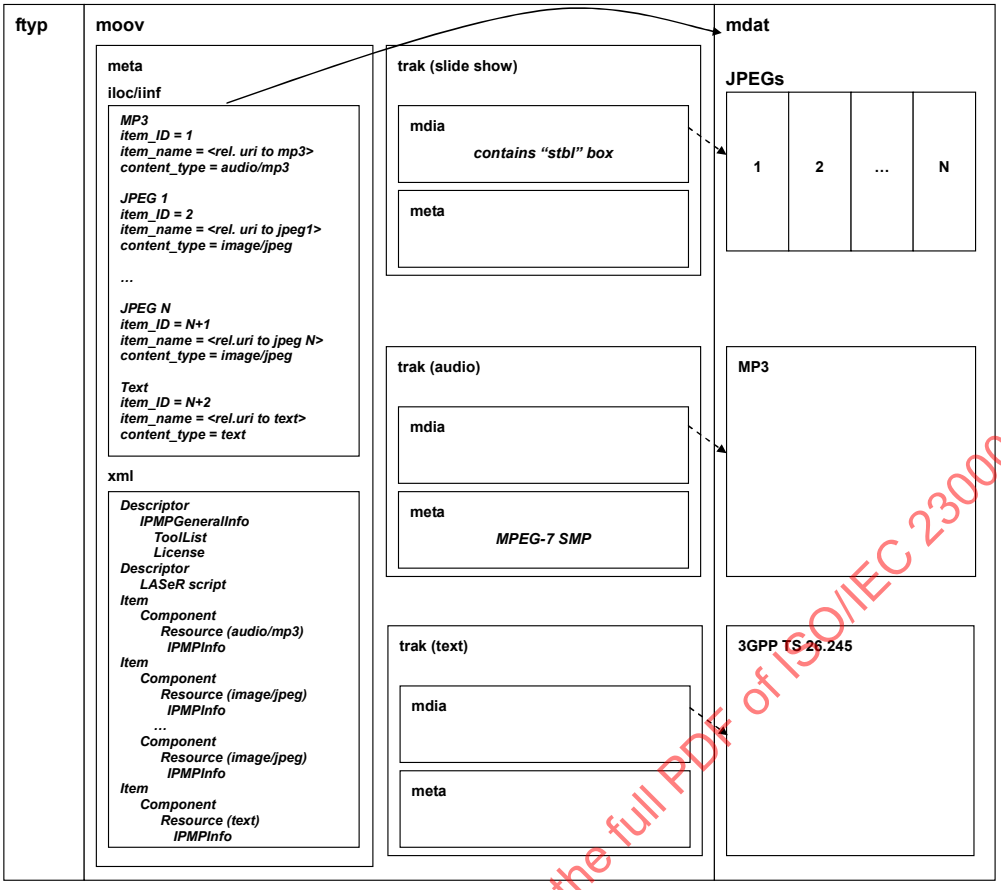


Figure 7 — The file structure of Musical slide show application format with IPMP DIDL

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.

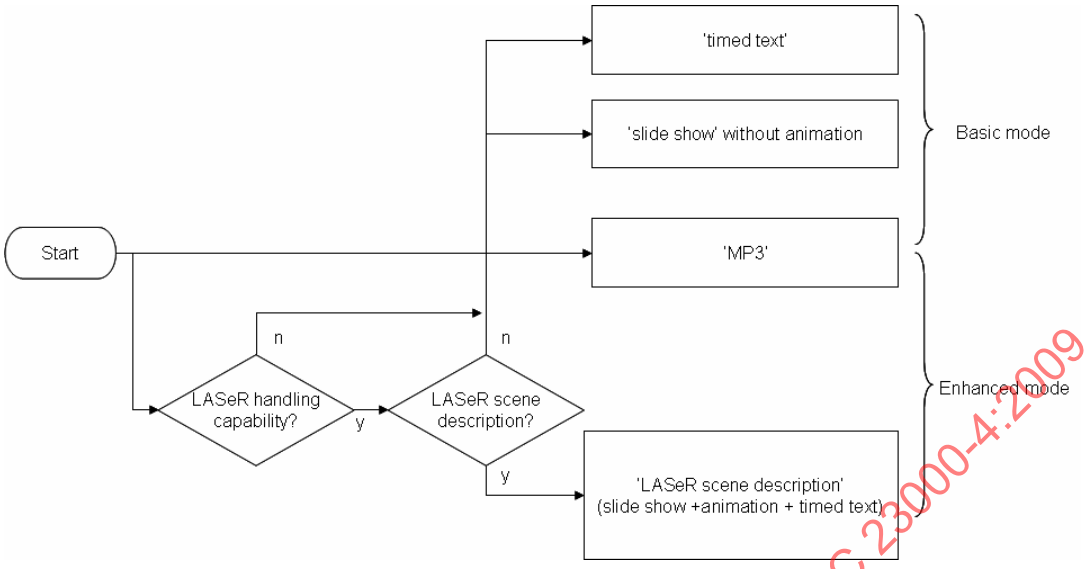


Figure 8 — Basic and Enhanced mode operational flow diagram

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.

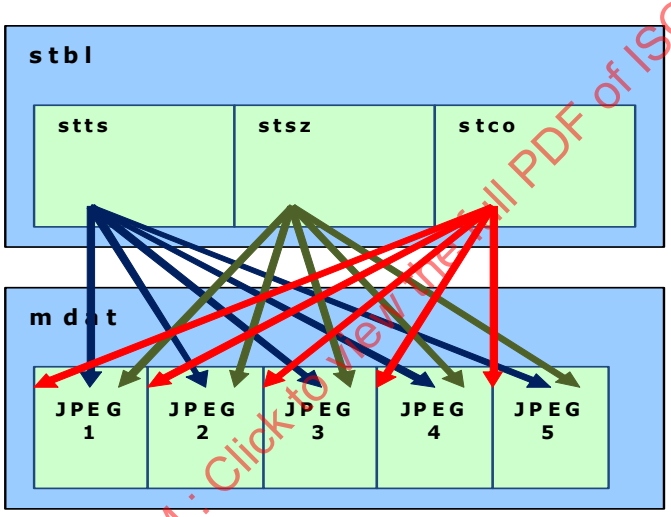


Figure 9 — Allocating several JPEGs as a collection of samples

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.

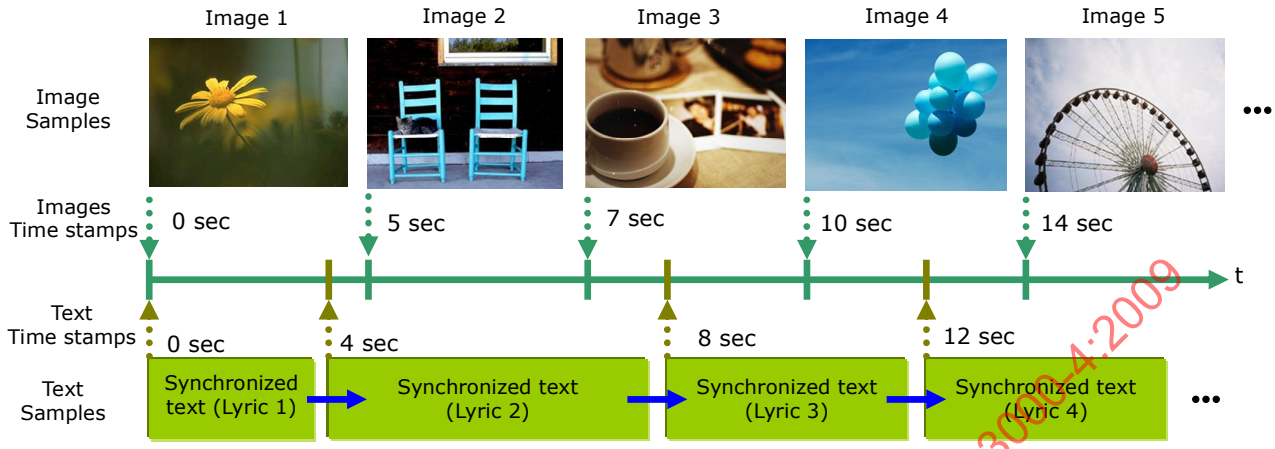


Figure 10 — Example of a synchronized Musical slide show application format presentation

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	*g*	Defined in subclause 6.8.15 of ISO/IEC 14496-20:2006
Image referencing	Image dimension	*image*	Defined in subclause 6.8.16 of ISO/IEC 14496-20:2006
	Referencing image		
Opacity control	Fade-in / Fade-out	*animate*	Defined in subclause 6.8.4 of ISO/IEC 14496-20:2006
Geometrical transformation	Translation	*animateTransform*	Defined in subclause 6.8.7 of ISO/IEC 14496-20:2006
	Scale		
	Rotation		
	Skew		

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

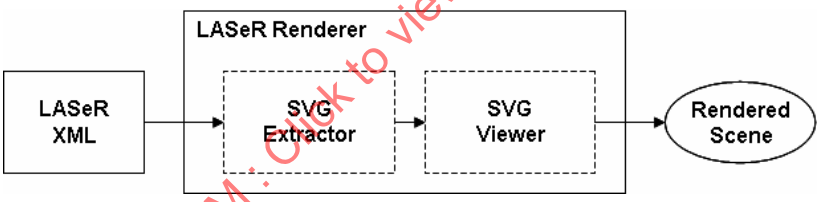


Figure 12 — LASeR renderer model for Musical slide show application format

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.

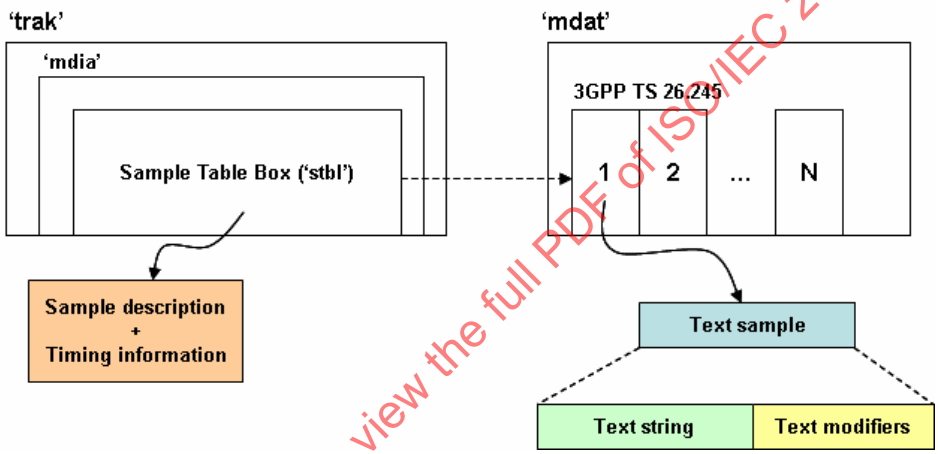


Figure 13 — Structural interpretation of text sample

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.

