IEC/TS 62592

Edition 2.0 2012-07

TECHNICAL SPECIFICATION

This document is a preview generated by EVS

Encoding guidelines for portable multimedia CE products using MP4 file format with AVC video codec and AAC audio codec

IEC/TS 62592:2012(E)

THIS PUBLICATION IS COPYRIGHT PROTECTED
Copyright © 2012 IEC, Geneva, Switzerland

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 IEC or IEC's member National Committee in the country of the requester.
If you have any questions about IEC copyright or have an enquiry about obtaining additional rights to this publication, please contact the address below or your local IEC member National Committee for further information.

IEC Central Office
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
Fax: +41 22 919 03 00
info@iec.ch
www.iec.ch

About the IEC

The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications

The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigenda or an amendment might have been published.

Useful links:

IEC publications search - www.iec.ch/searchpub

The advanced search enables you to find IEC publications by a variety of criteria (reference number, text, technical committee,…).
It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished

Stay up to date on all new IEC publications. Just Published details all new publications released. Available on-line and also once a month by email.

Electropedia - www.electropedia.org

The world's leading online dictionary of electronic and electrical terms containing more than 30 000 terms and definitions in English and French, with equivalent terms in additional languages. Also known as the International Electrotechnical Vocabulary (IEV) on-line.

Customer Service Centre - webstore.iec.ch/csc

If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: csc@iec.ch.

This document is a preview generated by EVS

IEC/TS 62592

Edition 2.0 2012-07

TECHNICAL SPECIFICATION

This document is a preview generated by EVS

Encoding guidelines for portable multimedia CE products using MP4 file format with AVC video codec and AAC audio codec

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

PRICE CODE XD

ICS 33.160.01

ISBN 978-2-83220-220-3

Warning! Make sure that you obtained this publication from an authorized distributor.

® Registered trademark of the International Electrotechnical Commission

CONTENTS

FOREWORD 6
0 Introduction 8
0.1 MP4 file format 8
0.2 Issues to be considered in implementation on portable CE devices 8
0.3 Encoding rules for open system and application 9
1 Scope 10
2 Normative references 10
3 Terms, definitions, abbreviations and conventions 11
3.1 Terms and definitions 11
3.2 Abbreviations 13
3.3 Conventions 13
3.3.1 Method of presenting box definition 13
3.3.2 Unit prefixes 14
3.3.3 Numeric representation 14
3.3.4 Character data 14
3.3.5 Nominal time length 14
3.3.6 Reserved fields and values 14
4 Design rules 14
4.1 General 14
4.2 File structure 15
4.2.1 General 15
4.2.2 Movie structure 16
4.2.3 Fragmented movie structure 17
4.3 Box order 18
4.4 Operational rules for boxes defined in ISO file format standards 20
4.4.1 General 20
4.4.2 Size fields 20
4.4.3 Versions and flags 20
4.4.4 File identification 20
4.4.5 File extension 21
4.4.6 Handler reference types 21
4.4.7 Movie header box 22
4.4.8 Track header box 22
4.4.9 Media header box 22
4.4.10 Sample size box 22
4.4.11 Time to sample boxes 22
4.4.12 Sample entries 22
4.4.13 Movie fragment boxes 22
4.4.14 Track fragment random access boxes – Version fields 23
4.4.15 Edit list entries 23
4.4.16 Template fields used 23
4.5 Additional definitions 24
4.5.1 Private extension box definitions 24
4.5.2 Profile configuration 30
4.5.3 Function flags 30

This document is a preview generated by EVS

4.5.4 Codec specific information 31
4.5.5 Attribute flags 33
4.6 Operational rules for tracks 33
4.6.1 General 33
4.6.2 Main audio track 34
4.6.3 Main video track 35
4.6.4 Thumbnail picture track 36
4.7 Operational rules for media data 38
4.7.1 MPEG-4 AAC elementary stream 38
4.7.2 AVC video elementary stream 41
4.7.3 JPEG image sequence for thumbnail pictures 66
4.8 Logical structure of media data 67
4.8.1 General 67
4.8.2 Interleaving of main audio and video 67
4.8.3 Arrangement of data chunks 68
4.8.4 Placement of thumbnail picture 68
4.8.5 Media data structure in movie fragment 68
4.8.6 Exception for logical structure of media data 68
4.9 Presentation arrangement for tracks 68
4.9.1 Alternate group support 68
4.9.2 Edit list support 69
4.10 Metadata 70
4.10.1 Metadata types 70
4.10.2 Metadata types definition 71
4.11 Pixel aspect ratios 74
5 PCE AV classes 76
5.1 Class 76
5.2 Class PT 76
5.2.1 General 76
5.2.2 Level 1 77
5.2.3 Level 2 77
5.2.4 Level 3 77
5.3 Class MB 77
5.3.1 General 77
5.3.2 Level 1 78
5.3.3 Level 2 78
5.3.4 Level 3 79
5.3.5 Level 4 79
5.4 Class SD 79
5.5 Class network SD 80
5.6 Class HD 81
5.6.1 General 81
5.6.2 Level 1 82
5.6.3 Level 2 82
5.7 Class network HD 82
5.7.1 General 82
5.7.2 Level 1 83
5.7.3 Level 2 83
5.7.4 Level 3 84

