

Australian Standard™

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

**Part 15: Advanced video coding (AVC)
file format**



This Australian Standard was prepared by Committee IT-029, Coded Representation of Picture, Audio and Multimedia/Hypermedia Information. It was approved on behalf of the Council of Standards Australia on 27 October 2004. This Standard was published on 25 November 2004.

The following are represented on Committee IT-029:

Australian Broadcasting Authority
Australian Broadcasting Corporation
Australian Consumer Association
Australian Subscription Television
CSIRO Information and Communication Technology Centre
DSTC
Department of Defence (Australia)
Free TV Australia
Special Broadcasting Service
The University of New South Wales
University of Sydney
University of Wollongong
Victoria University of Technology

Keeping Standards up-to-date

Standards are living documents which reflect progress in science, technology and systems. To maintain their currency, all Standards are periodically reviewed, and new editions are published. Between editions, amendments may be issued. Standards may also be withdrawn. It is important that readers assure themselves they are using a current Standard, which should include any amendments which may have been published since the Standard was purchased.

Detailed information about Standards can be found by visiting the Standards Web Shop at www.standards.com.au and looking up the relevant Standard in the on-line catalogue.

Alternatively, the printed Catalogue provides information current at 1 January each year, and the monthly magazine, *The Global Standard*, has a full listing of revisions and amendments published each month.

Australian Standards™ and other products and services developed by Standards Australia are published and distributed under contract by SAI Global, which operates the Standards Web Shop.

We also welcome suggestions for improvement in our Standards, and especially encourage readers to notify us immediately of any apparent inaccuracies or ambiguities. Contact us via email at mail@standards.org.au, or write to the Chief Executive, Standards Australia International Ltd, GPO Box 5420, Sydney, NSW 2001.

This Standard was issued in draft form for comment as DR 04416.

Australian Standard™

Information technology—Coding of audio-visual objects

Part 15: Advanced video coding (AVC) file format

First published as AS ISO/IEC 14496.15—2004.

COPYRIGHT

© Standards Australia International

All rights are reserved. No part of this work may be reproduced or copied in any form or by any means, electronic or mechanical, including photocopying, without the written permission of the publisher.

Published by Standards Australia International Ltd GPO Box 5420, Sydney, NSW 2001, Australia

ISBN 0 7337 6369 3

PREFACE

This Standard was prepared by the Standards Australia Committee IT-029, Coded Representation of Picture, Audio and Multimedia/Hypermedia Information.

This Standard is identical with, and has been reproduced from, ISO/IEC 14496-15:2004, *Information technology—Coding of audio-visual objects—Part 15: Advanced video coding (AVC) file format*.

The objective of this Standard is to provide the multimedia designer with the necessary information needed to use and implement Advanced Video Coding (AVC) video streams for streaming servers or other media streams such as audio. Its advantage compared to other video coding systems is that it suits both low, medium and high bit-rates. The Advanced Video Coding (AVC) file format defines a storage format for video streams compressed using AVC.

As this Standard is reproduced from an international standard, the following applies:

- (a) Its number appears on the cover and title page while the international standard number appears only on the cover.
- (b) In the source text ‘this International Standard’ should read ‘this Australian Standard’.
- (c) A full point substitutes for a comma when referring to a decimal marker.

References to International Standards should be replaced by references to Australian or Australian/New Zealand Standards, as follows:

<i>Reference to International Standard</i>		<i>Australian Standard/New Zealand Standard</i>	
ISO/IEC		AS/NZS	
14496	Information technology—Coding of audio-visual objects	14496	Information technology—Coding of audio-visual objects
14496-1	Part 1: Systems	14496.1	Part 1: Systems
		AS	
14496-10	Part 10: Advanced video coding	14496.10	Part 10: Advanced video coding
14496-12	Part 12: ISO base media file format	14496.12	Part 12: ISO base media file format

Only referenced documents that have been adopted as Australian or Australian/New Zealand Standards have been listed.

CONTENTS

Page

Introduction	iv
1 Scope.....	1
2 Normative references	1
3 Terms, definitions, symbols and abbreviated terms.....	1
3.1 Terms and definitions.....	1
3.2 Symbols and abbreviated terms.....	2
4 Extensions to the ISO Base Media File Format.....	2
4.1 Introduction	2
4.2 File identification.....	2
4.3 Independent and Disposable Samples Box	2
4.4 Sample groups	3
4.4.1 Introduction	3
4.4.2 SampleToGroup Box	4
4.4.3 SampleGroupDescription Box	5
4.5 Random access recovery points	6
4.5.1 Syntax.....	6
4.5.2 Semantics	6
4.6 Representation of new structures in movie fragments.....	7
5 AVC elementary streams and sample definitions	7
5.1 Elementary stream structure	7
5.2 Sample and Configuration definition	9
5.2.1 Introduction	9
5.2.2 Canonical order and restrictions.....	9
5.2.3 AVC sample structure definition	11
5.2.4 Decoder configuration information.....	11
5.3 Derivation from ISO Base Media File Format	13
5.3.1 Introduction	13
5.3.2 AVC File type and identification	13
5.3.3 AVC Track Structure	13
5.3.4 AVC Video Stream Definition	13
5.3.5 AVC parameter set stream definition	15
5.3.6 Template fields used.....	16
5.3.7 Visual width and height.....	16
5.3.8 Parameter sets	17
5.3.9 Decoding time (DTS) and composition time (CTS).....	17
5.3.10 Sync sample (IDR).....	17
5.3.11 Shadow sync	17
5.3.12 Layering and sub-sequences	18
5.3.13 Alternate streams and switching pictures.....	21
5.3.14 Random access recovery points	23
5.3.15 Hinting	23

