INTERNATIONAL STANDARD

ISO/IEC 23000-8

First edition
2008-12-01

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

Part 8: Portable video application format

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

Partie 8: Format pour application video portable

Reference number
ISO/IEC 23000-8:2008(E)

© ISO/IEC 2008

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

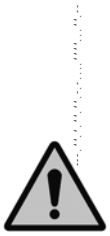
Not for Resale

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2008

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

© ISO/IEC 2008 – All rights reserved

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Overview of MPEG Standards 2
4.1 MPEG-4 Advanced Video Coding 2
4.2 MPEG-4 High Efficiency Advanced Audio Coding (HE-AAC) 2
4.3 MPEG-4 Bit Sliced Arithmetic Coding (BSAC) 2
4.4 ISO Base Media File Format 2
4.5 The ISO Base Media and MPEG-4 File Formats 3
4.6 MPEG-7 Multimedia Description Scheme 3
4.7 MPEG-4 LASeR 3

5 Using the Portable video application format 3
5.1 General 3
5.2 File Contents 3
5.3 Playback 5
5.4 File Structure 7

6 Technical features of Portable video application format 8
6.1 User Interface 8

Annex A (informative) **Use cases of Portable video application format** 11

Annex B (informative) **Usage of MPEG-7 MDS metadata in Portable video application format** 14

Annex C (informative) **Subtitle (captioning)** 26

Annex D (informative) **XML example of a menu scene description** 27

Bibliography 29

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale

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 23000-8 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 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*

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Introduction

The way people consume video content has evolved over the years from simply watching what is shown in television broadcasts and disc media (e.g. DVD) to consuming both commercial and non-commercial content over the network through file downloads and streaming services. New types of consumer electronic devices that are capable of decoding such files and data bitstreams are readily available today. With better processing power, bigger storage, and longer battery life, many of the portable multimedia players that we see in the market are capable of supporting various video and audio codecs, as well as file formats.

The portable video application format is designed for a mid-resolution video playback application in such portable devices. The file format is structured to contain audio and video data, and a scene description. The scene description contains data for graphics, animation, text, and temporal layout, and the data is rendered or generated as the user interface (e.g. a menu screen). Therefore, a single file may contain multiple video contents such as bonus tracks (e.g. “making-of” movie clips) and audio tracks for different languages, and a user can control them by using the menu.

With such capabilities, the portable video application format allows users to consume or create video contents that provide a richer multimedia experience.

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

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

Part 8:
Portable video application format

1 Scope

This part of ISO/IEC 23000 specifies a file format designed for mid-resolution “DVD-style” video applications. The file format provides the overall structure for storing video contents, images, metadata, and user interface in a single file.

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 14496-12:2005, *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 15938-5:2003, *Information technology — Multimedia content description interface — Part 5: Multimedia description schemes*

3GPP TS 26.245, *Transparent end-to-end Packet switched Streaming Service (PSS); Timed text format*, V6.1.0, December, 2004

3 Terms and definitions

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

3.1
menu scene description
LASeR scene description for composing a menu scene

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

4 Overview of MPEG Standards

4.1 MPEG-4 Advanced Video Coding

ISO/IEC 14496-10 Advanced Video Coding (AVC) is a digital video codec designed to achieve very high data compression. The standard was written by the Joint Video Team (JVT) which is a collaborative partnership between the ITU-T Video Coding Expert Group (VCEG) and the Motion Picture Experts Group (MPEG). ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10) are technically identical. The H.264/AVC project was intended to create a standard that would provide good video quality at substantially lower bit rates than the previous standards (i.e. relative to MPEG-2, H.263, or MPEG-4 Part 2).

4.2 MPEG-4 High Efficiency Advanced Audio Coding (HE-AAC)

ISO/IEC 14496-3 High Efficiency Advanced Audio Coding (HE-AAC) is a lossy digital audio compression scheme developed by MPEG. HE-AAC combines the following technologies:

- Advanced Audio Coding (AAC)
- Spectral Band Replication (SBR)
- Parametric Stereo (PS) – only in version 2

HE-AAC is known for improved compression that provides higher quality and resolution audio (e.g. sampling rates up to 96 kHz) with smaller file size, as well as the support for multichannel audio providing up to 48 full frequency channels. HE-AAC also provides improved decoding efficiency in terms of processing power.

4.3 MPEG-4 Bit Sliced Arithmetic Coding (BSAC)

ISO/IEC 14496-3 Bit Sliced Arithmetic Coding (BSAC) is a standard for scalable audio coding. BSAC uses an alternative noiseless coding to AAC, and all other AAC-based processing remains unchanged.

4.4 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.

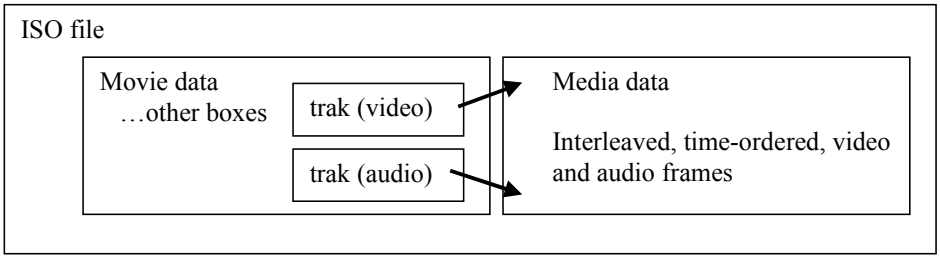


Figure 1 — Example of a simple ISO file used for interchange, containing two streams

The file structure is object-oriented as shown in Figure 1, 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.

--`,,```,,,,````-`-`,,`,,`,`,,`---
© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

4.5 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 (e.g. MP3onMP4) 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) meta-data. Figure 2 schematically illustrates the location of these un-timed MPEG-7 Metadata boxes.

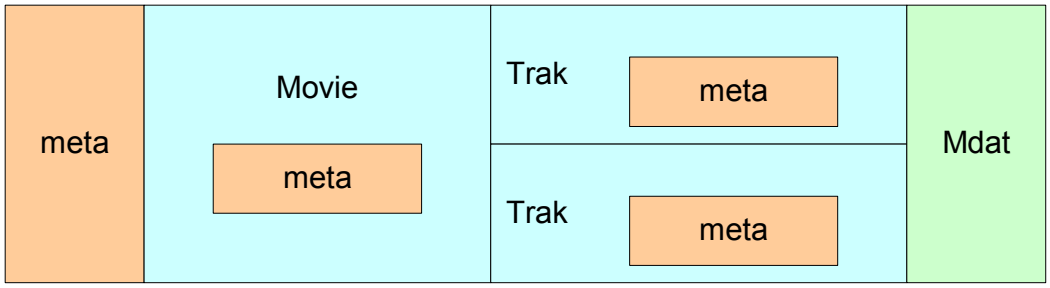


Figure 2 — Support of Static un-timed Metadata in ISO/MP4 Files

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 metadata specification system.

4.7 MPEG-4 LASeR

ISO/IEC 14496-20 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 and portable multimedia players (PMPs).

5 Using the Portable video application format

5.1 General

This clause provides necessary information for creating and playing a Portable video application format file.

5.2 File Contents

A Portable video application format file is comprised of:

- Video and audio data
- Menu scene description

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

- Metadata
- Image data
- Text data

The file format for the Portable video application format is based on the MPEG-4 file format. Hence, it is possible to store multiple video, audio, and text data. The “menu” scene description provides the necessary interface between the user and the media data contained inside the file. A normative Portable video file shall contain video, audio, and a LASeR scene description for composing a menu. Metadata and text are optional. Figure 3 shows how the above data items are converted and stored in a Portable video file.

A Portable video creator is responsible for structuring the file in an MP4 file format, and storing user-created contents (e.g. home videos) as separate video and audio tracks. The user interface (menu) is converted and stored as a LASeR scene description. Metadata is stored in the form of an MPEG-7 metadata. The file format also allows text data, such as movie subtitles, to be contained inside the file. However, both metadata and text data are optional features in the creation process.

A detailed description of the file structure is provided in 5.4.

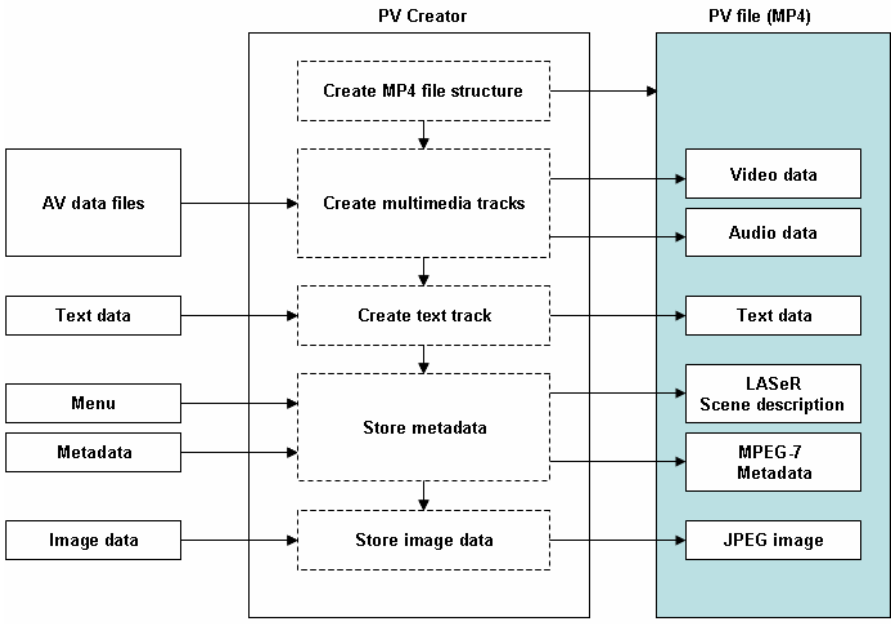


Figure 3 — Example of a Portable video creator system architecture

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

5.3 Playback

5.3.1 Portable video application format playback

For normal playback, the LASeR scene description responsible for the user interface is first extracted from the file, and displayed on screen as a menu. By using the menu options, a user is able to determine the contents of the file, and decide what he or she wants to watch. Figure 4 shows the basic system architecture for a Portable video player.