This document is a preview generated by EVE

6 Definitions for other files 84
Annex A (normative) Operation rules for compliant products 85
Annex B (normative) Capabilities required for compliant products 89
Annex C (informative) Recommended recording modes 91
Annex D (normative) Operation rules for rotated video file 96
Bibliography 98

Figure 1 – Example of a simple interchange file 15
Figure 2 – Examples of movies 17
Figure 3 – Example of a fragmented movie 18
Figure 4 – Example of AAC bitstream 38
Figure 5 – Example of a closed GOP 52
Figure 6 – Example of an open GOP 52
Figure 7 – Reference structure of a reference B-picture 60
Figure 8 – Reference structure of a non-reference B-picture 61
Figure 9 – Example of a closed GOP 61
Figure 10 – Example of an open GOP 62

Table 1 – Supported track types 16
Table 2 – Supported stream and data 16
Table 3 – Box types, structure, and cross-reference 19
Table 4 – brand identifier 21
Table 5 – handler_types 22
Table 6 – Common portion of the audio attribute flags 26
Table 7 – Common portion of the video attribute flags 27
Table 8 – encoding_type 29
Table 9 – Function flags 30
Table 10 – Data form type 31
Table 11 – Data form type 32
Table 12 – Codec specific portion of audio attribute flags 33
Table 13 – Codec specific portion of video attribute flags 33
Table 14 – The values of pic_struct and ct_type 54
Table 15 – Allowed values of frame_mbs_only_flag, pic_width_in_mbs_minus1 and pic_height_in_map_units_minus1 63
Table 16 – Allowed values of frame cropping related fields 63
Table 17 – data_type_ID 71
Table 18 – Presentation types 72
Table 19 – Track attribute flags 73
Table 20 – Usage of the metadata 74
Table 21 – Resolution values in visual sample entry 75
Table 22 – spacing values in pixel_aspect_ratio 75
Table C.1 – Recommended recording modes for Class PT 91
Table C.2 – Recommended recording modes for Class MB 92

This document is a preview generated by EVS

Table C.3 – Recommended recording modes for Class SD 92
Table C.4 – Parameters for progressive video 25[fps] and 29.97[fps] 93
Table C.5 – Recommended recording modes for Class network SD 93
Table C.6 – Recommended recording modes for Class HD 94
Table C.7 – Recommended recording modes for Class network HD 95
Table D.1 – Width and height in the track header box 97

This document is a preview generated by EVS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

ENCODING GUIDELINES FOR PORTABLE MULTIMEDIA CE PRODUCTS USING MP4 FILE FORMAT WITH AVC VIDEO CODEC AND AAC AUDIO CODEC

FOREWORD

1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
5) IEC provides no marking procedure to indicate its approval and cannot be rendered responsible for any equipment declared to be in conformity with an IEC Publication.
6) All users should ensure that they have the latest edition of this publication.
7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

The main task of IEC technical committees is to prepare International Standards. In exceptional circumstances, a technical committee may propose the publication of a Technical Specification when

- the required support cannot be obtained for the publication of an International Standard, despite repeated efforts, or
- the subject is still under technical development or where, for any other reason, there is the future but no immediate possibility of an agreement on an International Standard.

Technical Specifications are subject to review within three years of publication to decide whether they can be transformed into International Standards.

This Technical Specification cancels and replaces the first edition published in 2009. This edition constitutes a technical revision.

This edition includes the following significant technical changes with respect to the previous edition:

a) added 23,976 fps to all classes;

This document is provided by EVS

b) added new class (class network SD);

c) added some metadata for CE content;

d) inserted some editorial improvements and clarifications.

IEC/TS 62592, which is a Technical Specification, has been prepared by technical area 6: Storage media, data structures, equipment and systems, of IEC technical committee 100: Audio, video and multimedia systems and equipment.