INTRODUCTION

The Advanced Video Coding (AVC) standard, jointly developed by the ITU-T and ISO/IEC SC29/WG11 (MPEG), offers not only increased coding efficiency and enhanced robustness, but also many features for the systems that use it. To enable the best visibility of, and access to, those features, and to enhance the opportunities for the interchange and interoperability of media, this part of ISO/IEC 14496 defines a storage format for video streams compressed using AVC.

This part of ISO/IEC 14496 defines a storage format based on, and compatible with, the ISO Base Media File Format (ISO/IEC 14496-12 and ISO/IEC 15444-12), which is used by the MP4 file format (ISO/IEC 14496-14) and the Motion JPEG 2000 file format (ISO/IEC 15444-3) among others. This part of ISO/IEC 14496 enables AVC video streams to:

- be used in conjunction with other media streams, such as audio;
- be used in an MPEG-4 systems environment, if desired;
- be formatted for delivery by a streaming server, using hint tracks;
- inherit all the use cases and features of the ISO Base Media File Format on which MP4 and MJ2 are based.

This part of ISO/IEC 14496 may be used as a standalone specification; it specifies how AVC content shall be stored in an ISO Base Media File Format compliant format. However, it is normally used in the context of a specification, such as the MP4 file format, derived from the ISO Base Media File Format, that permits the use of AVC video.

The ISO Base Media File Format is becoming increasingly common as a general-purpose media container format for the exchange of digital media, and its use in this context should accelerate both adoption and interoperability.

Extensions to the ISO Base Media File Format are defined here to support the new systems aspects of the AVC codec.

AUSTRALIAN STANDARD

Information technology — Coding of audio-visual objects —**Part 15:****Advanced Video Coding (AVC) file format****1 Scope**

This part of ISO/IEC 14496 specifies the storage format for AVC (ISO/IEC 14496-10 | ITU-T Rec. H.264) video streams.

The storage of AVC content uses the existing capabilities of the ISO Base Media File Format but also defines extensions to support the following features of the AVC codec:

- **Switching pictures:** To enable switching between different coded streams and substitution of pictures within the same stream.
- **Sub-sequences and layers:** Provides a structuring of the dependencies of a group of pictures to provide for a flexible stream structure (e.g. in terms of temporal scalability and layering).
- **Parameter sets:** The sequence and picture parameter set mechanism decouples the transmission of infrequently changing information from the transmission of coded macroblock data. Each slice containing the coded macroblock data references the picture parameter set containing its decoding parameters. In turn, the picture parameter set references a sequence parameter set that contains sequence level decoding parameter information.

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 14496-1:2001, *Information technology — Coding of audio-visual objects — Part 1: Systems*

ISO/IEC 14496-10, *Information technology — Coding of audio-visual objects — Part 10: Advanced video coding* | ITU-T Rec. H.264, *Advanced video coding for generic audiovisual services*

ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format* (technically identical to ISO/IEC 15444-12)

3 Terms, definitions, symbols and abbreviated terms**3.1 Terms and definitions**

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-1, ISO/IEC 14496-10 | ITU-T Rec. H.264 and the following apply.

3.1.1**parameter set**

a sequence parameter set or a picture parameter set, as defined in ISO/IEC 14496-10

NOTE This term is used to refer to both types of parameter sets.

3.1.2**parameter set elementary stream**

elementary stream containing samples made up of only sequence and picture parameter set NAL units synchronized with the video elementary stream

3.1.3**video elementary stream**

elementary stream containing access units made up of NAL units for coded picture data

3.2 Symbols and abbreviated terms

AVC	Advanced Video Coding [ISO/IEC 14496-10]
HRD	Hypothetical Reference Decoder
IDR	Instantaneous Decoding Refresh
NAL	Network Abstraction Layer
PPS	Picture Parameter Set
SEI	Supplementary Enhancement Information
SPS	Sequence Parameter Set

4 Extensions to the ISO Base Media File Format**4.1 Introduction**

This clause documents technical additions to the ISO Base Media File Format, which can be used when storing AVC streams. However, these additions could also be used by other media, if they are defined to use them. They are therefore documented here separately.

4.2 File identification

The brand 'avc1' shall be used to indicate that extensions conformant with this section are used in a file. The use of 'avc1' as a major-brand may be permitted by specifications; in that case, that specification defines the file extension and required behaviour.