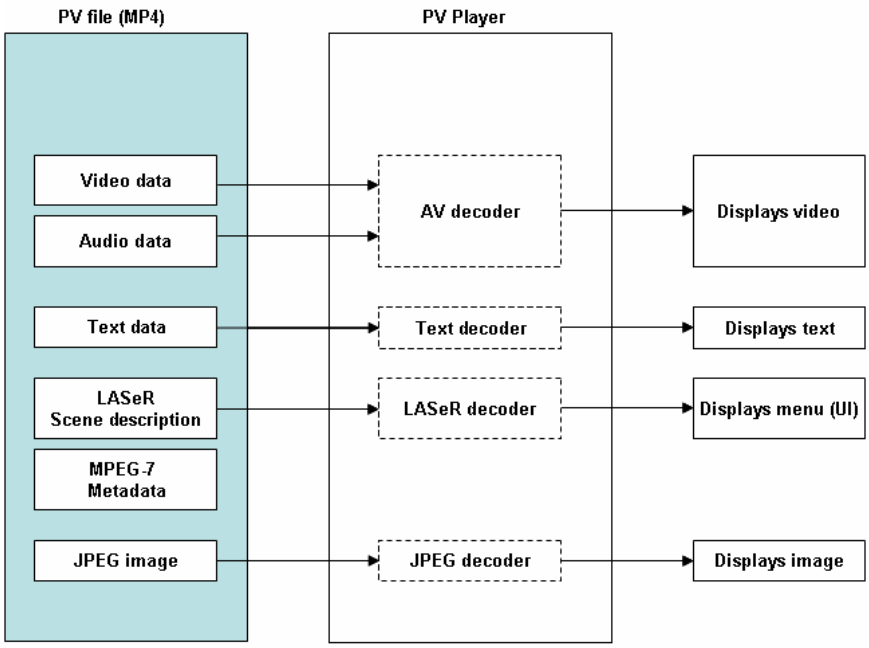


Figure 4 — Example of a Portable video player system architecture

A typical playback operation consists of:

- Extracting the LASeR scene description and displaying it on screen as a menu for user interface
- Extracting the audio and video data for playback on an AV decoder based on user commands

Metadata and text data are optional features, therefore, the player is not responsible for decoding these data. The players that are not capable of supporting such features shall simply ingnore these data types, if they exist in the file.

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

5.3.2 Regular MP4 playback

For regular MP4 players or devices that are not capable of decoding the LASeR scene description, the main video and audio tracks can still be played as a regular MP4 file (as shown in Figure 5).

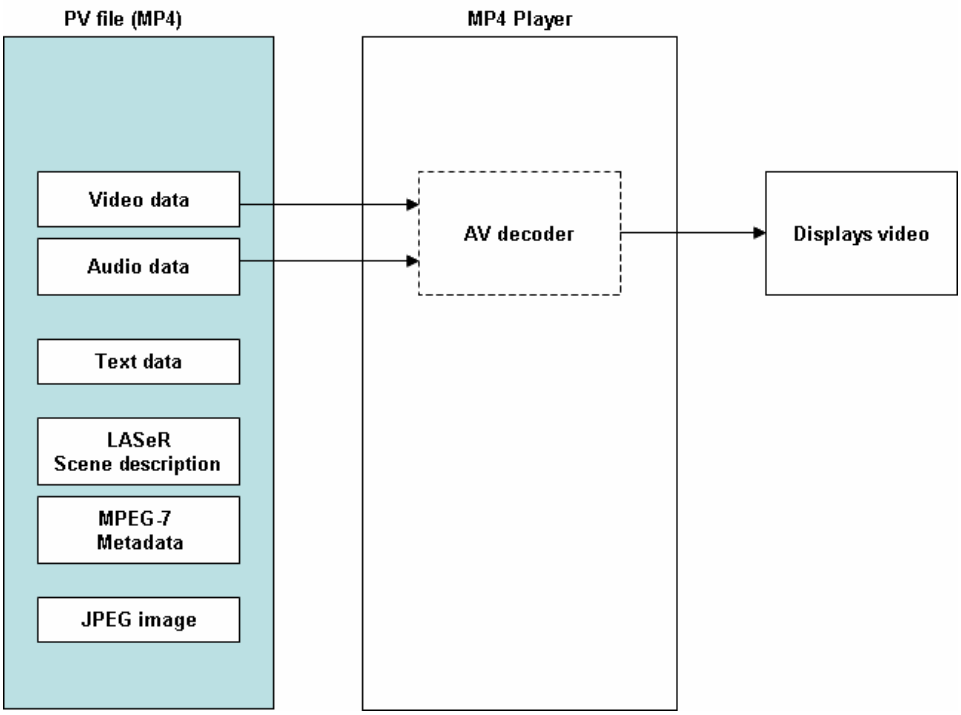


Figure 5 — Example of a Portable video file playback on a regular MP4 player

A general operational flow that outlines the two possible playback scenarios for the Portable video is demonstrated in Figure 6.

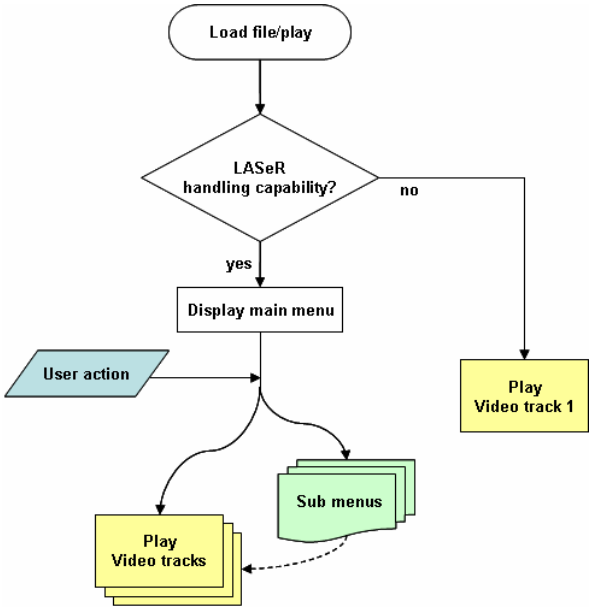


Figure 6 — Portable video application format operational flow diagram

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

5.4 File Structure

5.4.1 Structure Diagram

The file structure for the Portable video application format is derived from the MPEG-4 file format. The normative file structure consists of four file-level boxes (‘ftyp,’ ‘meta,’ ‘moov,’ and ‘mdat’).

An example of a file structure that contains multiple video, audio, and text tracks (e.g. different languages) is shown in Figure 7.

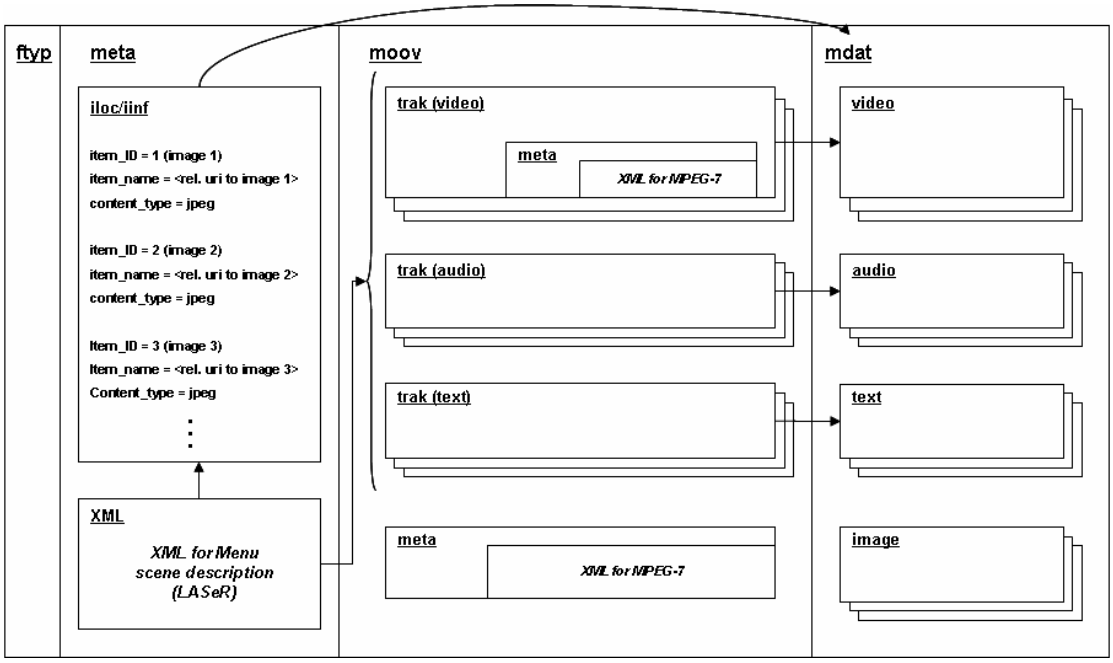


Figure 7 — Example of a Portable video application format file structure containing multiple media tracks

5.4.2 File Type Box (‘ftyp’)

The ‘ftyp’ box is used to identify the type of the file format that this file structure complies to. The brand that identifies the Portable video file structure is ‘pvp1.’

5.4.3 Meta Box (‘meta’)

The file-level ‘meta’ box contains the LASeR scene description for the menu. The description is stored inside the ‘xml’ box (normative). The ‘iloc’ and ‘iinf’ boxes are used for storing media resource information. The metadata handler-type is ‘lsr1'.

For each media item (e.g. video), the physical location (‘extent_offset’) and size (‘extent_length’) are stored in the ‘iloc’ box. The name of the media item (‘item_name’) and the type (‘content_type’) are contained in the ‘iinf’ box. An ‘item_ID’ is assigned to each media item for resource referencing.