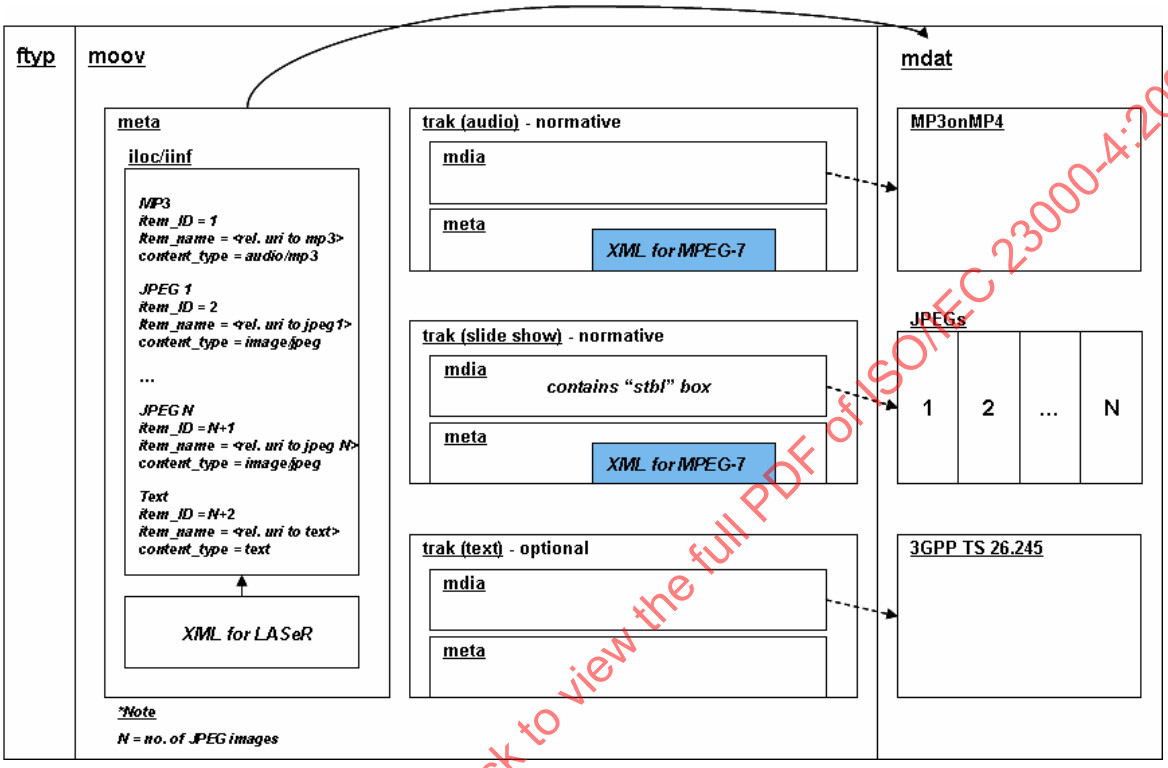


Figure 14 — Locations of metadata in the Musical slide show application format file structure

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.

ContentCollection/ CreationInformation/ Creation/ Date	Optional	The time (or period in time) when the photos in the collection were captured. The creation time of photos in any sub-collections should be included within the time period for the current collection.
ContentCollection/ CreationInformation/ Creation/ Location	Optional	The location where the photos in the collection were captured.
ContentCollection/ TextAnnotation/ FreeTextAnnotation	Optional	Summary text about the collection.
ContentCollection/ TextAnnotation/ KeywordAnnotation	Optional	Any keywords of the collection.
ContentCollection/ Content	Mandatory	Includes item level metadata.
ContentCollection/ ContentRef	Mandatory	The photos which are included in the collection. “idref” shall be employed to reference a *Content* element within the root collection.
ContentCollection/ ContentCollection	Optional	The sub-collections. Note that the root collection shall include all of the photos in the file. Hierarchical representation of collections is allowed.

Table 3 — Collection-level MPEG-7 metadata (normative)

Global Elements	**Name**	**Constraint**
Mpeg7	DescriptionUnit Description	xsi:type="ContentCollectionType" element excluded
Complex Types	**Element/Attribute Name**	**Constraint**
Mpeg7BaseType		
DSType	Header id timePropertyGrp mediaTimePropertyGrp	element excluded attributeGroup excluded attributeGroup excluded
HeaderType	id	
DescriptionMetadataType	Confidence Version LastUpdate Comment PublicIdentifier PrivateIdentifier Creator CreationLocation CreationTime Instrument Rights Package	element excluded element excluded minOccurs="1" element excluded element excluded element excluded element excluded element excluded element excluded element excluded
Mpeg7Type	DescriptionProfile DescriptionMetadata xml:lang timePropertyGrp mediaTimePropertyGrp	 minOccurs="1" attributeGroup excluded attributeGroup excluded
CollectionType	CreationInformation CreationInformationRef UsageInformation	minOccurs="0" element excluded element excluded

	UsageInformationRef TextAnnotation name	element excluded
ContentCollectionType	VisualFeature GofGopFeature AudioFeature Content ContentRef ContentCollection ContentCollectionRef	element excluded element excluded element excluded xsi:type=”ImageType” minOccurs="0" minOccurs="0" element excluded
CreationInformationType	Creation Classification RelatedMaterial	element excluded element excluded
CreationType	Title TitleMedia Abstract Creator CreationCoordinates Location Date CreationTool CopyrightString	maxOccurs="1" element excluded maxOccurs="1"
TimeType	TimePoint RelTimePoint RelIncrTimePoint Duration Incr Duration	element excluded element excluded minOccurs="0" element excluded
PlaceType	Name NameTerm PlaceDescription Role GeographicPosition Point datum AstronomicalBody Region AdministrativeUnit PostalAddress AddressLine PostingIdentifier xml:lang StructuredPostalAddress InternalCoordinates StructuredInternalCoordinates ElectronicAddress xml:lang	element excluded element excluded element excluded element excluded element excluded element excluded element excluded element excluded element excluded attribute excluded
GeographicPointType	longitude latitude altitude	
TitleType	type	attribute excluded
TitleMediaType	TitleImage TitleVideo TitleAudio	element excluded element excluded
TextAnnotationType	KeywordAnnotation FreeTextAnnotation StructuredAnnotation DependencyStructure	element excluded element excluded

	relevance confidence xml:lang	attribute excluded attribute excluded
KeywordAnnotationType	Keyword type xml:lang	
CreatorType	Character Instrument	element excluded element excluded
PersonNameType	GivenName FamilyName Title Numeration LinkingName Salutation dateFrom dateTo type xml:lang	 attribute excluded attribute excluded attribute excluded
NameComponentType	initial abbrev	attribute excluded attribute excluded
MediaAgentType	Role Agent AgentRef	 element excluded
InlineTermDefinitionType	Name preferred Definition Term	maxOccurs="1" attribute excluded element excluded element excluded
ControlledTermUseType	href	
AgentType	Icon	element excluded
PersonType	Name NameTerm Affiliation Organization OrganizationRef PersonGroup PersonGroupRef Citizenship Address AddressRef ElectronicAddress PersonDescription Nationality	maxOccurs="1" element excluded element excluded element excluded element excluded element excluded minOccurs="0" element excluded element excluded element excluded
OrganizationType	Name type NameTerm Kind Contact ContactRef Jurisdiction JurisdictionRef Address AddressRef ElectronicAddress	 attribute excluded element excluded element excluded element excluded element excluded element excluded element excluded minOccurs="0" element excluded

ElectronicAddressType	Telephone type Fax Email Url	attribute excluded
TextualBaseType	xml:lang phoneticTranscription phoneticAlphabet	attribute excluded attribute excluded
ImageLocatorType	MediaTimePoint MediaRelTimePoint MediaRelIncrTimePoint BytePosition	element excluded element excluded element excluded element excluded
InlineMediaType	MediaData16 MediaData64 type	
MediaLocatorType	MediaURI InlineMedia StreamID	element excluded minOccurs="0" element excluded
DescriptionProfileType	profileAndLevelIndication	
ReferenceType	mpeg7:referenceGroup	
TextualType		
Simple Types	**Element/Attribute Name**	**Constraints**
termURIReferenceType		
termAliasReferenceType		
termReferenceType		
basicDurationType		
durationType		
mimeType		
basicTimePointType		
timePointType		
Attribute Groups	**Attribute Name**	**Constraints**
referenceGrp	idref xpath href	attribute excluded

Table 4 — Semantics for the slide show track metadata (Item-level)

Tag Name		Semantics
Content/ Image/ MediaInformation/ MediaProfile/ MediaFormat	Optional	File properties of a resource. When appropriate, the FileSize and VisualCoding/Frame child elements should be used to specify the size of code-stream and pixel dimensions of the image, respectively.
Content/ Image/ MediaInformation/ MediaProfile/ MediaFormat/ Content	Mandatory (if MediaFormat is instantiated)	Describing the media. The term "Image" (registered in ContentCS) shall be used

Content/ Image/ MediaInformation/ MediaProfile/ MediaFormat/ FileFormat	Optional	Describing the format of the image. The terms from FileFormatCS should be used for this purpose.
Content/ Image/ MediaInformation/ MediaProfile/ MediaInstance/ MediaLocator/ MediaUri	Mandatory	Used for referencing individual image data that is contained inside the slide show track. URL reference shall be employed with MPEG-21 Fragment Identifier compliant format (ISO/IEC 21000-17).
Content/ Image/ CreationInformation/ Creation/ Title	Mandatory	The title of the resource. It may be obtained by referring to corresponding Exif tags of the resource.
Content/ Image/ CreationInformation/ Creation/ Creator	Optional	Describing a person or organization who relates to the creation process of the resource, 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. Regarding the photographer, such information may be obtained by referring to corresponding Exif tags of the resource. If the Role is set to “Actor”, this field shall describe the identity of persons who appear in the image.
Content/ Image/ TextAnnotation	Optional	Summary text of the resource. It may be obtained by referring to corresponding Exif tags of the resource.
Content/ Image/ CreationInformation/ Creation/ Location	Optional	The location where the resource was captured. GPS location information may be obtained by referring to corresponding Exif tags of the resource.
Content/ Image/ CreationInformation/ Creation/ Date	Optional	The time when the resource was captured. This may be obtained by referring to corresponding Exif tags of the resource.
Content/ Image/ VisualDescriptionScheme	Optional	Signal-level characteristics of the resource. Several elements may be instantiated.