4.3 Independent and Disposable Samples Box

Box Types: 'sdtb'
 Container: Sample Table Box ('stbl')
 Mandatory: No
 Quantity: Exactly one

This optional table answers three questions about sample dependency:

- 1) Does this sample depend on others (is it an I-picture)?
- 2) Do no other samples depend on this one?
- 3) Does this sample contain multiple (redundant) encodings of the data at this time-instant (possibly with different dependencies)?

In the absence of this table:

- 1) the sync sample table answers the first question; in most video codecs, I-pictures are also sync points,
- 2) the dependency of other samples on this one is unknown,
- 3) the existence of redundant coding is unknown.

When performing ‘trick’ modes, such as fast-forward, it is possible to use the first piece of information to locate independently decodable samples. Similarly, when performing random access, it may be necessary to locate the previous sync point or random access recovery point, and roll-forward from the sync point or the pre-roll starting point of the random access recovery point to the desired point. While rolling forward, samples on which no others depend need not be retrieved or decoded.

The value of ‘sample-is-depended-on’ is independent of the existence of redundant codings. However, a redundant coding may have different dependencies from the primary coding; if redundant codings are available, the value of ‘sample-depends-on’ documents only the primary coding.

The size of the table, `sample_count` is taken from the `sample_count` in the Sample Size Box (‘stsz’) or Compact Sample Size Box (‘stz2’).

4.3.1.1 Syntax

```
aligned(8) class SampleDependencyTypeBox
    extends FullBox('sdtp', version = 0, 0) {
    for (i=0; i < sample_count; i++){
        unsigned int(2) reserved = 0;
        unsigned int(2) sample-depends-on;
        unsigned int(2) sample-is-depended-on;
        unsigned int(2) sample-has-redundancy;
    }
}
```

4.3.1.2 Semantics

`sample-depends-on` takes one of the following four values:

- 0: the dependency of this sample is unknown;
- 1: this sample does depend on others (not an I picture);
- 2: this sample does not depend on others (I picture);
- 3: reserved.

`sample-is-depended-on` takes one of the following four values:

- 0: the dependency of other samples on this sample is unknown;
- 1: other samples depend on this one (not disposable);
- 2: no other sample depends on this one (disposable);
- 3: reserved.

`sample-has-redundancy` takes one of the following four values:

- 0: it is unknown whether there is redundant coding in this sample;
- 1: there is redundant coding in this sample;
- 2: there is no redundant coding in this sample;
- 3: reserved.

4.4 Sample groups

4.4.1 Introduction

This clause specifies a generic mechanism for representing a partition of the samples in a track. A *sample grouping* is an assignment of each sample in a track to be a member of one *sample group*, based on a

grouping criterion. A sample group in a sample grouping is not limited to being contiguous samples and may contain non-adjacent samples. As there may be more than one sample grouping for the samples in a track, each sample grouping has a type field to indicate the type of grouping. For example, a file might contain two sample groupings for the same track: one based on an assignment of sample to layers and another to sub-sequences.

Sample groupings are represented by two linked data structures: (1) a `SampleToGroupBox` represents the assignment of samples to sample groups; (2) a `SampleGroupDescription` box contains a *sample group entry* for each sample group describing the properties of the group. There may be multiple instances of the `SampleToGroupBox` and `SampleGroupDescription` boxes based on different grouping criteria. These are distinguished by a type field used to indicate the type of grouping.

One example of using these tables is to represent the assignments of samples to *layers*. In this case each sample group represents one layer, with an instance of the `SampleToGroupBox` describing which layer a sample belongs to. For more details, please refer to 5.3.12

4.4.2 SampleToGroupBox

4.4.2.1 Definition

Box Type: 'sbgp'
 Container: Sample Table Box ('stbl')
 Mandatory: No
 Quantity: Zero or more.

This table can be used to find the group that a sample belongs to and the associated description of that sample group. The table is compactly coded with each entry giving the index of the first sample of a run of samples with the same sample group descriptor. The sample group description ID is an index that refers to a `SampleGroupDescription` box, which contains entries describing the characteristics of each sample group.

There may be multiple instances of this box if there is more than one sample grouping for the samples in a track. Each instance of the `SampleToGroupBox` has a type code that distinguishes different sample groupings. Within a track, there shall be at most one instance of this box with a particular grouping type. The associated `SampleGroupDescription` shall indicate the same value for the grouping type.

4.4.2.2 Syntax

```
aligned(8) class SampleToGroupBox
  extends FullBox('sbgp', version = 0, 0)
{
  unsigned int(32)  grouping_type;
  unsigned int(32)  entry_count;
  for (i=1; i <= entry_count; i++)
  {
    unsigned int(32)  sample_count;
    unsigned int(32)  group_description_index;
  }
}
```

4.4.2.3 Semantics

`version` is an integer that specifies the version of this box.

`grouping_type` is an integer that identifies the type (i.e. criterion used to form the sample groups) of the sample grouping and links it to its sample group description table with the same value for grouping type. At most one occurrence of this box with the same value for `grouping_type` shall exist for a track.

`entry_count` is an integer that gives the number of entries in the following table.

`sample_count` is an integer that gives the number of consecutive samples with the same sample group descriptor.