5.4.4 Movie Box (‘moov’)

The ‘moov’ box contains ‘trak’ boxes for the media data (video, audio, text), and a ‘meta’ box. The ‘trak’ boxes contain temporal and spatial information of the media data. The movie-level ‘meta’ box is reserved for MPEG-7 metadata (e.g. creation information). Each ‘trak’ box can also utilize the ‘meta’ box for MPEG-7 metadata pertaining to its media data. The ‘moov’ box can contain multiple video, audio, and text tracks. For a file to be considered as a Portable video file, the file has to contain at least one video track, one audio track, and a LASeR script. Metadata and text track are optional.

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

5.4.5 Media Data Box ('mdat')

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

6 Technical features of Portable video application format

6.1 User Interface

6.1.1 Introduction

Movies and television shows released on DVD disks usually also contain various multimedia contents other than the main feature, such as "making of" clips, actor interviews, and image galleries. The Portable video application format is designed to accommodate similar multimedia contents in a single file.

In order to control and handle different multimedia contents in a file, an appropriate user interface is required. By applying a menu screen, a user is able to oversee the contents inside the file, and it provides directions and means for the user to choose what he or she wants to watch. If the user is editing several different home videos into a Portable video file, having a menu helps to represent the overall layout of the content.

In the Portable video application format, MPEG-4 Lightweight Application Scene Representation (LASeR) technology is used for creating a highly engaging menu screen. An example is shown in Figure 8.

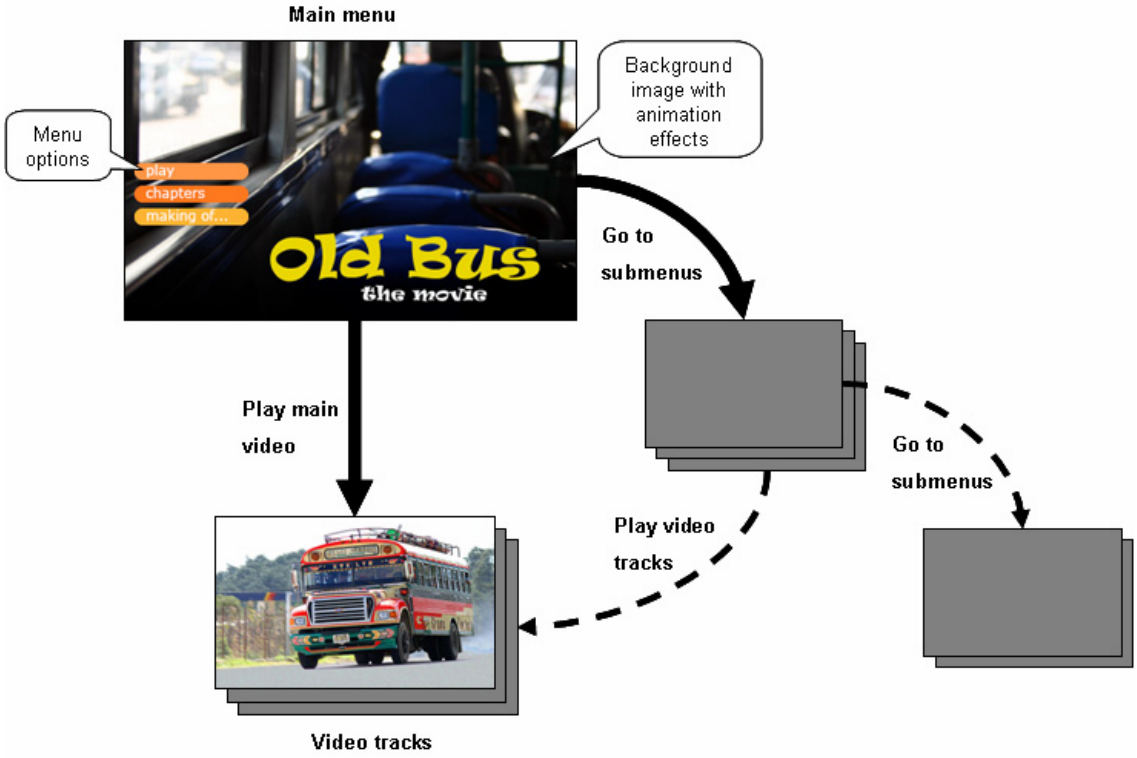


Figure 8 — Example of an animated menu screen for a Portable video file

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
© ISO/IEC 2008 – All rights reserved
Not for Resale

An XML example:

```
<lsru:NewScene>
    <svg width="400" height="250" viewBox="0 0 400 250">
        <!-- Main menu -->
        <g id="title_page" visibility="visible">
            <!-- Background image -->
            <image id="title_image" xlink:href="#ffp(item_ID = 1)" width="400" height="250" x="0"
            y="0"/>

            <!-- Menu options -->
            <rect x="10" y="109" width="100" height="15" rx="8" fill="rgb(255, 150, 70)"/>
            <text id="play_btn" x="20" y="120" font-family="Verdana" font-size="12" fill="white">
            play </text>
            ...
        </g>

        <!-- Video playback -->
        <g id="video_playback" visibility="hidden">
            <video xlink:href="#ffp(track_ID = 1)" begin="play_btn.click" type="video/mpeg" ...>
            <audio xlink:href="#ffp(track_ID = 2)" begin="play_btn.click" type="audio/mpeg" ...>
            <animation xlink:href="#ffp(track_ID = 3)" begin="play_btn.click" type="text/3gpp" ...>
        </g>
        ...
    </svg>
</lsru:NewScene>
```

6.1.2 Execution Model

Figure 9 shows the extraction process for a rendered menu. The LASeR scene description is extracted from the file, and passed onto a LASeR decoder, which is then displayed on screen as a menu.

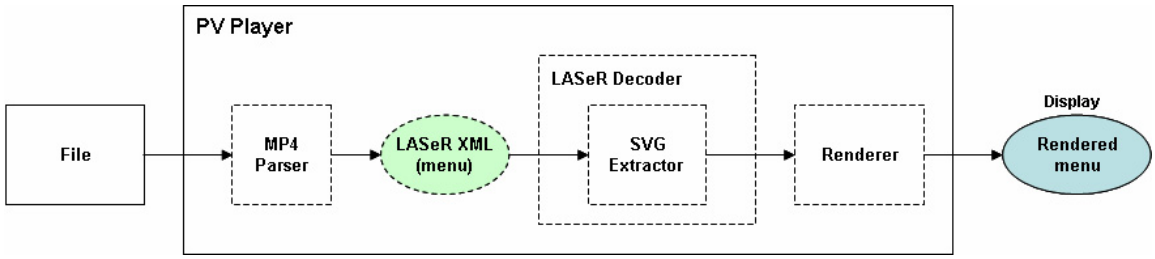


Figure 9 — Execution model for a rendered menu

Figure 10 shows the interaction between a user and the Portable video player.

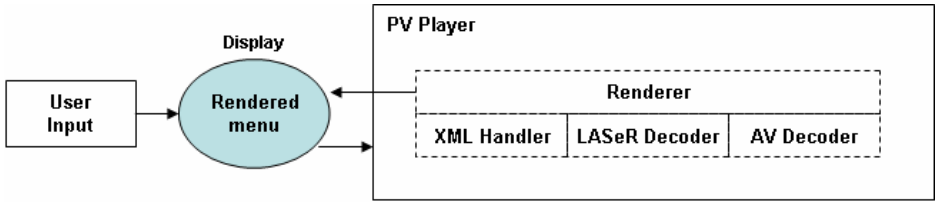


Figure 10 — Interaction model between a user and Portable video player

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

6.1.3 List of "LASeR" elements used for menu description

Table 1 shows the list of normative LASeR elements that are used for the menu description.

Table 1 — List of "LASeR" scene description elements

Effects	Description elements	Functionality
Element grouping	*g*	The *g* element is used for grouping of related graphics elements.
Text rendering	*text*	The *text* element is used for rendering text. For example, in Figure X, menu options and scrolling text are all rendered using the "text" element
Image referencing	*image*	The *image* element makes references to image data/file.
Video referencing	*video*	The *video* element makes references to video data/file.
Audio referencing	*audio*	The *audio* element makes references to audio data/file.
Text referencing	*animation*	The *animation* element makes references to text data/file (e.g. subtitles)
Animation	*animate*	The *animate* element is used for animating a single attribute over a predefined time, e.g. opacity control over time on a background image.
	animateTransform	The *animateTransform* element animates an attribute on a target element over a predefined time, where attributes can be any of the following: translate, scale, rotate, skewX, and skewY. For example, in Figure X, the scrolling text (e.g. song title) is rendered using the *animateTransform* element.
	animateColor	The *animateColor* element specifies a color transformation over a predefined time.
	animateMotion	The *animateMotion* element is used for allowing a referenced element to move along a predefined motion path over time.
Attribute control	*set*	The *set* element is used for setting the value of an attribute for a specified duration
Shapes & motion paths	*path*	The *path* element defines the outline of a shape (e.g. a star pattern). It can also be used for defining a motion path.
Basic shapes	*rect*	The basic shape elements are equivalent to the *path* element, except that they correspond to specific shapes. As an example, in the Portable video application format menu, the menu buttons can be constructed using the basic shape elements.
	circle	
	ellipse	
	line	

6.1.4 Supported Events

Supported events (e.g. mouse-click, key-input) are defined in ISO/IEC 14496-20:2008, 6.5, [Lightweight Application Scene Representation (LASeR)].

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Annex A
(informative)

Use cases of Portable video application format

Portable video players (PVP) today have the ability to play back good quality video on 3-5" size displays at the usual resolutions of 320x240 or 480x272 pixels. Flash memory or hard disk storage can hold many hours of video. Also small-size pre-recorded media like mini-DVD/UMD (1.8 GB capacity) are used in portable devices. As an example, 512MB of total movie file size correspond to 134 min at a bit rate of 500 kbps.

Three different scenarios apply:

- Playback of content the users generated on their own for PVP (downsize and encode), e.g. user created content ("UCC")
- Playback of content sold on disk media (mini-DVD), comparable to the DVD business
- Playback of content sold over the internet, comparable to the online music store business

In order to enable business of selling content for the PVP devices not only interoperability at the codec level is required but also interoperable digital rights management to protect the rights of the content owners.