Table 5 — Item-level MPEG-7 metadata (normative)

Complex Types	Element/Attribute Name	Constraint
ImageType	Image	
SegmentType	MediaInformation MediaInformationRef MediaLocator StructuralUnit CreationInformation CreationInformationRef	minOccurs="0" element excluded element excluded element excluded minOccurs="0" element excluded

	UsageInformation UsageInformationRef TextAnnotation type Semantic SemanticRef MatchingHint PointOfView Relation	element excluded element excluded attribute excluded element excluded element excluded element excluded element excluded element excluded
MediaInformationType	MediaIdentification MediaProfile	element excluded
MediaProfileType	ComponentMediaProfile MediaFormat MediaTranscodingHints MediaQuality MediaInstance master	element excluded element excluded element excluded maxOccurs="1" attribute excluded
MediaFormatType	Content Medium FileFormat FileSize System Bandwidth BitRate TargetChannelBitRate ScalableCoding VisualCoding Format Pixel Frame height width aspectRatio rate structure ColorSampling AudioCoding SceneCodingFormat GraphicsCodingFormat OtherCodingFormat	element excluded element excluded element excluded element excluded element excluded element excluded element excluded element excluded attribute excluded attribute excluded attribute excluded element excluded element excluded element excluded element excluded element excluded
MediaInstanceType	InstanceIdentifier MediaLocator LocationDescription	
UniqueIDType	type organization authority Encoding	 attribute excluded attribute excluded
MediaLocatorType	MediaUri InlineMedia StreamID	minOccurs="1" element excluded element excluded

StillRegionType	SpatialLocator SpatialMask MediaTimePoint MediaRelTimePoint MediaRelIncrTimePoint VisualDescriptor VisualDescriptionScheme GridLayoutDescriptors IllumininationInvariantColor MultipleView SpatialDecomposition	element excluded element excluded element excluded element excluded element excluded element excluded xsi:type="StillRegionFeatureType", maxOccurs="1" element excluded element excluded element excluded element excluded
VisualDType		
VisualDSType		
StillRegionFeatureType	DominantColor ScalableColor ColorStructure ColorLayout ColorTemperature IlluminationCompensatedColor Edge HomogeneousPattern TextureBrowsing ShapeMask Contour	element excluded element excluded element excluded element excluded element excluded
DominantColorType	ColorSpace ColorQuantization SpatialCoherency Value Percentage Index ColorVariance	element excluded element excluded
ScalableColor	Coeff numOfCoeff numOfBitplanesDiscarded	
ColorStructure	Values colorQuant	
ColorLayoutType	YDCCoeff CbDCCoeff CrDCCoeff YACCoeff2 YACCoeff5 YACCoeff9 YACCoeff14 YACCoeff20 YACCoeff27 YACCoeff63 CrACCoeff2 CrACCoeff5 CrACCoeff9 CrACCoeff14 CrACCoeff20 CrACCoeff27 CrACCoeff63 CbACCoeff2 CbACCoeff5	element excluded element excluded element excluded

	CbACCoeff9 CbACCoeff14 CbACCoeff20 CbACCoeff27 CbACCoeff63	
HomogeneousTextureType	Average StandardDeviation Energy EnergyDeviation	
EdgeHistogramType	BinCounts	
CreationInformationType	Creation Classification RelatedMaterial	 element excluded element excluded
CreationType	Title TitleMedia Abstract Creator CreationCoordinates Location Date CreationTool CopyrightString	maxOccurs="1" element excluded element excluded maxOccurs="1" element excluded element excluded
TextAnnotationType	KeywordAnnotation FreeTextAnnotation StructuredAnnotation DependencyStructure relevance confidence xml:lang	 element excluded element excluded attribute excluded attribute excluded
KeywordAnnotationType	Keyword type xml:lang	
PlaceType	Name NameTerm PlaceDescription Role GeographicPosition Point datum AstronomicalBody Region AdministrativeUnit PostalAddress AddressLine PostingIdentifier xml:lang StructuredPostalAddress InternalCoordinates StructuredInternalCoordinates ElectronicAddress xml:lang	 element excluded element excluded element excluded element excluded element excluded element excluded element excluded element excluded element excluded attribute excluded
GeographicPointType	Longitude Latitude altitude	

TitleType	type	attribute excluded
CreatorType	Character Instrument	element excluded element excluded
PersonNameType	GivenName FamilyName Title Numeration LinkingName Salutation dateFrom dateTo type xml:lang	 attribute excluded attribute excluded attribute excluded
NameComponentType	Initial abbrev	attribute excluded attribute excluded
MediaAgentType	Role Agent AgentRef	 element excluded
InlineTermDefinitionType	Name preferred Definition Term	maxOccurs="1" attribute excluded element excluded element excluded
ControlledTermUseType	href	
AgentType	Icon	element excluded
PersonType	Name NameTerm Affiliation Organization OrganizationRef PersonGroup PersonGroupRef Citizenship Address AddressRef ElectronicAddress PersonDescription Nationality	maxOccurs="1" element excluded element excluded element excluded element excluded element excluded minOccurs="0" element excluded element excluded element excluded
OrganizationType	Name type NameTerm Kind Contact ContactRef Jurisdiction JurisdictionRef Address AddressRef ElectronicAddress	 attribute excluded element excluded element excluded element excluded element excluded element excluded element excluded minOccurs="0" element excluded
ElectronicAddressType	Telephone type Fax Email Url	 attribute excluded

TextualBaseType	xml:lang phoneticTranscription phoneticAlphabet	 attribute excluded attribute excluded
TextualType		
TimeType	TimePoint RelTimePoint RelIncrTimePoint Duration Incr Duration	 element excluded element excluded element excluded element excluded
Simple Types	**Element/Attribute Name**	**Constraints**
termURIReferenceType		
termAliasReferenceType		
termReferenceType		
textureListType		
basicTimePointType		
timePointType		
unsigned1		
unsigned3		
unsigned5		
unsigned6		
unsigned8		
unsigned12		
integerVector		

6.5.1.1 Example of a slide show track metadata

```
<?xml version="1.0" encoding="UTF-8"?>
<Mpeg7 xmlns="urn:mpeg:mpeg7:schema:2004" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:mpeg:mpeg7:schema:2004 version2schema_unbounded_stillregionfix.xsd">
  <DescriptionMetadata>
    <LastUpdate>2006-11-06T19:20:25+00:00</LastUpdate>
    <Creator>
      <Role href="urn:mpeg:mpeg7:cs:RoleCS:2001:AUTHOR">
        <Name>Author</Name>
      </Role>
      <Agent xsi:type="PersonType">
        <Name>
          <GivenName>Akio</GivenName>
          <FamilyName>Yamada</FamilyName>
        </Name>
      </Agent>
    </Creator>
    <CreationTime>2006-11-06T19:20:25+00:00</CreationTime>
  </DescriptionMetadata>
  <DescriptionUnit xsi:type="ContentCollectionType" name="Yuto's 6th birthday">
    <!-- ########## ROOT COLLECTION ########## !-->
    <CreationInformation>
      <Creation>
        <Title>
        </Title>
        <TitleMedia>
          <TitleImage>
            <!-- Thumbnail image (<InlineMedia>) here  -->
          </TitleImage>
```

```
        </TitleMedia>
      </Creation>
    </CreationInformation>
    <TextAnnotation>
      <FreeTextAnnotation>Birthday Party</FreeTextAnnotation>
    </TextAnnotation>
    <Content xsi:type="ImageType" id="photo1">
      <!-- Item level metadata for photo 1 -->
      <Image>
        <MediaInformation>
          <MediaProfile>
            <MediaFormat>
              <Content href="urn:mpeg:mpeg7:cs:ContentCS:2001:4.1">
                <Name>Image</Name>
              </Content>
              <FileFormat href="urn:mpeg:mpeg7:cs:FileFormatCS:2001:1">
                <Name>jpeg</Name>
              </FileFormat>
              <FileSize>138474</FileSize>
              <VisualCoding><Frame height="320" width="240"/></VisualCoding>
            </MediaFormat>
            <MediaInstance>
              <InstanceIdentifier></InstanceIdentifier>
              <MediaLocator>
                <MediaUri>#ffp(item_ID=1)</MediaUri>
              </MediaLocator>
            </MediaInstance>
          </MediaProfile>
        </MediaInformation>
        <CreationInformation>
          <Creation>
            <Title>Birthday cake</Title>
            <CreationCoordinates>
              <Date>
                <TimePoint>2005-11-05T15:20:37+09:00</TimePoint>
              </Date>
            </CreationCoordinates>
          </Creation>
        </CreationInformation>
        <TextAnnotation>
          <KeywordAnnotation>
            <Keyword>Cake</Keyword>
          </KeywordAnnotation>
        </TextAnnotation>
        <VisualDescriptionScheme xsi:type="StillRegionFeatureType">
          <ColorLayout>
            <YDCCoeff>48</YDCCoeff>
            <CbDCCoeff>34</CbDCCoeff>
            <CrDCCoeff>32</CrDCCoeff>
            <YACCoeff5>12 10 13 9 10</YACCoeff5>
            <CbACCoeff5>14 15 8 7 3</CbACCoeff5>
            <CrACCoeff5>16 12 9 6 6</CrACCoeff5>
          </ColorLayout>
          <Edge>
            <BinCounts>2 6 4 4 2 1 7 5 3 2 1 6 4 2 2 2 5 4 5 3 1 5 5 6 5
              2 6 5 4 4 1 6 4 4 4 0 6 3 5 2 1 5 5 6 6 4 2 3 6 7 3 2 5 5 7
              3 2 4 4 7 1 5 6 4 6 1 5 7 4 5 1 6 4 6 5 1 3 4 7 6
            </BinCounts>
          </Edge>
        </VisualDescriptionScheme>
      </Image>
    </Content>
    <Content xsi:type="ImageType" id="photo2">
      <!-- Item level metadata for photo 2 -->
      <Image>
        <MediaInformation>
          <MediaProfile>
            <MediaInstance>
              <InstanceIdentifier></InstanceIdentifier>
              <MediaLocator>
                <MediaUri>#ffp(item_ID=2)</MediaUri>
              </MediaLocator>
            </MediaInstance>
          </MediaProfile>
        </MediaInformation>
        <CreationInformation>
```

```
          <Creation>
            <Title>Blowing out candles on the cake </Title>
            <Creator>
              <!-- Person appearing in the photo  -->
              <Role href="urn:mpeg:mpeg7:cs:RoleCS:2001:ACTOR"/>
              <Agent xsi:type="PersonType">
                <Name>
                  <GivenName>Yuto</GivenName>
                  <FamilyName>Yamada</FamilyName>
                </Name>
              </Agent>
            </Creator>
            <CreationCoordinates>
              <Date>
                <TimePoint>2005-11-05T15:23:40+09:00</TimePoint>
              </Date>
            </CreationCoordinates>
          </Creation>
        </CreationInformation>
        <TextAnnotation>
          <KeywordAnnotation>
            <Keyword>Cake</Keyword>
          </KeywordAnnotation>
        </TextAnnotation>
      </Image>
    </Content>
    <Content xsi:type="ImageType" id="photo3">
      <!-- Item level metadata for photo 3 -->
      <Image>
        <MediaInformation>
          <MediaProfile>
            <MediaInstance>
              <InstanceIdentifier></InstanceIdentifier>
              <MediaLocator>
                <MediaUri>#ffp(item_ID=3)</MediaUri>
              </MediaLocator>
            </MediaInstance>
          </MediaProfile>
        </MediaInformation>
        <CreationInformation>
          <Creation>
            <Title>Yuto's smile</Title>
            <Creator>
              <!-- Person appearing in the photo  -->
              <Role href="urn:mpeg:mpeg7:cs:RoleCS:2001:ACTOR"/>
              <Agent xsi:type="PersonType">
                <Name>
                  <GivenName>Yuto</GivenName>
                  <FamilyName>Yamada</FamilyName>
                </Name>
              </Agent>
            </Creator>
            <CreationCoordinates>
              <Date>
                <TimePoint>2005-11-05T15:27:50+09:00</TimePoint>
              </Date>
            </CreationCoordinates>
          </Creation>
        </CreationInformation>
        <TextAnnotation>
          <KeywordAnnotation>
            <Keyword>Smile</Keyword>
          </KeywordAnnotation>
        </TextAnnotation>
      </Image>
    </Content>
    <Content xsi:type="ImageType" id="photo4">
      <!-- Item level metadata for photo 4 -->
      <Image>
        <MediaInformation>
          <MediaProfile>
            <MediaInstance>
              <InstanceIdentifier></InstanceIdentifier>
              <MediaLocator>
                <MediaUri>#ffp(item_ID=4)</MediaUri>
              </MediaLocator>
```

```
            </MediaInstance>
          </MediaProfile>
        </MediaInformation>
        <CreationInformation>
          <Creation>
            <Title>Tinted autumn leaves</Title>
            <CreationCoordinates>
              <Date>
                <TimePoint>2004-11-20T10:10:50+09:00</TimePoint>
              </Date>
            </CreationCoordinates>
          </Creation>
        </CreationInformation>
      </Image>
    </Content>
    <Content xsi:type="ImageType" id="photo5">
      <!-- Item level metadata for photo 5 -->
      <Image>
        <MediaInformation>
          <MediaProfile>
            <MediaInstance>
              <InstanceIdentifier></InstanceIdentifier>
              <MediaLocator>
                <MediaUri>#ffp(item_ID=5)</MediaUri>
              </MediaLocator>
            </MediaInstance>
          </MediaProfile>
        </MediaInformation>
        <CreationInformation>
          <Creation>
            <Title>A deer running through woods</Title>
            <CreationCoordinates>
              <Date>
                <TimePoint>2004-11-20T13:04:00+09:00</TimePoint>
              </Date>
            </CreationCoordinates>
          </Creation>
        </CreationInformation>
      </Image>
    </Content>
    <ContentCollection name="Birthday Party1">
      <!-- ########## SUB COLLECTION 1 ########## -->
      <CreationInformation>
        <Creation>
          <Title>
          </Title>
          <TitleMedia>
            <TitleImage>
              <!-- Thumbnail image (<InlineMedia>) here  -->
            </TitleImage>
          </TitleMedia>
          <Creator>
            <Role href="urn:mpeg:mpeg7:cs:RoleCS:2001:AUTHOR"/>
            <Agent xsi:type="PersonType">
              <Name>
                <GivenName>Akio</GivenName>
                <FamilyName>Yamada</FamilyName>
              </Name>
            </Agent>
          </Creator>
          <CreationCoordinates>
            <Date>
              <TimePoint>2005-11-05T15:20+09:00</TimePoint>
              <Duration>PT10M</Duration>
            </Date>
          </CreationCoordinates>
        </Creation>
      </CreationInformation>
      <TextAnnotation>
        <FreeTextAnnotation> Birthday Party</FreeTextAnnotation>
        <KeywordAnnotation>
          <Keyword>EventCollection</Keyword>
        </KeywordAnnotation>
        <KeywordAnnotation>
          <Keyword>Party</Keyword>
        </KeywordAnnotation>
```

```
      </TextAnnotation>
      <ContentRef idref="photo1" />
      <ContentRef idref="photo2" />
      <ContentRef idref="photo3" />
    </ContentCollection>
    <ContentCollection name="Nature in autumn season">
      <!-- ########## SUB COLLECTION 2 ########## -->
      <CreationInformation>
        <Creation>
          <Title></Title>
          <Creator>
            <Role href="urn:mpeg:mpeg7:cs:RoleCS:2001:AUTHOR"/>
            <Agent xsi:type="PersonType">
              <Name>
                <GivenName>John</GivenName>
                <FamilyName>Smith</FamilyName>
              </Name>
            </Agent>
          </Creator>
        </Creation>
      </CreationInformation>
      <ContentRef idref="photo4" />
      <ContentRef idref="photo5" />
    </ContentCollection>
  </DescriptionUnit>
</Mpeg7>
```