The text of this Technical Specification is based on the following documents:

Enquiry draft	Report on voting
100/1926/DTS	100/1970/RVC

Full information on the voting for the approval of this Technical Specification can be found in the report on voting indicated in the above table.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

The committee has decided that the contents of this publication will remain unchanged until the stability date indicated on the IEC web site under "http://webstore.iec.ch" in the data related to the specific publication. At this date, the publication will be

- transformed into an International standard,
- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

A bilingual version of this publication may be issued at a later date.

This document is a preview generated by EVS

0 Introduction

0.1 MP4 file format

ISO base media file format, ISO/IEC 14496-12, has been developed by ISO/IEC JTC 1/SC 29/WG 11 and WG1 as a common base media file format of audio, video and image applications. It provides a file format to contain timed media information for a presentation in a flexible, extensible format. The extensions to support specific codecs and systems are specified as the series of standard ISO/IEC 14496. The MP4 file format (MPEG-4 file format), ISO/IEC 14496-14, is an extension to support MPEG-4 systems in ISO base media file format. The extension to support Advanced Video Coding (AVC), ISO/IEC 14496-10 and ITU-T H.264[1], is standardized as AVC file format, ISO/IEC 14496-15. In this Technical Specification, the family of ISO base media file format is referred to as MP4 file format, which is the name widely used in the industry. The MP4 file format is designed very flexibly so that the series of the standard can be applied to various kinds of applications and can bring the maximum performances for the applications.

MP4 file format is adopted by various Consumer Electronics (CE) devices, e.g. broadcasting receivers, disc recorders / players, AV content distribution, portable AV recorders / players, and so on. In the CE audio, video and multimedia applications, MP4 file format with Advanced Video Coding (AVC), (ISO/IEC 14496-10 and ITU-T H.264) and Advanced Audio Coding (AAC), (ISO/IEC 13818-3 and ISO/IEC 14496-3), is employed most popularly.

AVC (ISO/IEC 14496-10 and ITU-T H.264) specifies profiles and levels to ensure interoperability of decoding procedure of a video elementary stream, which fits various levels of application. AVC, ISO/IEC 14496-10 and ITU-T H.264, specifies the bitstream syntax and its decoding process.

0.2 Issues to be considered in implementation on portable CE devices

To implement applications using MP4 file format into CE products, the characteristics of CE industry must be considered. The characteristics of CE industry and products are different from IT equipment especially personal computers (PCs) which have strong computational power and flexible software solution. That is:

- CE devices are designed with limited resources, processing power and memory size;
- most CE devices can not update software or hardware afterwards; and
- manufacturers have to assure the capability and the quality of the product for consumers, who are not familiar with processes inside the device;
- CE devices need to support functionalities, e.g. fast forward / backward play which are widely supported by existing CE devices. Manufacturers have to ensure such functionalities regardless of creators of the content (bitstreams).

Given these criteria, manufacturers need to check all the performance of the products because any defect in the products may generate serious complaints among consumers. Since the MP 4 file format is so flexible, the number of combination of setting parameters is very large and it takes tremendous workload and cost to check the performance for each combination of parameters. Even AVC (ISO/IEC 14496-10 and ITU-T H.264) specifies the decoding procedure and conformance point by a profile and a level. Furthermore, the constraints on bitstreams must be specified in order to ensure functionalities, which are widely supported by CE devices. Therefore, in most CE systems, the combination of encoding parameters is limited and the dedicated encoding rules are specified for the system, which enable designing and manufacturing process practical to guarantee the quality of the product. In general, these encoding rules are proprietary to the involved parties.

[1] ITU-T H.264 is equivalent to ISO/IEC 14496-10.

This preview is generated by IEC

0.3 Encoding rules for open system and application

Currently, the content of audio, video and multimedia products are provided by commercial content providers through specific sales channels such as optical discs, CDs, DVDs and Blu-ray Discs. Additional content is provided by broadcasters and commercial content providers. However, more recently end-user generated content is increasing and placed on many Internet sites enabled by the availability of digital video cameras and Internet related technology. In this situation, portable CE devices are required to store and reproduce such content for consumer satisfaction. Accordingly, portable CE devices need to guarantee the capability to decode MP4 files with AVC (ISO/IEC 14496-10 and ITU-T H.264) and AAC (ISO/IEC 13818-3 and ISO/IEC 14496-3) which are most commonly used as the file format and codecs.