Premium content will support advanced features like multiple audio tracks in different languages and subtitles (as known by DVDs). Metadata like the film title or the names of the actors should be included. The movie poster can be included as still image.

Simple past TV program "download & play" application

John's favorite television show is "*Friends*." He decides to download last season's episodes from the website (owned by the television network) and store it in his portable media player so that he can watch it during his daily commute to work in the subway. John performs the following:

- He logs on to the website owned by the television network
- He chooses the episodes of "*Friends*" that he wants to download
- He pays for his selections
- He downloads the files to his computer
- He moves the files to his portable media player for later viewing

On his way to work the next morning, John takes out his portable media player and opens the file "friends01_season8.mp4." Once the file loads, he sees a visually enhanced, interactive menu screen. He performs the following:

- He browses the menu options ("Main feature" and "Extra features")
- He chooses the "Extra features" option and presses the "play" button
- He sees another sub-menu screen that has links to "previews" and "behind-the-scene" clips
- He presses the "cancel" button and returns to the main menu screen
- He chooses the "Main feature" option and presses the "play" button
- He begins to watch the episode

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Simple playback application utilizing "Usage History Description"

John has an appointment with his clients at their office, and he has to take the subway, which is the most reliable means of transportation in the city. He has to make two transfers in order to get to his clients' office. John takes out his portable media player once again, and opens another file, "friends02_season8.mp4." He performs the following:

- He chooses the "Main feature" option and presses the "play" button

John arrives at a station where he has to make his first transfer.

- He presses the "stop" button, and he exits the program.

The program saves John's *usage history description* onto a separate file. He transfers onto another subway, and then reloads the file "friends02_season8.mp4" on his PMP. The system automatically loads the *usage history description.*

- He chooses the "Continue" option from the menu screen and presses the "play" button
- He begins to watch the episode from where he left off previously

Simple playback application utilizing "Hierarchical Summary Description"

Susan has missed several episodes of her favorite television newsmagazine program "*60 Minutes*" while she was away on a business trip. She decides to log on to the TV network's website and download the past episodes and play them on her portable media player during her commute to work (by using the subway). Susan performs the following:

- She logs on to the website owned by the television network.
- She chooses the episodes of "*60 minutes*" that she wants to download.
- She pays for her selections.
- She downloads the files to her computer.
- She moves the files to her portable media player for later viewing.

On her way to work the next day, Susan loads the file "60min_050510.mp4" on her PMP. Among many of the options provided in the menu screen, she chooses the "Summary Segments" option, and presses the "play" button. She notices a sub-menu screen that displays three key frame images with key words/key phrases which represent the three stories that are presented in the episode of "*60 minutes"* that she has chosen. The key phrases are "Climate change", "X-treme sports", and "Violence in television".

- She chooses "Climate change" from the menu and presses the "play" button.
- She browses a series of key frames related to the "Climate change" story. The key frames are labeled as "Global warming", "Water pollution", and "Rain forest".
- She chooses one of the key frames labeled "Global warming" and presses the "play" button.
- She begins to watch a clip on global warming instead of watching the whole story on "Climate change."

Simple mobile broadcast "save & replay" application

John is a high school student who frequently watches educational programs pertaining to the 10th grade curriculum on television. He used to record them on VHS tapes and watch them again whenever he needed to. But, now that he has a mobile phone capable of receiving DMB signal, he chose his mobile phone over his television set. He performs the following:

- John selects the DMB function on his mobile phone
- He chooses the EBS (Educational Broadcasting System) channel
- He begins watching a program on 10th grade "Advanced Mathematics"

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

- He presses the button that corresponds to the “SAVE” menu on his phone
- He presses the button that corresponds to the “BOOKMARK” menu, and makes notes using the memo pad function on his phone

After 20 minutes, the program ends, and John performs the following:

- He presses the button that corresponds to the “END” menu on his phone

The media content is saved as a Portable video file along with the metadata that contains bookmarks associated with his notes for later viewing. As long as it is saved as a file, John can always watch them later, hopping back and forth using the bookmarks, on any multimedia device that supports this format, and with the metadata, it is easier to browse when there are lots of files.

Reconsumption of e-learning content

John is scheduled to take his “TOEFL” exam in three months, and he prefers to study on his daily commute to school with his portable multimedia player. He downloads a file called, “toefl070712.mp4” from the EBS web server. He performs the following:

- He presses the “play” button

- He consumes the content with user actions such as “record”, “fast forward”, “repeat”, and, etc.

The system automatically records the *usage history and browsing preference* descriptions.

Several days later, he decides to watch “toefl070712.mp4” again. The material that he wants to watch is positioned 20 minutes into the program. He turns the portable multimedia player on, and selects “toefl070712.mp4” for a replay. The system automatically loads the *usage history* and *browsing preference* descriptions.

- He chooses the “replay” option from the menu screen, and enters “20” min (position in time)

- He presses the “play” button

- He begins to watch a segment that has the highest preference value at a pre-defined time range (e.g. 00:12:00 ~ 00:32:00)

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Annex B
(informative)

Usage of MPEG-7 MDS metadata in Portable video application format

B.1 Introduction

In the Portable video application format, metadata is used to provide additional information about the content. The developers may utilize this information to implement extra functionality in their device applications.

The Portable video application format supports the following MPEG-7 metadata:

- Creation DS (Information on content creation)
- HierarchicalSummary DS (Information on chaptering and video summary)
- UsageHistory DS (Information on user actions)
- AudioVisualSegment DS (Bookmarks)
- BrowsingPreferences DS
- ContentCollection DS

Figure B.1 shows an example of an execution model for extracting metadata from a Portable video file.

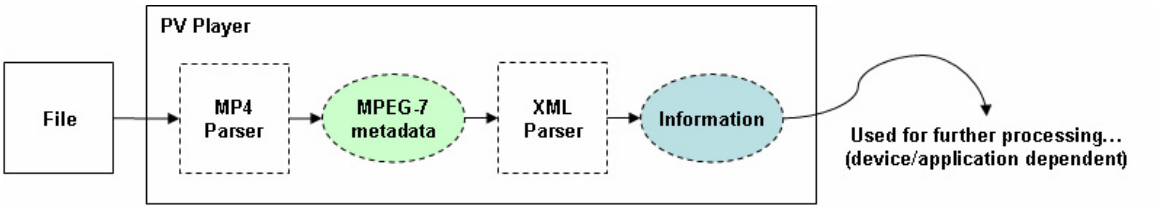


Figure B.1 — Execution model for extracting metadata

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

B.2 Metadata structure

The metadata shall be structured and contained within two different levels of the file. The metadata pertaining to the file, or a collection of contents as a whole are contained inside the movie-level ‘meta’ box. Figure B.2 shows the metadata structure.

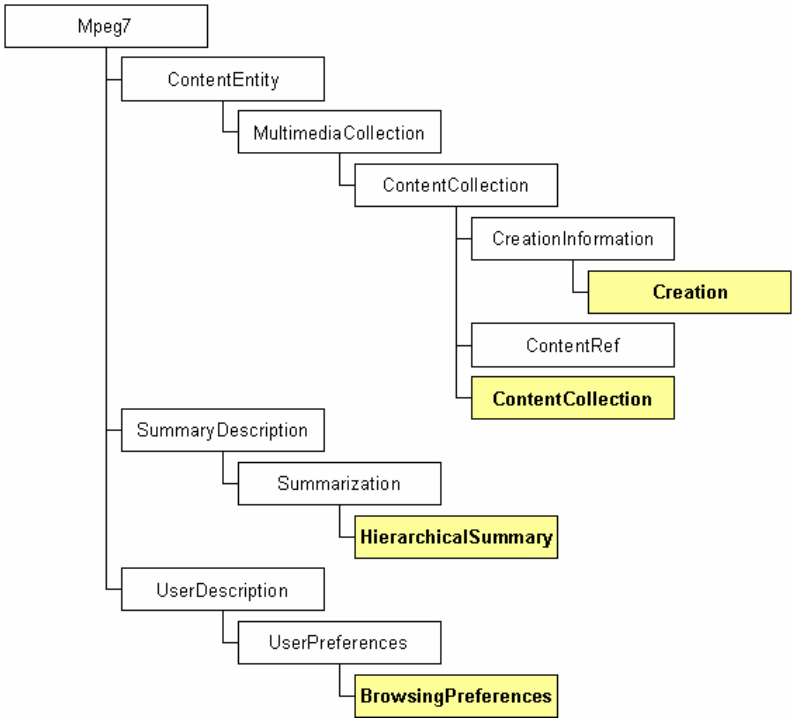


Figure B.2 — Movie-level metadata structure

The metadata that describes an individual content (e.g. a video clip) is stored inside the corresponding track-level ‘meta’ box. The structure is shown in Figure B.3.

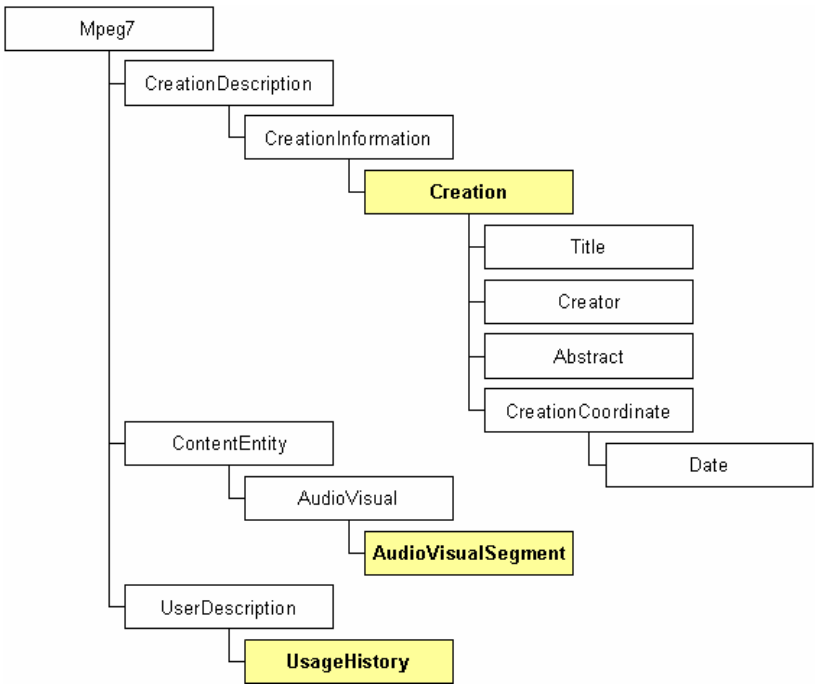


Figure B.3 — Track-level metadata structure

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

B.3 Creation DS

The *Creation DS* is an MPEG-7 MDS tool that organizes content creation information, including the following metadata:

- Title (title of the movie)
- Artist (the creator who took the clip/movie)
- Clip description (general description of the clip/movie)
- File date/time (the date/time in which the clip was taken)

B.4 HierarchicalSummary DS

B.4.1 Introduction

The *HierarchicalSummary DS* is an MPEG-7 MDS tool that is used to summarize a video content at different levels of detail. In the Portable video application format, the *HierarchicalSummary DS* is specifically used for chaptering and video summary.