6.5.2 Metadata for the audio track

The metadata for the audio track provides the song title, name of the artist, album title, year and genre of the audio content.

For the Musical slide show application format, MPEG-7 Creation and Semantics description tools are used for data description. Equivalent metadata are used in current MP3 applications (using ID3 tags) as shown in Table 6.

Table 6 — Semantics for the audio track metadata

Tag Name	ID3 V1	Semantics
CreationInformation/Creation/Creator	Artist	Artist performing the song
CreationInformation/Creation/Title[@type=”albumTitle”]	Album	Title of the album
CreationInformation/Creation/Title[@type=”songTitle”]	Song Title	Title of the song
CreationInformation/CreationCoordinates/Date/TimePoint (*Recording date.*)	Year	Year of the recording
CreationInformation/Creation/Abstract/FreeTextAnnotation	Comment	Any comment of any length
Semantics/SemanticBase[@xsi:type=”SemanticStateType”]/AttributeValuePair	Track	CD track number of song
CreationInformation/Classification/Genre[@href=”urn:id3:v1:4”] CreationInformation/Classification/Genre[@href=”urn:id3:v1:4”]/Term[@termID=”urn:id3:v2:Eurodisco”] CreationInformation/Classification/Genre[@href=”urn:id3:v1:4”] CreationInformation/Classification/Genre[@type=”secondary][@href=”urn:id3:v2:Eurodisco”]	Genre	ID 3 V1.1 Genre ID 3 V2 Genre(4)(Eurodisco)

6.5.2.1 Example of an audio track metadata

```
<Description xsi:type="CreationDescriptionType">
    <CreationInformation>
    <Creation>
        <Title type="songTitle">I am three years old</Title>
        <Title type="albumTitle">Birthday Songs</Title>
        <Abstract>
            <FreeTextAnnotation></FreeTextAnnotation>
        </Abstract>
        <Creator>
            <Role href="urn:mpeg:mpeg7:RoleCS:2001:PERFORMER"/>
            <Agent xsi:type=”PersonType”>
                <Name>
                    <FamilyName>Cha</FamilyName>
                    <GivenName>Soeun</GivenName>
                </Name>
            </Agent>
        </Creator>
        <CreationCoordinates>
            <Date>
                <TimePoint>2007-02-22T14:21+00:00</TimePoint>
            </Date>
        </CreationCoordinates>
    </Creation>
    <Classification>
        <Genre href="urn:id3:cs:ID3genreCS:v1:12">
            <Name>Other</Name>
        </Genre>
    </Classification>
    </CreationInformation>
</Description>
<Description xsi:type="SemanticDescriptionType">
    <Semantics>
        <SemanticBase xsi:type="SemanticStateType">
            <AttributeValuePair>
                <Attribute>
                    <TermUse href="urn:mpeg:maf:cs:musicplayer:CollectionElementsCS:2007:assetNum"/>
                </Attribute>
                <IntegerValue>2</IntegerValue>
            </AttributeValuePair>
            <AttributeValuePair>
                <Attribute>
                    <TermUse href="urn:mpeg:maf:cs:musicplayer:CollectionElementsCS:2007:assetTot"/>
                </Attribute>
                <IntegerValue>5</IntegerValue>
            </AttributeValuePair>
        </SemanticBase>
    </Semantics>
</Description>
```


7 Scope of Protected Musical slide show application format

7.1 Overview

The “Protected Musical slide show application format”, described in the following clauses 7 and 8, builds on the Musical slide show application format as described above. It adds content protection for MP3 audio, JPEG images, 3GPP Timed Text, and LASeR script animation with flexible protection tool selection and key management components.

7.2 Creating Protected Musical slide show application format

Creating a protected Musical slide show application format file involves formatting different types of media data, defining the protection and license information, and storing them into an MPEG-4 file format. Based on the Musical slide show application format system architecture described in subclause 5.2, the protection module is included in the Creator to protect the resources based on the protection and license description.

Figure 15 shows an example of protected 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. These resources are described in structured way using MPEG-21 Digital Item Declaration Language (DIDL) (ISO/IEC 21000-2), while the protection and license information for the protected resource is described using MPEG-21 Intellectual Property Management and Protection (IPMP) Components (ISO/IEC 21000-4) and MPEG-21 Rights Expression Language (REL) (ISO/IEC 21000-5).

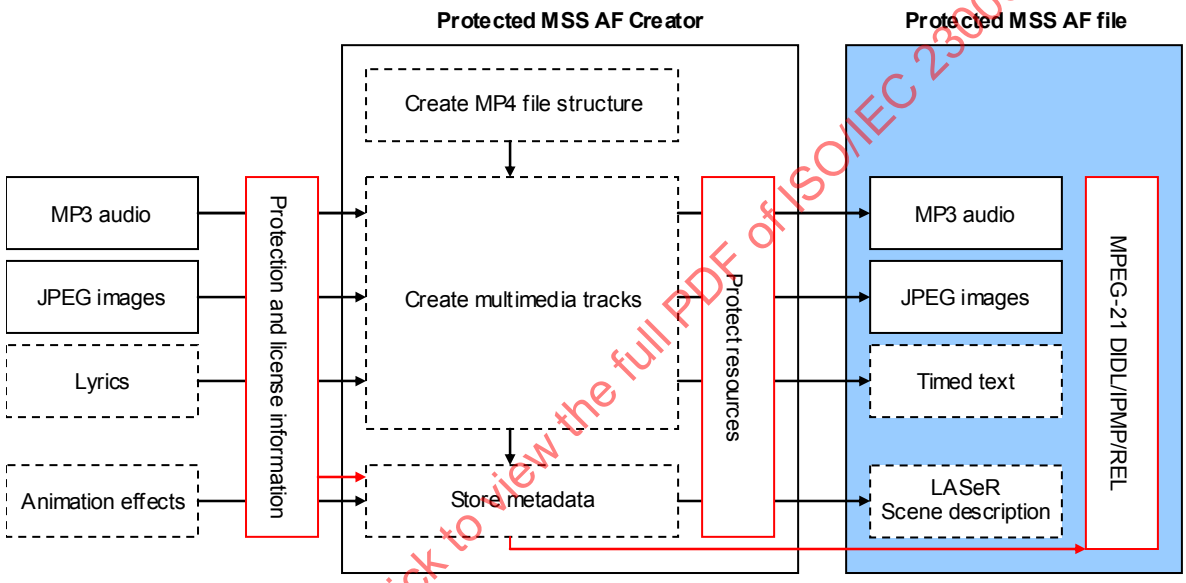


Figure 15 — Example of Protected Musical slide show application format creator system architecture

7.3 Metadata for Protection

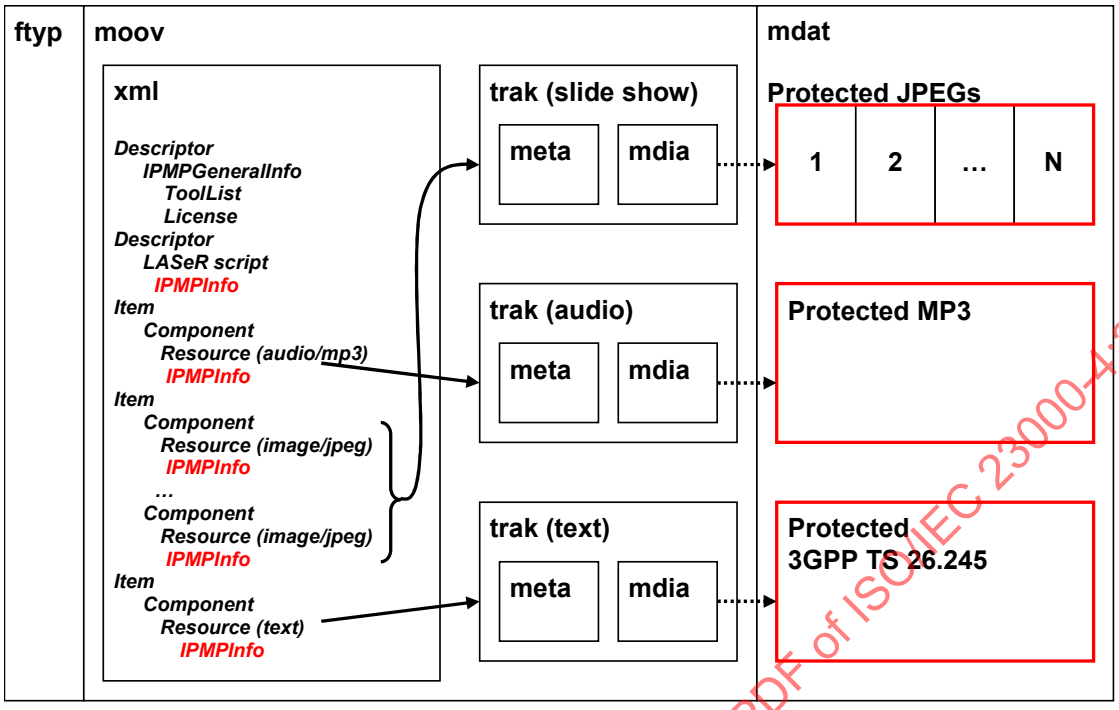


Figure 16 — How the metadata points to the protected resources

Figure 16 shows the information described by IPMP and what resources are protected. The mechanism of signalling protection for the resources with the metadata is as follows: each of MP3 audio and 3GPP Timed text is described as one item; the collection of JPEG images (the slide show) is described as one item that contains individual JPEG image as a component; and LASeR script for animation is described as *Descriptor*. If the resource is protected, its description will be described as *ProtectedAsset* using IPMP DIDL description scheme.