- To assure decoding of such MP4 files, CE device manufacturers have to check the performance for the files encoded by unknown parameters. However, it is impossible to achieve this with limited manpower and technical resources as mentioned above. This situation may cause a problem in decoding quality and is not beneficial for both end-users and manufacturers. CE device manufacturers also have to ensure functionalities, e.g. fast forward / backward play, regardless of the creators of the content (bitstreams).

To solve this issue, this Technical Specification specifies the encoding rules for the MP4 files targeted to be stored and reproduced by portable CE devices so that the decoding of the content encoded by these guidelines is guaranteed by the portable CE product manufacturer with reasonable cost and resources.

The encoding rules are specified considering the capability of current portable CE products. However, the rules should be revised depending on the progress of CE technologies in an appropriate time frame.

This document is a preview generated by EVS

ENCODING GUIDELINES FOR PORTABLE MULTIMEDIA CE PRODUCTS USING MP4 FILE FORMAT WITH AVC VIDEO CODEC AND AAC AUDIO CODEC

This document is a preview generated by EVS

1 Scope

This Technical Specification specifies encoding guidelines for portable multimedia CE products using ISO base media file format and its family (ISO/IEC 14496-12, ISO/IEC 14496-14, and ISO/IEC 14496-15) with AVC[2] (ISO/IEC 14496-10 and ITU-T H.264) and AAC[2] (ISO/IEC 13818-3 and ISO/IEC 14496-3). These guidelines may also be applicable to portable non-CE products.

This Technical Specification is applicable to the creation of MP4 files with AVC and AAC which are intended for storage, reproduction and display by portable CE products. It is applicable to both content generation software and hardware.

These guidelines are intended to encourage global interoperability between portable CE products by recommending the use of specific parameters to allow efficient interworking of devices which may have limited resources so that the decoding of content encoded in accordance with these guidelines is assured. In addition, the guidelines provide for simplified testing and verification of the interoperability of portable CE products.

2 Normative references

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

ISO 639-2:1998, *Codes for the representation of names of languages – Part 2: Alpha-3 code*

ISO/IEC 10646:2012, *Information technology – Universal Coded Character Set (UCS)*

ISO/IEC 10918-1, *Information technology – Digital compression and coding of continuous-tone still images – Part 1: Requirements and guidelines*

ISO/IEC 13818-3, *Information technology – Generic coding of moving pictures and associated audio information – Part 3: Audio*

ISO/IEC 14496-1:2010, *Information technology – Coding of audio-visual objects – Part 1: Systems*

ISO/IEC 14496-2:2004, *Information technology – Coding of audio-visual objects – Part 2: Visual*

ISO/IEC 14496-3:2009, *Information technology – Coding of audio-visual objects – Part 3: Audio*

ISO/IEC 14496-10:2012, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*

---

2 The full wording for AAC and AVC can be found in 3.2.

ISO/IEC 14496-12:2008, *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 14496-15:2010, *Information technology – Coding of audio-visual objects – Part 15: Advanced Video Coding (AVC) file format*

ISO/IEC 15444-12, *Information technology – JPEG 2000 image coding system – Part 12: ISO base media file format*

3 Terms, definitions, abbreviations and conventions

3.1 Terms and definitions

For the purposes of this document, the terms and definitions specified in ISO/IEC 14496-1, ISO/IEC 14496-2, ISO/IEC 14496-3, ISO/IEC 14496-10, ISO/IEC 14496-12 and ISO/IEC 14496-15, as well as the following apply.

3.1.1
B-field
field consisting only of B slices

3.1.2
B-frame
frame consisting only of B slices

3.1.3
bitrate
rate at which the coded bitstream is delivered to the input of a decoder

Note 1 to entry: Bitrate is an index to indicate how many bits are transmitted per second.

Note 2 to entry: The unit is bps (bits per second).

3.1.4
B-picture
frame or a complementary field pair consisting only of B slices

3.1.5
class
video and audio data formats and constraints on the values of the syntax

3.1.6
fragmented movie
movie that consists of a single Movie Box (‘moov’), Movie Fragment Boxes (‘moof’), and Media Data Boxes (‘mdat’) that contain fragmented media data

Note 1 to entry: The Movie Box of a fragmented movie is referred to as an initial Movie Box to distinguish it from that of an ordinary movie in this Technical Specification. An initial ‘moov’ box contains a Movie Extends Box (‘mvex’) that indicates that the movie is possibly fragmented, which means ‘moof’s should be found and used in the file.

3.1.7
frame rate
rate at which frames are to be output from the composition process

Note 1 to entry: Frame rate is an index to indicate how many frames can be drawn per second in video playback.

This document is a preview generated by EVS