B.4.2 Chaptering

For chaptering, each temporal segment of the video data that corresponds to a scene is described by a summary segment, and the summary segments are organized into a summary segment group, which corresponds to a group of scenes (e.g. scene 1 – 10). Figure B.4 shows an example of chaptering.

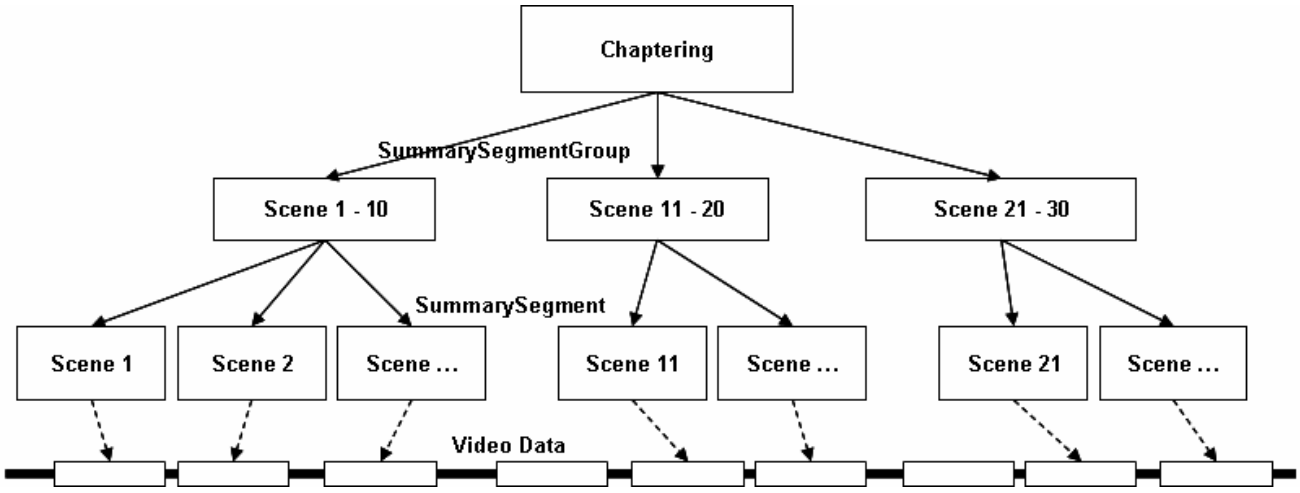


Figure B.4 — Example of chaptering

The main benefit of using such description scheme is that a user can obtain necessary information that he or she only requires, without having to search through the whole content by fast-forwarding and rewinding constantly during playback.

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

B.4.3 Video summary

For video summary, *HierarchicalSummary DS* is used for organizing the video content thematically. Each temporal segment of the video data is described by a summary segment, and each summary segment locates the associated key clips or key frames. The summary segments are grouped into summary segment groups, and summary segment groups are grouped into a hierarchical summary description. Figure B.5 shows an example of a video summary.

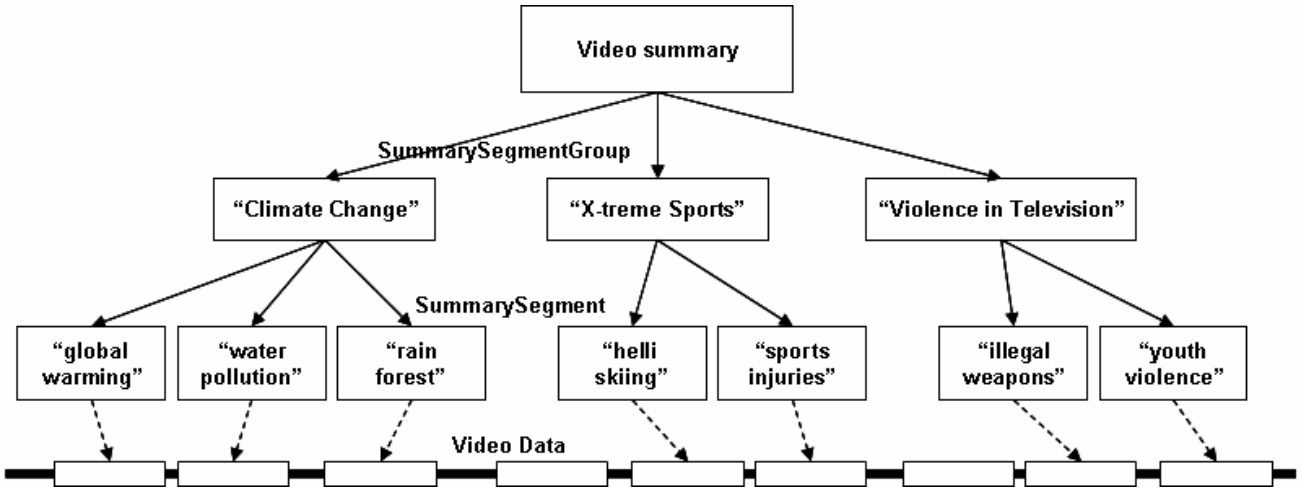


Figure B.5 — Example of a video summary

A user can search for specific topics in a library of contents, e.g. if Susan had collected 125 episodes of "*60 minutes*", and if she wanted to search for all the episodes relating to "global warming", she could easily proceed with the search (e.g. keyword search), if all the episodes were summarized with the hierarchical summary description scheme.

B.5 UsageHistory DS

The *UsageHistory DS* is a tool that describes a user's actions carried out over a specified period of time on consuming a multimedia content, and such descriptions can be exchanged between the user and the service providers. With the Portable video, the UsageHistory DS is used as means to store user history description, and a tool to compensate for interruptions that are likely to occur by being used on a portable application. Figure B.6 shows an example of a Portable video application that utilizes user history description.

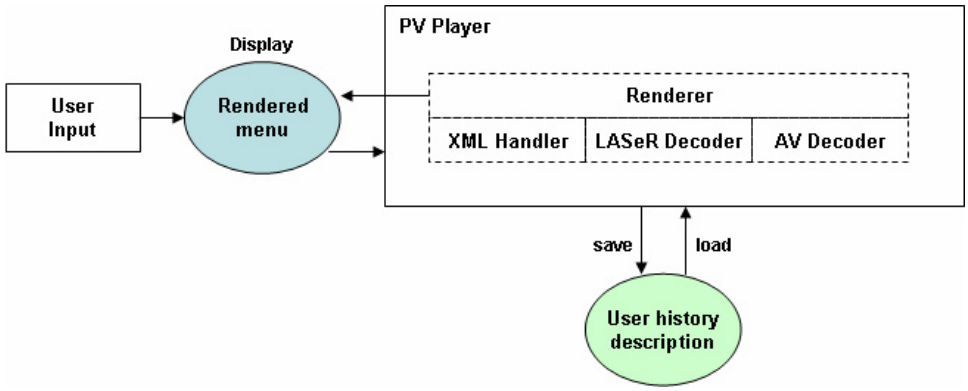


Figure B.6 — Interaction model for user history description

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

If a user decides to watch a 2-hour movie in the comforts of his own living room, then he is more likely to finish watching the whole movie than, say, a user who is in a subway watching the same movie on his portable multimedia player (PMP) during his commute to work. If a user history description exists, he can always watch the movie from the point where he previously left off.

B.6 AudioVisualSegment DS

Another usage of *UsageHistory DS*, described in 4.2.4, is to carry bookmark information a user adds while watching a video program. Timing information of bookmarks and reference to their associated text are stored in this DS. The associated text can be contained as text annotation in *AudioVisualSegment DS*, as depicted in Figure B.7. This allows users to attach notes to some specific time points of the program, which is useful especially for educational video programs.