The IPMPDIDL metadata contains *Descriptor* that contains *IPMPGeneralInfo*. It is recommended that the *Descriptor* is defined at the beginning of the IPMPDIDL metadata. The *IPMPGeneralInfo* contains:

– ToolList, as defined in MPEG-21 IPMP Components Base Profile.

– Container for licenses. The license information is as described by MPEG-21 REL MAM Profile

With this specification, the collection and item level metadata will be left intact. The MPEG-21 will not be used to describe any information that has been described by the MPEG-7 metadata in both audio and slide show 'trak's. Instead, it will be only used to describe the structure and the governance of the multimedia content inside the AF.

Several protection scenarios are possible:

7.3.1 Protecting individual item

Each resource: MP3 audio, JPEG image, 3GPP Timed Text, and LASeR script animation can be protected individually. Figure 17 shows how the protection can be described in the metadata for MP3 audio only, while Figure 18 shows how the protection can be described for the LASeR script only.

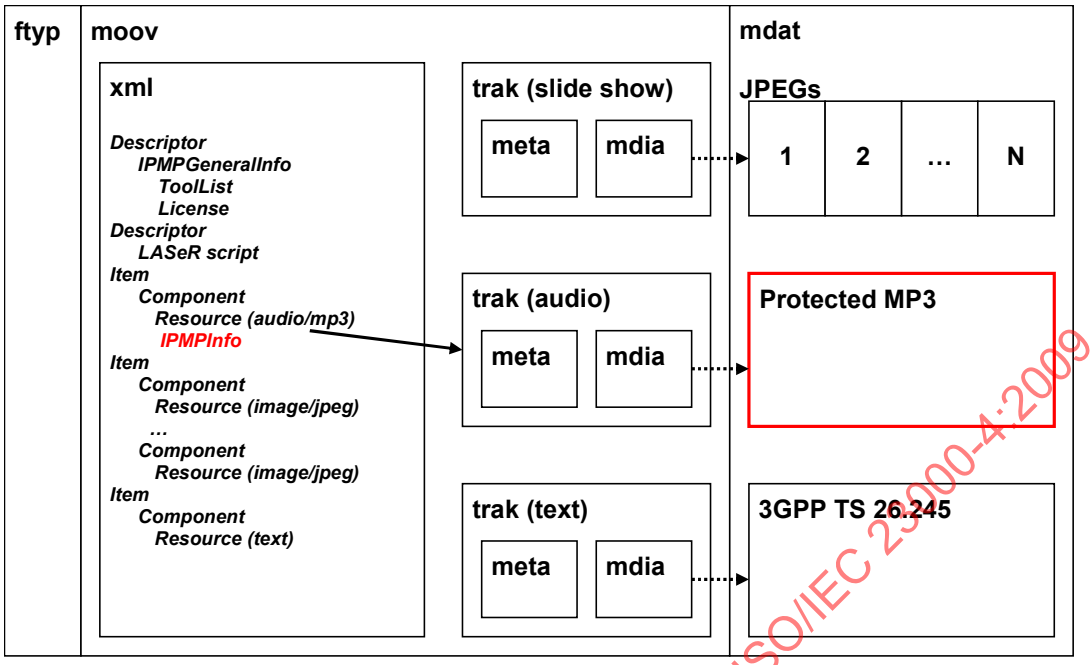


Figure 17 — Protecting MP3 audio only

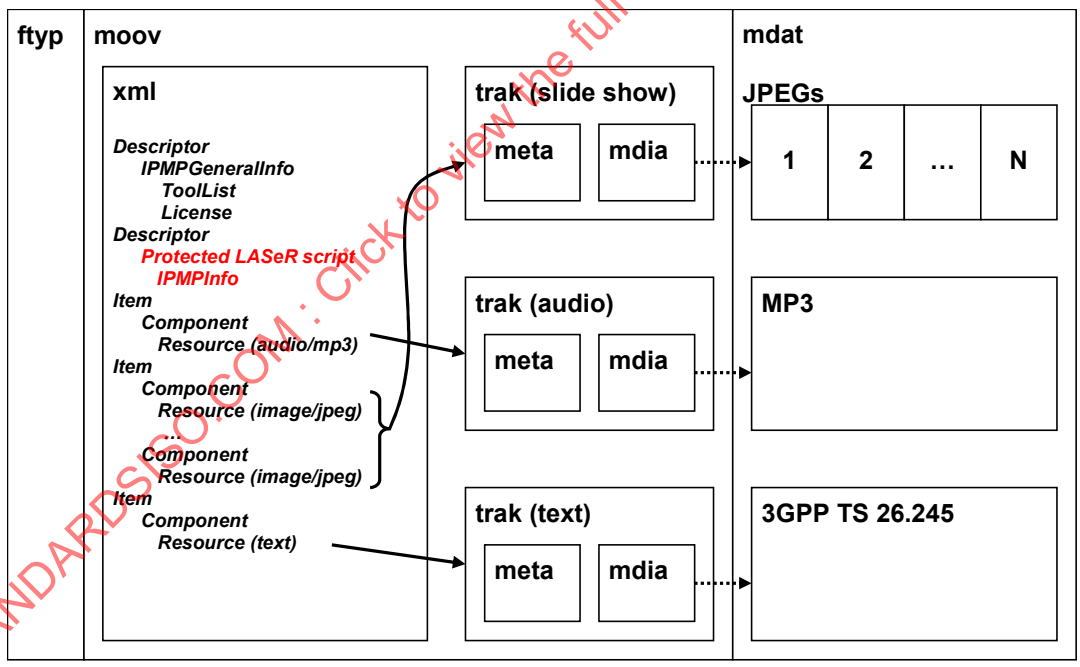


Figure 18 — Protecting LASeR script only

7.3.2 Protecting combination of individual items

The whole resource or its combination (e.g. MP3 audio and JPEG images, or JPEG images and its animation) can be protected at the same time. Figure 19 shows the protection description to protect JPEG images and slide show animation.

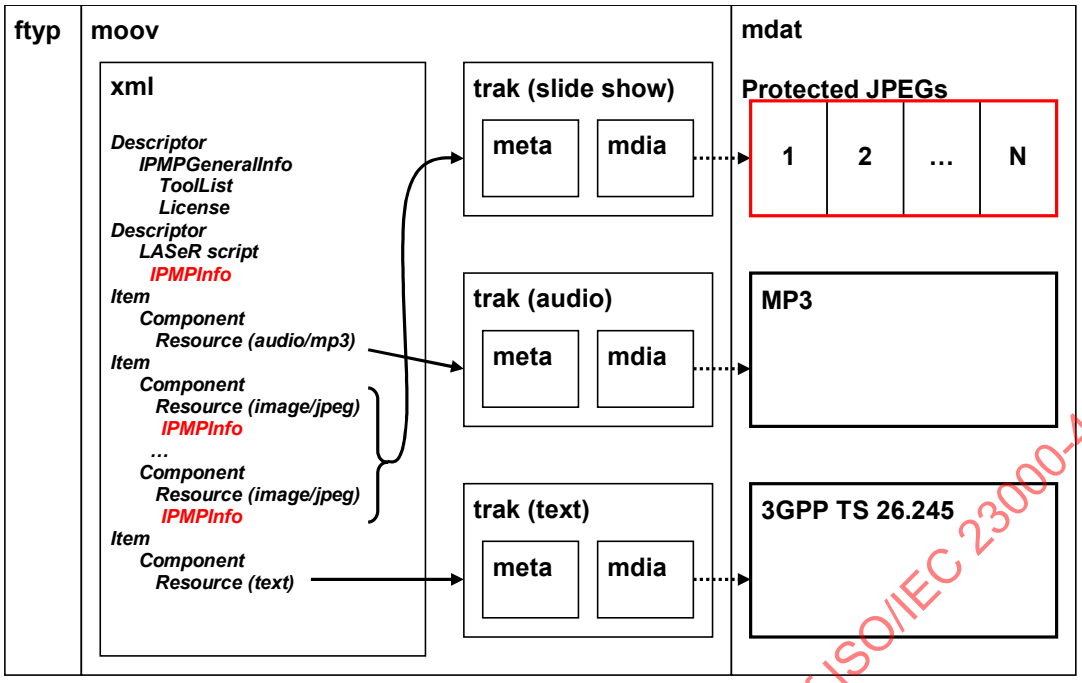


Figure 19 — Protecting JPEG images and slide show animation

7.3.3 Protecting one or more JPEG images

Each JPEG images inside the image track can be protected individually or collectively. Figure 20 shows the protection description for protecting two JPEG images.

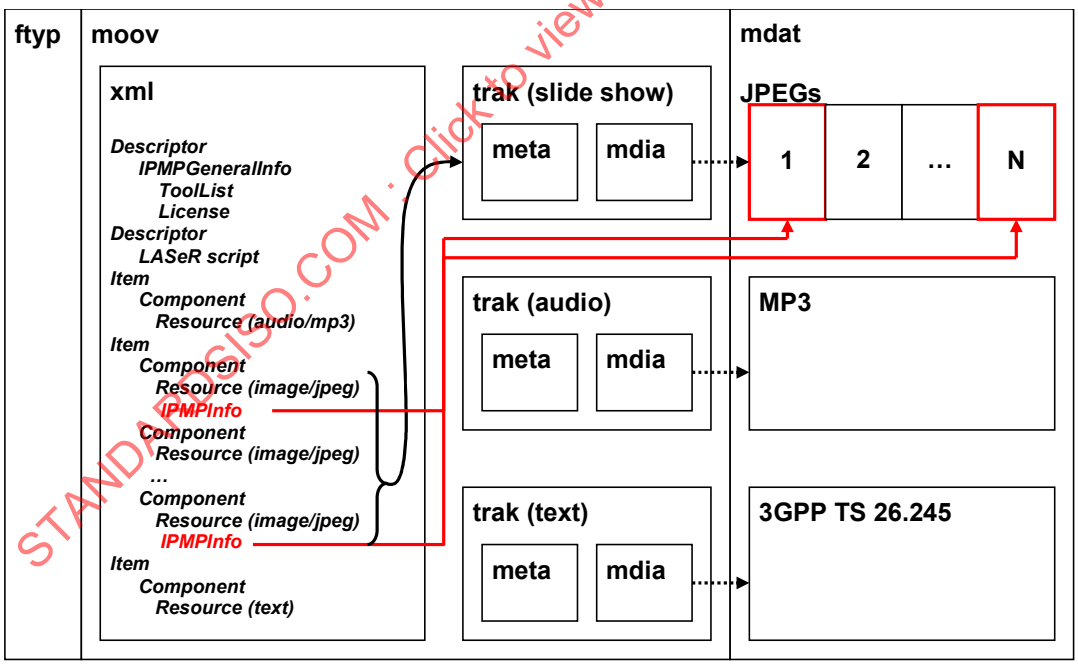


Figure 20 — Protecting two JPEG images

7.3.4 Protecting certain segment of resource

Using MPEG-21 Fragment Identifier (ISO/IEC 21000-17), it is possible to protect specific segment of the content, such as specifically defined rectangle region of JPEG image or specific segment of MP3 audio which bytes duration are defined, as shown in Figure 21 and Figure 22, respectively.