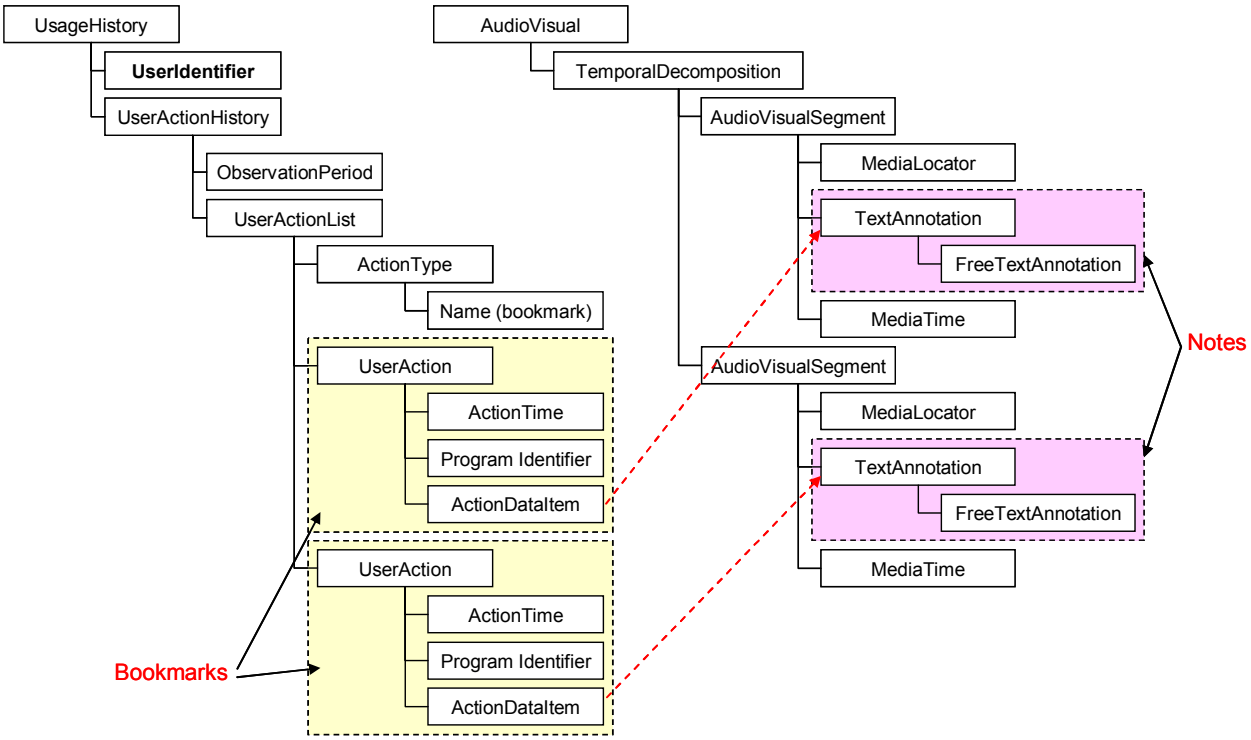


Figure B.7 — Bookmarks and their associated notes

B.7 BrowsingPreferences DS

The *BrowsingPreferences DS* describes user preferences on specific themes or segments pertaining to media content in a file. The preferences are expressed using preference values, and they may be conditioned on various attributes such as themes, genre, and types of content. The basic functionality of this tool is to provide users with the ability to search for their favorite segments based on individual preferences. Figure B.8 shows the relationship between summary segments and corresponding preference values.

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

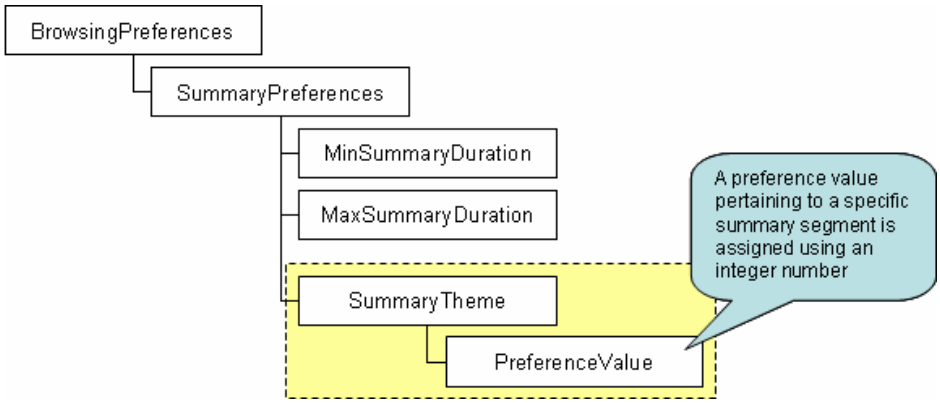


Figure B.8 — Summary segments and corresponding preference descriptions

B.8 ContentCollection DS

The *ContentCollection DS* describes the collection of media resources contained inside a Portable video file. The description provides information on a collection, or collections of individual media data. This tool provides users with extra means to search for media data in Portable video files. Figure B.9 shows the *ContentCollection DS* layout in Portable video application format.

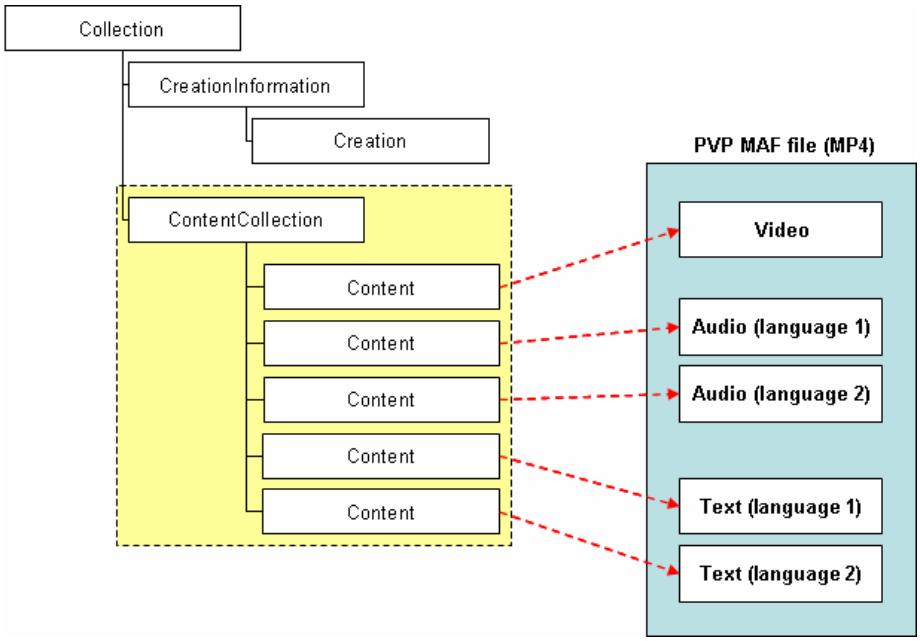


Figure B.9 — Media resources and content collection description

B.9 Description examples

Examples of using MPEG-7 multimedia description schemes in the Portable video application format are shown below.

Table B.1 shows the example of a “UsageHistory” description. It keeps records of user’s actions; therefore it is used as a tool to compensate for interruptions that are likely to occur by being used on a portable application.

--`,,```,,,,````-`-`,,`,,`,`,,`---
© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Table B.1 — Example of a *UsageHistory* description

```
<UsageHistory id="usage-history-001" allowCollection="true">
  <UserIdentifier protected="true">
      <Name xml:lang="en">John Smith</Name>
  </UserIdentifier>
  <UserActionHistory id="useraction-history-001 protected="false">
      <ObservationPeriod>
          <TimePoint>2006-05-17T07:10</TimePoint>
          <Duration>T01:00:00</Duration>
      </ObservationPeriod>
      <ObservationPeriod>
          <TimePoint>2006-05-17T18:30</TimePoint>
          <Duration>T01:00:00</Duration>
      </ObservationPeriod>
      <UserActionList id="useraction-list-01" numOfInstances="2" totalDuration="T00:50:00">
          <ActionType href="urn:mpeg:mpeg7:cs:ActionTypeCS:2001:1.3">
              <Name>Play</Name>
          </ActionType>
          <UserAction>
              <ActionTime>
                  <MediaTime>
                      <MediaTimePoint>T00:00:00</MediaTimePoint>
                      <MediaTimeDuration>T00:25:00</MediaTimeDuration>
                  </MediaTime>
                  <GeneralTime>
                      <TimePoint>2006-05-17T07:10</TimePoint>
                      <Duration>T00:25:00</Duration>
                  </GeneralTime>
              </ActionTime>
              <ProgramIdentifier organization="org" type="typ">soeun_cha_ep01</ProgramIdentifier>
              <ActionDataItem href="#episode01"/>
          </UserAction>
          <UserAction>
              <ActionTime>
                  <MediaTime>
                      <MediaTimePoint>T00:25:00</MediaTimePoint>
                      <MediaTimeDuration>T00:25:00</MediaTimeDuration>
                  </MediaTime>
                  <GeneralTime>
                      <TimePoint>2006-05-17T18:05</TimePoint>
                      <Duration>T00:25:00</Duration>
                  </GeneralTime>
              </ActionTime>
              <ProgramIdentifier organization="org" type="typ">soeun_cha_ep01</ProgramIdentifier>
              <ActionDataItem href="#episode01"/>
          </UserAction>
      </UserActionList>
      <UserActionList id="useraction-list-01" numOfInstances="2" totalDuration="T00:00:02">
          <ActionType href="urn:mpeg:mpeg7:cs:ActionTypeCS:2001:1.3">
              <Name>Stop</Name>
          </ActionType>
          <UserAction>
              <ActionTime>
                  <MediaTime>
                      <MediaTimePoint>T00:25:00</MediaTimePoint>
                      <MediaTimeDuration>T00:00:01</MediaTimeDuration>
                  </MediaTime>
                  <GeneralTime>
                      <TimePoint>2006-05-17T07:37</TimePoint>
                      <Duration>T00:00:01</Duration>
                  </GeneralTime>
              </ActionTime>
              <ProgramIdentifier organization="org" type="typ">soeun_cha_ep01</ProgramIdentifier>
              <ActionDataItem href="#episode01"/>
          </UserAction>
          <UserAction>
              <ActionTime>
                  <MediaTime>
                      <MediaTimePoint>T00:50:00</MediaTimePoint>
                      <MediaTimeDuration>T00:00:01</MediaTimeDuration>
                  </MediaTime>
                  <GeneralTime>
                      <TimePoint>2006-05-17T18:30</TimePoint>
                      <Duration>T00:00:01</Duration>
```


© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    </GeneralTime>
                </ActionTime>
                <ProgramIdentifier organization="org" type="typ">soeun_cha_ep01</ProgramIdentifier>
                <ActionDataItem href="#episode01"/>
            </UserAction>
        </UserActionList>
        <UserActionList id="useraction-list-01" numOfInstances="1" totalDuration="T00:02:00">
            <ActionType href="urn:mpeg:mpeg7:cs:ActionTypeCS:2001:1.3">
                <Name>Pause</Name>
            </ActionType>
            <UserAction>
                <ActionTime>
                    <MediaTime>
                        <MediaTimePoint>T00:14:23</MediaTimePoint>
                        <MediaTimeDuration>T00:02:00</MediaTimeDuration>
                    </MediaTime>
                    <GeneralTime>
                        <TimePoint>2006-05-17T07:24</TimePoint>
                        <Duration>T00:02:00</Duration>
                    </GeneralTime>
                </ActionTime>
                <ProgramIdentifier organization="org" type="typ">soeun_cha_ep01</ProgramIdentifier>
                <ActionDataItem href="#episode01"/>
            </UserAction>
        </UserActionList>
    </UserActionHistory>
 </UsageHistory>
```

Table B.2 shows the example of carrying bookmarks. It keeps records of user's actions; therefore it is used as a tool to compensate for interruptions that are likely to occur by being used on a portable application.

Table B.2 — Example of a *UsageHistory* description that contains bookmark information

```
<Mpeg7>
  <Description xsi:type="ContentEntityType">
    <MultimediaContent xsi:type="AudioVisualType">
      <AudioVisual>
        <MediaTime>
          <MediaTimePoint>PT0S</MediaTimePoint>
          <MediaDuration>PT30M</MediaDuration>
        </MediaTime>

        <!-- Decomposition related to bookmarks -->
        <TemporalDecomposition criteria="manual" gap="true">

          <!-- note1 -->
          <AudioVisualSegment id="bookmark-note1">
            <TextAnnotation>
              <FreeTextAnnotation>
                Explanation of Pythagorean theorem
              </FreeTextAnnotation>
            </TextAnnotation>
            <MediaTime>
              <MediaRelTimePoint>PT1M30S</MediaRelTimePoint>
              <MediaDuration>PT0S</MediaDuration>
            </MediaTime>
          </AudioVisualSegment>

          <!-- note2 -->
          <AudioVisualSegment id="bookmark-note2">
            <TextAnnotation>
              <FreeTextAnnotation>
                Exercise 1
              </FreeTextAnnotation>
            </TextAnnotation>
            <MediaTime>
              <MediaRelTimePoint>PT10M15S</MediaRelTimePoint>
              <MediaDuration>PT0S</MediaDuration>
            </MediaTime>
          </AudioVisualSegment>

        </TemporalDecomposition>
```

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
      </AudioVisual>
    </MultimediaContent>
  </Description>

  <!-- Description of usage history  -->
  <Description xsi:type="UserDescription">
    <UsageHistory id="usage-history-004" allowCollection="true">
      <UserIdentifier protected="true">
        <Name xml:lang="en">John Doe</Name>
      </UserIdentifier>
      <UserActionHistory id="useraction-history-bookmark-001" protected="false">
        <ObservationPeriod>
          <TimePoint> 2007-04-16T18:00-8:00</TimePoint>
          <Duration>PT6H</Duration>
        </ObservationPeriod>

        <!-- bookmark list -->
        <UserActionList id="ua-bookmark-list-001" numOfInstances="2" totalDuration="">
          <ActionType>
            <Name>Bookmarks</Name>
          </ActionType>

          <!-- bookmark 1  -->
          <UserAction>
            <ActionTime>
              <MediaTime>
                <MediaRelTimePoint>PT1M30S</MediaTimePoint>
                <MediaDuration>PT0S</MediaDuration>
              </MediaTime>
              <GeneralTime>
                <TimePoint>2007-04-16T18:10-8:00</TimePoint>
              </GeneralTime>
            </ActionTime>
            <ProgramIdentifier>01-mnf-100900</ProgramIdentifier>
            <ActionDataItem idref="bookmark-note1"/>
          </UserAction>

          <!-- bookmark 2  -->
          <UserAction>
            <ActionTime>
              <MediaTime>
                <MediaRelTimePoint>PT10M15S</MediaRelTimePoint>
                <MediaDuration>PT0S</MediaDuration>
              </MediaTime>
              <GeneralTime>
                <TimePoint>2007-04-16T18:20-8:00</TimePoint>
              </GeneralTime>
            </ActionTime>
            <ProgramIdentifier>01-mnf-100900</ProgramIdentifier>
            <ActionDataItem idref="bookmark-note2"/>
          </UserAction>

          <!-- more bookmarks . . . -->

        </UserActionList>
      </UserActionHistory>
    </UsageHistory>
  </Description>
</Mpeg7>
```

Table B.3 shows the example of a “HierarchicalSummary” description where the user can acquire necessary information that he only needs, without having to search through the whole content by fast-forwarding and rewinding constantly during playback.

--`,,```,,,,````-`-`,,`,,`,`,,`---
© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Table B.3 — Example of a HierarchicalSummary description

```
<HierarchicalSummary components="keyFrames" hierarchy="independent">
    <SourceInformation href="#episode01"/>
    <SummaryThemeList>
        <SummaryTheme id="theme00">Introduction</SummaryTheme>
        <SummaryTheme id="theme01">Global warming</SummaryTheme>
        <SummaryTheme id="theme02">Water pollution</SummaryTheme>
        <SummaryTheme id="theme03">Rain forest</SummaryTheme>
        <SummaryTheme id="theme04">Helli skiing</SummaryTheme>
        <SummaryTheme id="theme05">Sports injuries</SummaryTheme>
        <SummaryTheme id="theme06">Illegal weapons</SummaryTheme>
        <SummaryTheme id="theme07">Youth violence</SummaryTheme>
    </SummaryThemeList>
    <SummarySegmentGroup level="0" numOfKeyFrames="1">
        <SummarySegment id="Clip00" themeIDs="theme00">
            <Name>Introduction</Name>
            <KeyAudioVisualClip>
                <MediaTimePoint>T00:00:00</MediaTimePoint>
                <MediaDuration>T00:01:45</MediaDuration>
            </KeyAudioVisualClip>
            <KeyFrame>
                <MediaTimePoint>T00:00:00</MediaTimePoint>
            </KeyFrame>
        </SummarySegment>
        <SummarySegmentGroup level="1" numOfKeyFrames="1">
            <SummarySegment id="Keyframe01">
                <Name>Climate Change</Name>
                <KeyFrame>
                    <MediaTimePoint>T00:01:43</MediaTimePoint>
                </KeyFrame>
            </SummarySegment>
            <SummarySegmentGroup level="2" numOfKeyFrames="3">
                <SummarySegment id="Clip01" themeIDs="theme01">
                    <Name>Global warming</Name>
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:01:45</MediaTimePoint>
                        <MediaDuration>T00:06:12</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:01:45</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
                <SummarySegment id="Clip02" themeIDs="theme02">
                    <Name>Water pollution</Name>
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:08:03</MediaTimePoint>
                        <MediaDuration>T00:06:25</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:08:03</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
                <SummarySegment id="Clip03" themeIDs="theme03">
                    <Name>Rain forest</Name>
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:14:57</MediaTimePoint>
                        <MediaDuration>T00:06:00</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:14:57</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
            </SummarySegmentGroup>
        </SummarySegmentGroup>
        <SummarySegmentGroup level="1" numOfKeyFrames="1">
            <SummarySegment id="Keyframe02">
                <Name>Extreme sports</Name>
                <KeyFrame>
                    <MediaTimePoint>T00:20:58</MediaTimePoint>
                </KeyFrame>
            </SummarySegment>
            <SummarySegmentGroup level="2" numOfKeyFrames="2">
                <SummarySegment id="Clip04" themeIDs="theme04">
                    <Name>Helli skiing</Name>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:21:00</MediaTimePoint>
                        <MediaDuration>T00:10:05</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:01:45</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
                <SummarySegment id="Clip05" themeIDs="theme05">
                    <Name>Sports injuries</Name>
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:32:00</MediaTimePoint>
                        <MediaDuration>T00:09:10</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:32:00</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
            </SummarySegmentGroup>
        </SummarySegmentGroup>
        <SummarySegmentGroup level="1" numOfKeyFrames="1">
            <SummarySegment id="Keyframe03">
                <Name>Violence on television</Name>
                <KeyFrame>
                    <MediaTimePoint>T00:41:55</MediaTimePoint>
                </KeyFrame>
            </SummarySegment>
            <SummarySegmentGroup level="2" numOfKeyFrames="2">
                <SummarySegment id="Clip06" themeIDs="theme06">
                    <Name>Illegal weapons</Name>
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:42:00</MediaTimePoint>
                        <MediaDuration>T00:08:50</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:42:00</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
                <SummarySegment id="Clip07" themeIDs="theme07">
                    <Name>Youth violence</Name>
                    <KeyAudioVisualClip>
                        <MediaTimePoint>T00:50:52</MediaTimePoint>
                        <MediaDuration>T00:09:08</MediaDuration>
                    </KeyAudioVisualClip>
                    <KeyFrame>
                        <MediaTimePoint>T00:50:52</MediaTimePoint>
                    </KeyFrame>
                </SummarySegment>
            </SummarySegmentGroup>
        </SummarySegmentGroup>
    </SummarySegmentGroup>
</HierarchicalSummary>
```

Table B.4 shows the example of a "BroswsingPreference" description. It describes the user preference on different summary segment themes using preference values. The tool allows users to easily access video segments in a file, based on their preferences.

Table B.4 — Example of a *BrowsingPreferences* description

```
<BrowsingPreferences protected="true">
     <SummaryPreferences>
          <SummaryType>audioVisual</SummaryType>
          <MinSummaryDuration>PT05M</MinSummaryDuration>
          <MaxSummaryDuration>PT10M</MaxSummaryDuration>
          <SummaryTheme preferenceValue="12">Introduction       </SummaryTheme>
          <SummaryTheme preferenceValue="16">Global warming     </SummaryTheme>
          <SummaryTheme preferenceValue="27">Water pollution    </SummaryTheme>
          <SummaryTheme preferenceValue="25">Rain forest        </SummaryTheme>
          <SummaryTheme preferenceValue="16">Helli skiing       </SummaryTheme>
          <SummaryTheme preferenceValue="14">Sports injuries    </SummaryTheme>
     </SummaryPreferences>
</BrowsingPreferences>
```


© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Table B.5 shows the example of a “ContentCollection” description. It describes a collection of individual media data. This tool provides users with extra means to search for different resources in Portable video files.