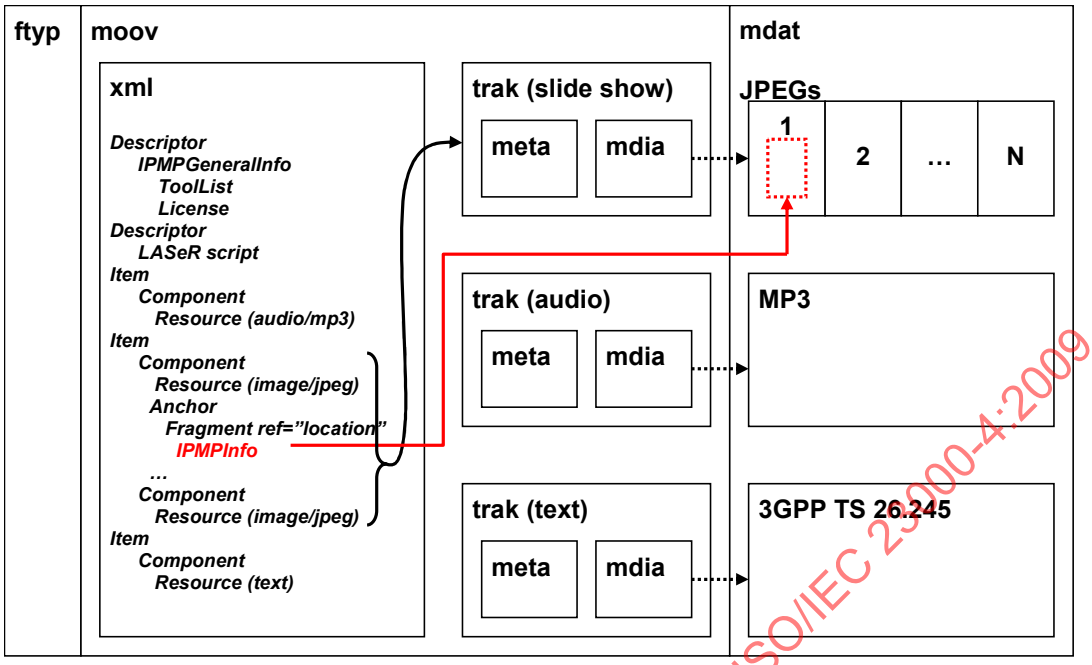


Figure 21 — Protecting specific region in JPEG image

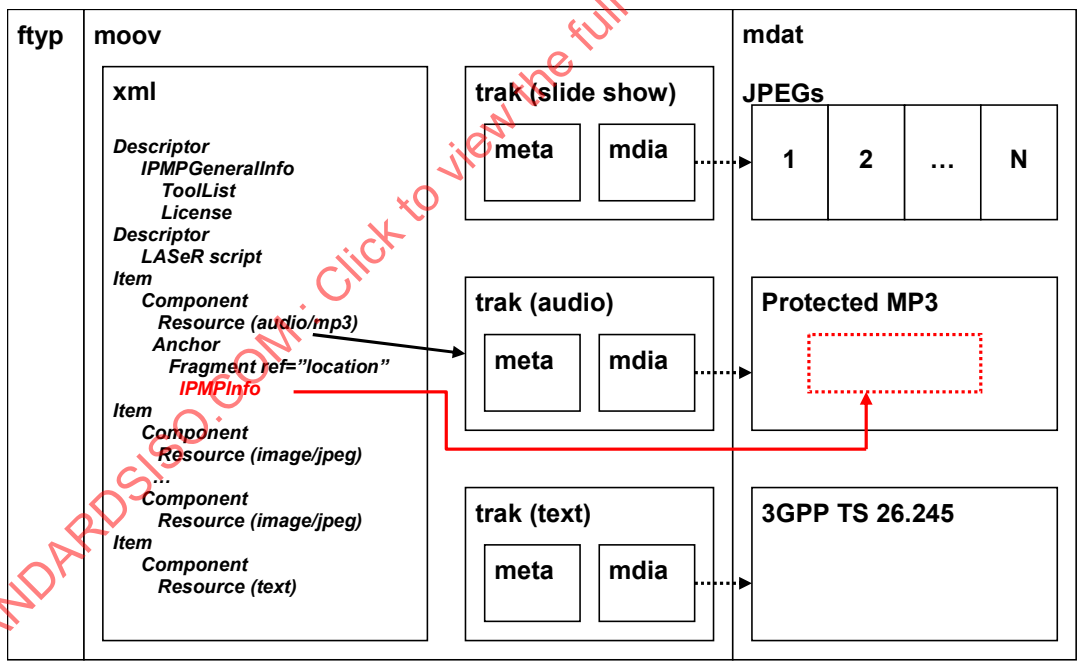


Figure 22 — Protecting specific segment in MP3 audio

Moreover, it is also possible to combine the protection to certain segment of resource with protection of other content, e.g., protecting certain segment of MP3 audio and protecting JPEG images that are synchronized to the timestamp of the protected segment of MP3.

Annex C presents the metadata instantiation examples for all protection scenarios. In the examples, the use of metadata to signal AES-128 encryption to protect the contents is presented.

7.4 Playback

The two rendering modes for Musical slide show application format: “Basic” mode and “Enhanced” mode as described in subclause 5.4 are still available in protected Musical slide show application format. In this case, the resource shall be unprotected prior to rendering.

As the player in Basic mode may ignore the LASeR script for animation, the protection for LASeR script may also be ignored in this mode.

8 Overview of Basic Standards for Protection

8.1 MPEG-21 IPMP Components Base Profile

MPEG-21 IPMP Components Base Profile is aimed at supporting use cases in widespread use in the area of commercial content distribution. The Base Profile purposely provides a limited scope in order to facilitate the implementation in devices with limited computational/storage capabilities. It provides sufficient functionality to support current and emerging practices for distribution of commercial content, with a special focus on entertainment content such as movies and music, while reducing the requirements on end devices (e.g. footprint, memory usage, computational power, storage).

The MPEG-21 IPMP Components Base Profile adopts and in some cases restricts the MPEG-21 IPMP Components specification.

8.1.1 IPMP Digital Item Declaration Language

The MPEG-21 IPMP Components Base Profile includes all the elements in the IPMP DIDL schema. A set of these elements are a representation of the DID model that allows for inclusion of governance information as defined in [9].

8.1.2 IPMP General Info Descriptor

The IPMP General Info Descriptor in the MPEG-21 IPMP Components Base Profile restricts the corresponding element defined in ISO/IEC 21000-4 as follows:

- The *ToolList* shall contain at most one instance of *ToolDescription* (i.e. cardinality is 1).
- The *ToolDescription* element shall include the *IPMPToolID* element and optionally the Remote element, i.e. this Base Profile provides no support for Inline tool.
- The tool is assumed to be ready-to-be-used on the terminal; hence there is no need to carry the *ConfigurationSettings* element.
- The *LicenseCollection* element can contain any number of *RightsDescriptor* elements (in case there are multiple assets in the digital item), although in most usage instances a single *RightsDescriptor* element is likely to be used. The *RightsDescriptor* in the Base Profile excludes the possibility of having an *IPMPInfoDescriptor* child.

8.1.3 IPMP Info Descriptor

The IPMP Info Descriptor in the MPEG-21 IPMP Components Base Profile restricts the corresponding elements defined in ISO/IEC 21000-4 as follows:

- The *Tool* element shall have no attributes (since there is at most one *Tool*, order is no longer relevant).
- There is no support to describe the tool in the Info Descriptor. The tool will only refer to the one that has been defined in the IPMP General Info Descriptor; hence a Tool shall include a *ToolRef* element.

- There is no license information place holder under the IPMP Info Descriptor as all the necessary license shall be put under *LicenseCollection* in the IPMP General Info Descriptor.
- There is no support for hierarchical protection to the *InitializationSettings* element.

8.2 MPEG-21 REL MAM Profile

The MPEG-21 REL MAM Profile (Mobile And optical Media Profile) is meant to be used in simple multimedia application domains. This simple application may sometimes but not always run on limited devices/systems. To this end, the MAM Profile restricts the MPEG-21 Rights Expression Language specification which includes REL Core, REL Standard Extension, and REL Multimedia Extension.

9 Usage of File Format Brands

The 'ftyp' box of the ISO Base Media File Format contains a list of "brands" that are used as identifiers in the file format. To enable player applications to easily identify files which are compliant to this AF specification, specific brand identifiers are defined. These brands are used in the compatible-brands list in addition to other appropriate brand types, like "iso2", "mp42" or "mp21".

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

- 'mss2'

It follows the brand of the Musical slide show application format specified in subclause 5.3, by altering the number of brands into '2' to specify the '2nd Edition'. The conformance of the AF application is described in clause 10.

10 Conformance and List of Technologies

This clause describes the conformance from the Music player application to parse the Musical slide show application format file with brands described in clause 9.

The application shall be able to read the MPEG-21 DID metadata stored in movie-level 'meta' box as hierarchical structure described in sub-clause 7.3. By parsing and executing the protection scheme to unprotect the protected contents (if any), the output of the devices shall be the same as the output of Musical slide show application format without the MPEG-21 DID metadata, for "Basic" or "Enhanced" mode.

Table 7 shows the comparison of technologies used in Musical slide show application format compared to this part of ISO/IEC 23000.

Table 7 – Technologies used in Musical slide show application format

Brand / Technologies	'mss1'	'mss2'
MPEG-1/2 Layer 3 Audio	✓	✓
JPEG Images	✓	✓
MPEG-7 metadata	✓	✓
LASeR script for animation	✓	✓
3GPP TS 26.245 timed text data	✓	✓
MPEG-21 DID		✓
MPEG-21 IPMP Components Base Profile		✓
MPEG-21 REL MAM Profile		✓
MPEG-21 Fragment Identifier		✓

Annex A
(informative)

Use cases of Musical slide show application format

A.1 General

The following section describes the use cases for the Musical slide show application format.

A.2 Personal slide show application

Various single still images are used as a source in the personal slideshow application scenario. The user arranges the images on a timeline and generates the timing information for the presentation of the still images.

The images may be rescaled in size to optimize the resolution for the target output device: higher resolutions for presentations optimized for computer displays, lower resolutions optimized for mobile devices. Either no transitions between the images are used, or transition effects are applied at the presentation time.

An audio soundtrack and a text track (for sub-titles, annotations or intro text) are used to enrich the presentation. If metadata describing the images is available, it is stored so that the relationship of the metadata and the image is maintained.

The timed and optionally rescaled images, the metadata, the audio soundtrack and the timed text are stored to a file.

A.3 Photo-Music album application

Many of the portable multimedia devices have display capabilities; therefore, people are able to listen to their favourite songs while viewing images (e.g. album covers, photos inside the CD booklet) of their favourite singers.

The images are synchronized to the MP3 song tracks, and the lyrics are displayed in a synchronized way on the display window, so the people are better informed about the lyrics, and if they choose to sing-along to their favourite tunes, they could easily do so. Overall, more enjoyable music experiences are possible with such improved functionalities.

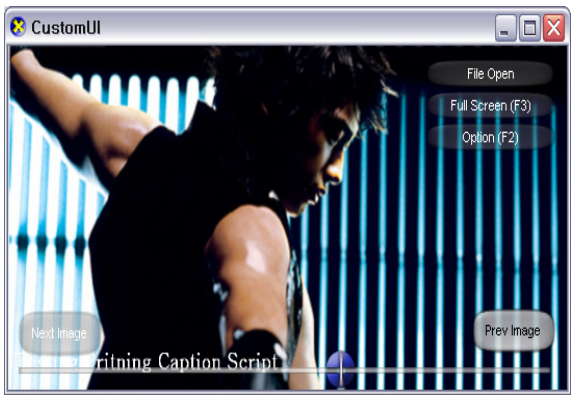


Figure A.1 — Example of a photo-music album application

A.4 Foreign language exercise materials

Many of the educational materials (e.g. foreign language tapes) that used to exist as tapes and CDs are now being replaced with the MP3 audio format. In schools, lecture materials are created by using MP3, and they are often downloaded by students for later use (e.g. for doing homework). Such materials are better understood if visual information is also conveyed.

In foreign language exercise applications, pictures are used to describe certain situations pertaining to the dialogue (e.g. TOEFL, TOEIC), and often questions are asked in forms of overlaid text on the screen.

Visual information is also helpful in learning how to pronounce a word. People are able to refer to the images that show the position of the tongue, or the precise location of the accent.

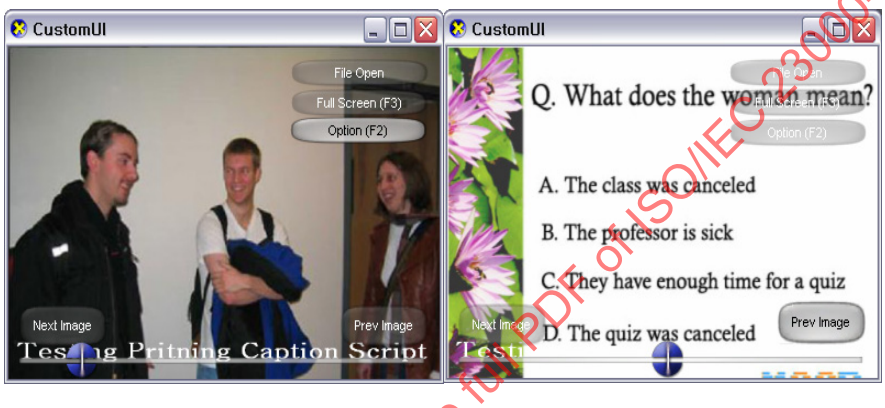


Figure A.2 — Example of a foreign language exercise material

A.5 Storytelling application

Children's story books (e.g. Aesop's fables) are often sold with audio tapes. In this digital version of a story book, the storytelling application utilizes multiple JPEG images with a synchronized MP3 audio track, and text data to create and deliver a visually and aurally enhanced story.

Figure A.3 — Example of a storytelling application

A.6 Karaoke application

Mary's favourite song is "Blue" by Leann Rimes. She wishes to listen, as well as sing along to her favourite song. Mary uses the karaoke function (= the device is removing the original voice from the music as a post-processing step) on her mobile phone. She performs the following:

- She chooses one of Leann Rimes' slideshows with animation effects
- She hears "Blue" without the voice of Leann Rimes
- She sees the lyrics
- She reads the lyrics and sings along to the music

A.7 Slide show + Karaoke application

Now, she wishes to listen, as well as sing along to her favourite song while viewing her images. Mary performs the following:

- She visits the mobile service provider's website
- She chooses "Blue" from the service provider's music selection
- She uploads her vacation photos
- She chooses the animation effects for the photos
- She synchronizes the photos to the music using a web-based tool provided by the service provider
- The service provider stores the synchronized photos as in the file together with the music track and the lyrics of the song
- She downloads the file to her mobile phone

After the download process, Mary plays the MSS file on her mobile phone. She performs the following:

- She listens to "Blue" by Leann Rimes
- She watches her slideshow with animation effects
- She sees the lyrics
- She reads the lyrics and sings along to the music

A.8 Online Musical slide show application format store

An online music store provides Protected Musical slide show application format files that contain the mp3 music, its album cover, a slide show showing the artist or some screenshots from the song's music video with animation, and the lyric of the songs. The music store has registered users in different user profiles. Each user profile has it own protection scheme to the content resources in the Protected Musical slide show application format files. The higher the user profile, the more content can be executed.

The following example shows how the music store handles the user's profile to define the protection scheme to the content.

John logs in to the music store's website. He has user's profile with privileges to play all content in the Protected Musical slide show application format file (MP3, JPEG slide show with LASeR animation and lyrics).

- John selects the file "I Wanna See You Now" by Leon Haines Band with album cover, JPEG slide show of the music's music video with LASeR animation and the song's lyric
- The application checks his user's profile
- Based on his user's profile, the application generates IPMP scheme to let John plays all the content with his un-protection key
- The generated IPMP scheme in form of metadata is packaged into the file
- John downloads the file and plays the content

Meanwhile, Maggie also logs in to the music store's website. She has user's profile to play only the MP3, and JPEG slide show without LASeR animation

- Maggie selects the same file as John did
- The application checks her user's profile
- Based on her user's profile, the application generates IPMP schema to let Maggie only to play MP3 and the JPEG slide show without animation.
- The generated IPMP scheme in form of metadata is packaged into the same file
- Maggie downloads the file and plays the content. However, since her un-protection key only valid to the MP3 and the JPEG slide show, she cannot see the animation and the lyrics

A.9 Storytelling content provider

A storytelling company distributes their children story contents using Protected Musical slide show application format in their website. Users can freely download the Protected Musical slide show application format file and play it in their device. However, the company can limit the number of images to be displayed in the file unless for registered users who already have the un-protection key to the whole contents of the slide show.

In this way, the company can introduce their children story content for the Musical slide show application format application by letting the users to evaluate the content before they are decided to register to obtain the un-protection key to see the whole slide show.

The following example shows how users evaluate the story-telling content

- James browses the storytelling company's website
- He selects the "Red Little Riding Hood" file, and downloads it
- He plays the file that shows first three images from the story while listening to the story.
- After three images, the application asks for un-protection key
- Since James had register for the full content, he unlocks the protection with his un-protection key and sees the remaining images

A.10 Protected foreign language exercise

A Protected Musical slide show application format file is used in a foreign language exercise applications that contains pictures related to the dialogue or questions. Users can obtain the file for free, but can only evaluate 10 or 15 questions in the exercise (that is, users can only see 15 pictures and hear the corresponding 15 dialogue or questions). To obtain the remaining contents and opportunities to see the total score for answering the questions, users can obtain the un-protection key from the file provider.

The following example shows how users evaluate the exercise file

- Maggie downloads the TOEFL exercise file from the website. It has 30 dialogue questions with corresponding images
- She plays the file for 15 dialogue questions with images
- After 15 questions, the application asks for her un-protection key
- Since Maggie has the key, she puts the key in the application, and plays the remaining 15 questions and got her score in the end

Annex B
(informative)

Examples of LASeR description element usage

B.1 General

LASeR description elements allow attributes of a referenced object (e.g. an image) to be animated over a specified period of time.

The following section contains XML examples of using LASeR description elements related to the basic transition effects defined for the Musical slide show application format.

B.2 Examples of basic transition effects

B.2.1 Example of opacity control using *animate*

```
<lsru:NewScene>
    <svg width="267" height="200" viewBox="0 0 267 200" version="1.1" baseProfile="tiny">
        <image width="267" height="200" xlink:href="balloons.jpg">
            <animate attributeName="opacity" calcMode="linear" values="1.00; 0.00"
                dur="3" repeatDur="indefinite"/>
        </image>
    </svg>
</lsru:NewScene>
```

The opacity level of the referenced image ("balloons") gradually decreases over a period of 3 seconds.

B.2.2 Example of scaling using *animateTransform*

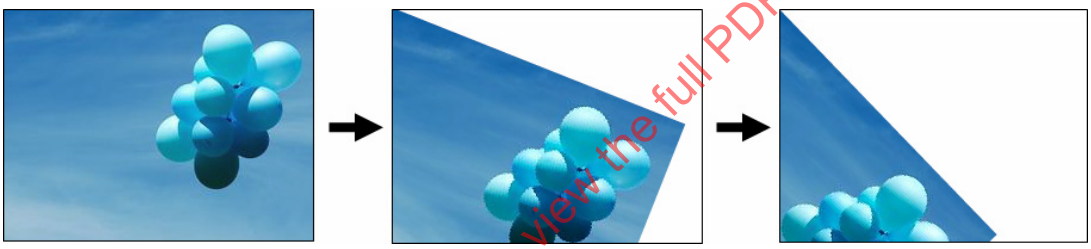
```
<lsru:NewScene>
    <svg width="267" height="200" viewBox="0 0 267 200" version="1.1" baseProfile="tiny">
        <image width="267" height="200" xlink:href="balloons.jpg">
            <animateTransform attributeName="transform" type="scale" additive="sum"
                calcMode="linear" values="1.00 1.00; 0.00 1.00" dur="5" repeatDur="indefinite"/>
        </image>
    </svg>
</lsru:NewScene>
```

The scale factor decreases in the x-direction (from 1.0 to 0.0) over a period of 5 seconds.

B.2.3 Example of rotation using *animateTransform*

```
<lsru:NewScene>
    <svg width="267" height="200" viewBox="0 0 267 200" version="1.1" baseProfile="tiny">
        <image width="267" height="200" xlink:href="balloons.jpg">
            <animateTransform attributeName="transform" type="rotate" additive="sum"
                calcMode="linear" values="0; 90" dur="3" repeatDur="indefinite"/>
        </image>
    </svg>
</lsru:NewScene>
```

The image rotates in the clockwise direction (90 degrees) over a period of 3 seconds.



B.2.4 Example of skewing using *animateTransform*

```
<lsru:NewScene>
    <svg width="267" height="200" viewBox="0 0 267 200" version="1.1" baseProfile="tiny">
        <image width="267" height="200" xlink:href="balloons.jpg">
            <animateTransform attributeName="transform" type="skewX" additive="sum"
                calcMode="linear" values="0; 45" dur="3" repeatDur="indefinite"/>
        </image>
    </svg>
</lsru:NewScene>
```

The image is skewed in the x-direction (from 0 to 45 degrees) over a period of 3 seconds.

B.2.5 Example of translation using *animateTransform*

```
<lsru:NewScene>
    <svg width=”267” height=”200” viewBox=”0 0 267 200” version=”1.1” baseProfile=”tiny”>
        <image width=”267” height=”200” xlink:href=”balloons.jpg”>
            <animateTransform attributeName=”transform” type=”translate” additive=”sum”
                    calcMode=”linear” from=”0 0” to=”100 100” dur=”5” repeatDur=”indefinite”/>
        </image>
    </svg>
</lsru:NewScene>
```

The image is translated from (x1, y1) = (0, 0) to (x2, y2) = (100, 100) over a period of 5 seconds.

B.2.6 Example of color transformation using *animateColor*

```
<lsru:NewScene>
    <svg width=”267” height=”200” viewBox=”0 0 267 200” version=”1.1” baseProfile=”tiny”>
        <!-- shape of a star -->
        <path id="star" d="M 61.42857,152.36218 L 53.00962,139.59064 L 37.7159,139.28559
            L 47.26076,127.33206 L 42.82486,112.6926 L 57.14286,118.07646
            L 69.69504,109.33383 L 68.99919,124.61476 L 81.19277,133.85098
            L 66.44471,137.91125 L 61.42857,152.36218 z " fill="blue">
            <animateColor attributeName="fill" from="blue" to="red" dur="3" repeatDur="indefinite"/>
        </path>
    </svg>
</lsru:NewScene>
```

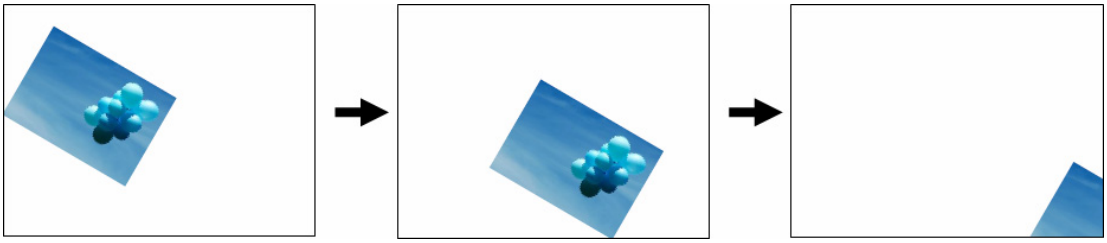
The “star” shape is formed using the *path* element. The color gradually changes from blue to red over a period of 3 seconds.

B.2.7 Example of object motion using *animateMotion*

```
<lsru:NewScene>
    <svg width="267" height="200" viewBox="0 0 267 200">
        <!-- motion path -->
        <path id="path1" d="M 200,0 C 0,0 200,100 250,120
                M 50,0 L 10,0 267,150
                M 50,200 L 50,200 200,0" fill="none" />

        <!-- image in motion -->
        <image width="120" height="86" xlink:href="balloons.JPG">
            <animateMotion dur="6" repeatCount="indefinite" rotate="auto">
                <mpath xlink:href="#path1"/>
            </animateMotion>
        </image>
    </svg>
</lsru:NewScene>
```

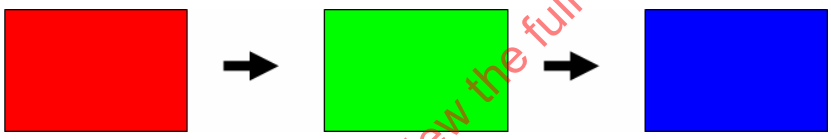
The image moves along a motion path created using the *path* element.



B.2.8 Example of setting attribute values using *set*

```
<lsru:NewScene>
    <svg width="267" height="200" viewBox="0 0 267 200" version="1.1" baseProfile="tiny">
        <rect width="267" height="200" fill="#000000">
            <set id="r" attributeName="fill" to="#ff0000" begin="0.0s; b.end" end="0.5s;b.end+0.5s"/>
            <set id="g" attributeName="fill" to="#00ff00" begin="r.end" end="r.end+0.5s"/>
            <set id="b" attributeName="fill" to="#0000ff" begin="g.end" end="g.end+0.5s"/>
        </rect>
    </svg>
</lsru:NewScene>
```

The color of the rectangular shape is first set to red, then to green, and finally to blue, every 0.5 second.



B.2.9 Example of using “shape” elements

```
<lsru:NewScene>
    <svg width="800" height="300" viewBox="0 0 800 300" version="1.1" baseProfile="tiny">
        <rect x="30" y="30" width="100" height="50" fill="rgb(0, 100, 190)" />
        <circle cx="200" cy="55" r="40" fill="rgb(0, 100,190)"/>
        <ellipse cx="320" cy="55" rx="50" ry="30" fill="rgb(0, 100,190)"/>
        <polygon points="450,20 400,90 500,90" fill="rgb(0, 100,190)"/>
        <polyline points="550,80 600,80 600,30, 650,30" stroke="rgb(0,100,190)" stroke-width="10"
            fill="none"/>
        <line x1="700" y1="30" x2="750" y2="90" stroke="rgb(0,100,190)" stroke-width="5"/>
    </svg>
</lsru:NewScene>
```

Annex C
(informative)

Examples for MPEG-21 Metadata

C.1 Protecting all resources

The protection to all resources in this part of ISO/IEC 23000 can be done by protecting each resource and describing the protection in the IPMP DIDL. Table C.1 shows the example of IPMP DIDL instantiation. The tool list is carried at the top of IPMP DIDL with necessary license collection. As described in `ipmpinfo:IPMPToolID` element, the AES-128 encryption tool can be signalled by defining the tool name as the identification tag. The linkage that refers the tool used for the content protection is described in the `localID` attribute in the `ToolList` element in `IPMPGeneralInfoDescriptor` and `localidref` attribute in the `IPMPInfoDescriptor`. An item in the IPMP DIDL represents a collection of MP3 audio, slide show (JPEG images) and timed text. The protected resources are described by the Resource element. All examples in this Annex are validated.

Table C.1 — IPMP DIDL instantiation example to protect all resources

```
<?xml version="1.0" encoding="UTF-8"?>
<DIDL xmlns="urn:mpeg:mpeg21:2002:02-DIDL-NS" xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS"
xmlns:ipmpdidl="urn:mpeg:mpeg21:2004:01-IPMPDIDL-NS" xmlns:ipmpinfo="urn:mpeg:mpeg21:2004:01-IPMPINFO-
BASE-NS" xmlns:mx="urn:mpeg:mpeg21:2003:01-REL-MX-NS" xmlns:r="urn:mpeg:mpeg21:2003:01-REL-R-NS"
xmlns:sx="urn:mpeg:mpeg21:2003:01-REL-SX-NS" xmlns:enc="http://www.w3.org/2001/04/xmlenc#"
xmlns:dsig="http://www.w3.org/2000/09/xmldsig#" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:mpeg:mpeg21:2003:01-REL-R-NS rel-r.xsd
urn:mpeg:mpeg21:2003:01-REL-MX-NS rel-mx.xsd
urn:mpeg:mpeg21:2004:01IPMPINFO-BASE-NS IPMPInfo-Profilev0.4.xsd
urn:mpeg:mpeg21:2004:01-IPMPDIDL-NS IPMPDIDL.xsd
urn:mpeg:mpeg21:2002:01-DII-NS dii.xsd
urn:mpeg:mpeg21:2002:02-DIDL-NS DIDL.xsd">
    <Container>
        <Descriptor>
            <Statement mimeType="text/xml">
                <ipmpinfo:IPMPGeneralInfoDescriptor>
                    <ipmpinfo:ToolList>
                        <ipmpinfo:ToolDescription localID="10">
                            <ipmpinfo:IPMPToolID>AES-128-CBC</ipmpinfo:IPMPToolID>
                        </ipmpinfo:ToolDescription>
                    </ipmpinfo:ToolList>
                    <ipmpinfo:LicenseCollection>
                        <ipmpinfo:RightsDescriptor>
                            <ipmpinfo:License>
                                <!-- license information -->
                            </ipmpinfo:License>
                        </ipmpinfo:RightsDescriptor>
                    </ipmpinfo:LicenseCollection>
                </ipmpinfo:IPMPGeneralInfoDescriptor>
            </Statement>
        </Descriptor>
        <Descriptor>
            <Statement mimeType="application/ipmp">
                <ipmpdidl:ProtectedAsset mimeType="application/laser">
                    <ipmpdidl:Identifier>
                        <dii:Identifier>IPMPId001</dii:Identifier>
                    </ipmpdidl:Identifier>
                    <ipmpdidl:Info>
                        <ipmpinfo:IPMPInfoDescriptor>
                            <ipmpinfo:Tool>
                                <ipmpinfo:ToolRef localidref="10"/>
                            </ipmpinfo:Tool>
                        </ipmpinfo:IPMPInfoDescriptor>
```

```
                </ipmpdidl:Info>
            <ipmpdidl:Contents>
                    <![CDATA[---------- animation script or code ---------]]>
            </ipmpdidl:Contents>
                </ipmpdidl:ProtectedAsset>
            </Statement>
        </Descriptor>
        <Item id="1">
            <Component>
                <Resource mimeType="application/ipmp">
                    <ipmpdidl:ProtectedAsset mimeType="audio/mp3">
                        <ipmpdidl:Identifier>
                            <dii:Identifier>IPMPId002</dii:Identifier>
                        </ipmpdidl:Identifier>
                        <ipmpdidl:Info>
                            <ipmpinfo:IPMPInfoDescriptor>
                                <ipmpinfo:Tool>
                                    <ipmpinfo:ToolRef localidref="10"/>
                                </ipmpinfo:Tool>
                            </ipmpinfo:IPMPInfoDescriptor>
                        </ipmpdidl:Info>
                        <ipmpdidl:Contents ref="#mp (/byte(0, 4550000)"/>
                    </ipmpdidl:ProtectedAsset>
                </Resource>
            </Component>
        </Item>
        <Item id="2">
            <Component>
                <Resource mimeType="application/ipmp">
                    <ipmpdidl:ProtectedAsset mimeType="image/jpeg">
                        <ipmpdidl:Identifier>
                            <dii:Identifier>IPMPId003</dii:Identifier>
                        </ipmpdidl:Identifier>
                        <ipmpdidl:Info>
                            <ipmpinfo:IPMPInfoDescriptor>
                                <ipmpinfo:Tool>
                                    <ipmpinfo:ToolRef localidref="10"/>
                                </ipmpinfo:Tool>
                            </ipmpinfo:IPMPInfoDescriptor>
                        </ipmpdidl:Info>
                        <ipmpdidl:Contents ref="#mp (/byte(4550000, 4550370)"/>
                    </ipmpdidl:ProtectedAsset>
                </Resource>
            </Component>
            <Component>
                <Resource mimeType="application/ipmp">
                    <ipmpdidl:ProtectedAsset mimeType="image/jpeg">
                        <ipmpdidl:Identifier>
                            <dii:Identifier>IPMPId004</dii:Identifier>
                        </ipmpdidl:Identifier>
                        <ipmpdidl:Info>
                            <ipmpinfo:IPMPInfoDescriptor>
                                <ipmpinfo:Tool>
                                    <ipmpinfo:ToolRef localidref="10"/>
                                </ipmpinfo:Tool>
                            </ipmpinfo:IPMPInfoDescriptor>
                        </ipmpdidl:Info>
                        <ipmpdidl:Contents ref="#mp (/byte(4550370, 4550892)"/>
                    </ipmpdidl:ProtectedAsset>
                </Resource>
            </Component>
        </Item>
        <Item id="3">
            <Component>
                <Resource mimeType="application/ipmp">
                    <ipmpdidl:ProtectedAsset mimeType="text/txt">
                        <ipmpdidl:Identifier>
                            <dii:Identifier>IPMPId005</dii:Identifier>
                        </ipmpdidl:Identifier>
                        <ipmpdidl:Info>
                            <ipmpinfo:IPMPInfoDescriptor>
                                <ipmpinfo:Tool>
                                    <ipmpinfo:ToolRef localidref="10"/>
                                </ipmpinfo:Tool>
                            </ipmpinfo:IPMPInfoDescriptor>
```