Table B.5 — Example of a *ContentCollection* description

```
<Collection xsi:type="ContentCollectionType" name="Collection - Golf lesson">
     <CreationInformation>
          <Creation>
               <Title type="popular"> Basic Golf </Title>
               <Creator>
                    <Role href="RoleCS">
                         <Name> Anchor </Name>
                    </Role>
                    <Agent xsi:type="PersonType">
                         <Name>
                              <GivenName> John </GivenName>
                              <FamilyName> Smith </FamilyName>
                         </Name>
                    </Agent>
               </Creator>
          </Creation>
     </CreationInformation>
     <ContentCollection name="History of Golf">
          <Content xsi:type="VideoType">
               <Video>
                    <MediaLocator>
                         <MediaUri>#ffp(track_ID = 1)</MediaUri>
                    </MediaLocator>
                    <CreationInformation>
                         <Creation>
                              <Title> History of Golf with Green Jacket </Title>
                         </Creation>
                    </CreationInformation>
               </Video>
          </Content>
          <Content xsi:type="AudioType">
               <Audio>
                    <MediaLocator>
                         <MediaUri>#ffp(track_ID = 2)</MediaUri>
                    </MediaLocator>
                    <CreationInformation>
                         <Creation>
                              <Title> History of Golf with Green Jacket </Title>
                         </Creation>
                    </CreationInformation>
               </Audio>
          </Content>
          <Content xsi:type="ImageType">
               <Image>
                    <MediaLocator>
                         <MediaUri>#ffp(track_ID = 3)</MediaUri>
                    </MediaLocator>
                    <CreationInformation>
                         <Creation>
                              <Title> Tiger Woods with Green Jacket</Title>
                         </Creation>
                    </CreationInformation>
               </Image>
          </Content>
     </ContentCollection>
</Collection>
```

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Annex C
(informative)

Subtitle (captioning)

C.1 Introduction

Subtitles are textual information, usually displayed at the bottom of the screen to help viewers understand the dialogue of a movie or television shows. They are rendered as written translation of a foreign dialogue, but they are also used for dialogues in the native language (e.g. closed captioning) to help viewers who have hearing impairment, or difficulty in following the dialogue.

Many video player applications require the use of external files to contain subtitles. In the Portable video application format, subtitles are provided within the file structure as text tracks.

C.2 Timed text

3GPP TS 26.245 timed text format is used for storing subtitles inside a Portable video file as text tracks.

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.

Figure 20 shows the structural interpretation of text sample.

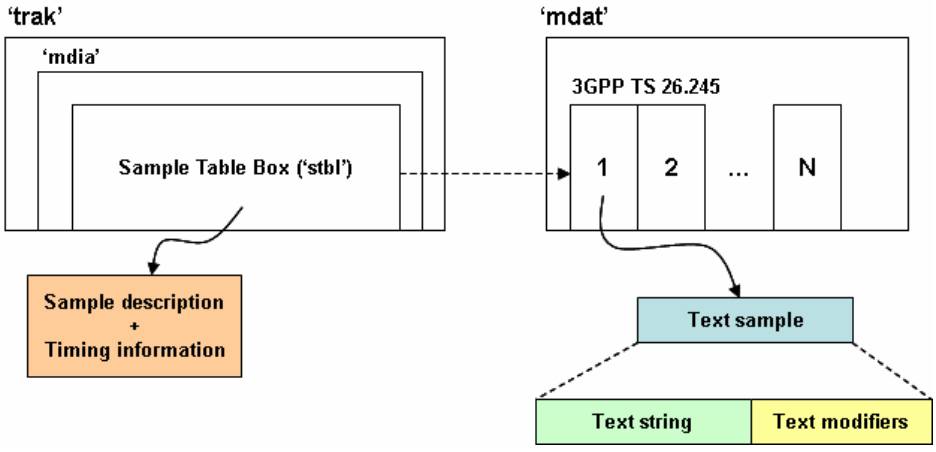


Figure C.1 — Structural interpretation of text sample

The supported functionalities of timed text are:

- Characters and glyphs support
- Font Support
- Color Support
- Text rendering position and composition

© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Annex D
(informative)

XML example of a menu scene description

An XML example for a menu scene description containing submenus is provided in table D.1.

Table D.1 — Example of a menu scene description containing submenus

```
<lsru:NewScene>
    <svg xmlns="http://www.w3.org/2000/svg" version="1.2"
        baseProfile="tiny" xmlns:xlink="http://www.w3.org/1999/xlink"
        width="400" height="250" viewBox="0 0 400 250" >

        <rect width="400" height="250" fill="black"/>

        <!-- Main menu screen -->
        <g id="title_page" visibility="visible">
            <!-- Background image -->
            <image id="title_image" xlink:href="#ffp(item_ID = 1)" width="400"
                height="250" x="0" y="0"/>

            <!-- Scrolling title -->
            <text x="-250" y="210" font-family="Verdana" font-size="12" fill="green" >
                "The Other Side - starring Soeun C"
                <animateTransform attributeName="transform" type="translate" additive="sum"
                    calcMode="linear" from="0 0" to="450 0" dur="7" repeatDur="indefinite"/>
            </text>

            <!-- Operation menu -->
            <text id="play_btn" x="20" y="120" font-family="Verdana"
                font-size="11" fill="white" > play </text>
            <text id="gallery_btn" x="20" y="160" font-family="Verdana"
                font-size="11" fill="white" > gallery </text>

            <!-- ON/OFF switch -->
            <set attributeName="visibility" attributeType="CSS" to="visible"
                begin="video_bk2m_btn.click; gallery_bk2m_btn.click"/>
            <set attributeName="visibility" attributeType="CSS" to="hidden"
                begin="play_btn.click; gallery_btn.click"/>
        </g>

        <!—- Gallery menu screen -->
        <g id="gallery_page" visibility="hidden">
            <!-- Background image -->
            <image id="title_image" xlink:href="#ffp(item_ID = 2)" width="400"
                height="250" x="0" y="0">
                <animate attributeName="opacity" calcMode="linear" values="0.20" />
            </image>

            <!-- Page title text -->
            <text x="160" y="30" font-family="Verdana" font-size="20" fill="white" > gallery </text>

            <!-- image set 1 -->
            <image id="image1" xlink:href="#ffp(item_ID = 3)" width="50"
                height="50" x="20" y="50"/>
            <image id="image2" xlink:href="#ffp(item_ID = 4)" width="50"
                height="50" x="80" y="50"/>
            <image id="image3" xlink:href="#ffp(item_ID = 5)" width="50"
                height="50" x="140" y="50"/>
            <image id="image4" xlink:href="#ffp(item_ID = 6)" width="50"
                height="50" x="200" y="50"/>
            <image id="image5" xlink:href="#ffp(item_ID = 7)" width="50"
                height="50" x="260" y="50"/>
            <image id="image6" xlink:href="#ffp(item_ID = 8)" width="50"
                height="50" x="320" y="50"/>

            <!-- image set 2 -->
```

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
            <image id="image7" xlink:href="#ffp(item_ID = 9)" width="50"
                height="50" x="20" y="110"/>
            <image id="image8" xlink:href="#ffp(item_ID = 10)" width="50"
                height="50" x="80" y="110"/>
            <image id="image9" xlink:href="#ffp(item_ID = 11)" width="50"
                height="50" x="140" y="110"/>
            <image id="image10" xlink:href="#ffp(item_ID = 12)" width="50"
                height="50" x="200" y="110"/>
            <image id="image11" xlink:href="#ffp(item_ID = 13)" width="50"
                height="50" x="260" y="110"/>
            <image id="image12" xlink:href="#ffp(item_ID = 14)" width="50"
                height="50" x="320" y="110"/>

            <!-- Go back -->
            <text id="gallery_bk2m_btn" x="360" y="240" font-family="Verdana"
                font-size="11" fill="white" > back </text>

            <!-- "ON/OFF" switch -->
            <set attributeName="visibility" attributeType="CSS" to="visible"
                begin="gallery_btn.click; imagedisplay_bk2g_btn.click" />
            <set attributeName="visibility" attributeType="CSS" to="hidden"
                begin="play_btn.click; gallery_bk2m_btn.click; image1.click"/>
        </g>

        <!-- Video playback -->
        <g id="video_playback" visibility="hidden">
            <video xlink:href="#ffp(track_ID = 1)" begin="play_btn.click" audio-level=".8"
                type="video/mpeg" width="320" height="136" x="43" y="50" repeatCount="0"/>
            <audio xlink:href="#ffp(track_ID = 1)" begin="play_btn.click" audio-level=".8"
                type="audio/mpeg" width="320" height="136" x="43" y="50" repeatCount="0"/>


            <!-- Go back -->
            <text id="video_bk2m_btn" x="360" y="240" font-family="Verdana"
                font-size="11" fill="white" > back </text>

            <!-- "ON/OFF" switch -->
            <set attributeName="visibility" attributeType="CSS" to="visible" begin="play_btn.click"/>
            <set attributeName="visibility" attributeType="CSS" to="hidden"
                begin="video_bk2m_btn.click; gallery_btn.click"/>
        </g>

        <!-- Image01 display -->
        <g id="image_display" visibility="hidden">
            <image xlink:href="#ffp(item_ID = 3)" width="210" height="250" x="90" y="0" />

            <!-- Go back -->
            <text id="imagedisplay_bk2g_btn" x="360" y="240" font-family="Verdana"
                font-size="11" fill="white" > back </text>

            <!-- "ON/OFF" switch -->
            <set attributeName="visibility" attributeType="CSS" to="visible" begin="image1.click"/>
            <set attributeName="visibility" attributeType="CSS" to="hidden"
                begin="imagedisplay_bk2g_btn.click"/>
        </g>
    </svg>
</lsru:NewScene>
```


© ISO/IEC 2008 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Bibliography

[1] ISO/IEC 14496-3:2005, *Information technology — Coding of audio-visual objects — Part 3: Audio*

[2] ISO/IEC 14496-10:2003, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

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

[4] ISO/IEC 15938-3, *Information technology — Multimedia content description interface — Part 3: Visual*

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

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

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

ICS 35.040

Price based on 29 pages

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2008 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale