INTERNATIONAL STANDARD

ISO/IEC 23002-4

Third edition
2018-03

Information technology — MPEG video technologies —

Part 4: Video tool library

Technologies de l'information — Technologies vidéo MPEG —

Partie 4: Bibliothèque d'outils vidéo

Reference number
ISO/IEC 23002-4:2018(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

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

This third edition cancels and replaces the second edition (ISO/IEC 23002-4:2014), which has been technically revised.

It also incorporates ISO/IEC 23002-4:2014/Amd.1:2014, ISO/IEC 23002-4:2014/Amd.2:2015, and ISO/IEC 23002-4:2014/DAmd.3[1].

The changes compared to the previous edition are as follows.

— FU description convention described in Clause 4 has been technically updated.

— Functional unit (FU) and FU network (FN) descriptions for graphics tool library (GTL) have been added according to ISO/IEC 23002-4:2014/Amd.1:2014. Clause 5, Clause 9, Annexes D, E, F, and G have been added.

— FU and FN descriptions for high efficiency video coding (HEVC) have been added according to ISO/IEC 23002-4:2014/Amd.2:2015. Clause 10 has been added and Annex B has been updated.

[1] Draft amendment approved in DAM ballot but not published.

— FU and FN descriptions for parser instantiation from bitstream syntax description (BSD) have added. Clause 5 has updated and Annex I has added.

— A patent statement has been added in Annex H.

A list of all parts in the ISO/IEC 23002 series can be found on the ISO website.

Introduction

This document defines the MPEG video tool library, which contains tools drawn from existing MPEG coding standards, such as ISO/IEC 14496-2 and ISO/IEC 14496-10, and ISO/IEC 23001-4 defines the methods capable of describing codec configurations in the reconfigurable video coding (RVC) framework.

This document primarily addresses reconfigurable video aspects and will only focus on the description of representation of video codec configurations under the RVC framework, but could be extended to a more generic reconfigurable media coding (RMC) framework.

The objective of RVC is to offer a framework that is capable of configuring and specifying video codecs as a collection of "higher level" modules by using video coding tools. The video coding tools are defined in video tool libraries. This d defines the MPEG video tool library. The RVC framework principle could also support non-MPEG tool libraries, provided that their developers have taken care to obey the appropriate rules of operation.

For the purpose of framework deployment, an appropriate description is needed to describe configurations of decoders composed of or instantiated from a subset of video tools from either one or more libraries. As illustrated in Figure 1, the configuration information consists of

— bitstream syntax description, and

— network of functional units (FUs) description (also referred to as the decoder configuration)

that together constitute the entire decoder description.

Bitstreams of existing MPEG standards are specified by specific syntax structures and decoders are composed of various coding tools. Therefore, RVC includes support for bitstream syntax descriptions as well as video coding tools. As depicted in Figure 1, a typical RVC decoder requires two types of information, namely the decoder description and the encoded media (e.g. video bitstreams) data.

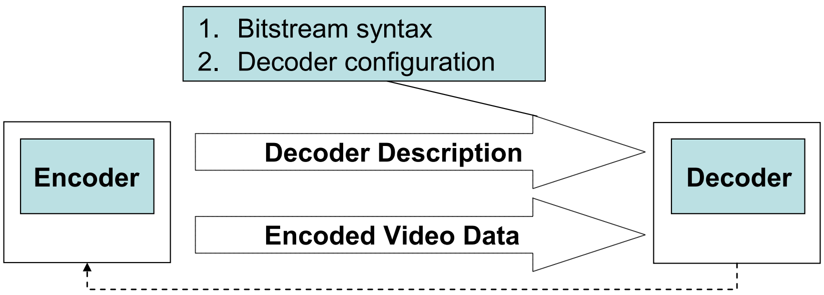

Figure 1 — Conceptual diagram of RVC

A more detailed description of the RVC decoder is illustrated in Figure 2. As shown in Figure 2, the decoder description is required for the configuration of a RVC decoder. The Bitstream Syntax Description (BSD) and FU Network Description (FND) (which compose the Decoder Description) are used to configure or compose an abstract decoder model (ADM) which is instantiated through the selection of FUs from tool libraries optionally with proper parameter assignment. Such ADM constitutes the behavioral reference model used in setting up a decoding solution under the RVC framework. The process of yielding a decoding solution may vary depending on the technologies used for the desired

implementations. Examples of the instantiation of an ADM and generation of proprietary decoding solutions can be found in ISO/IEC 23001-4.

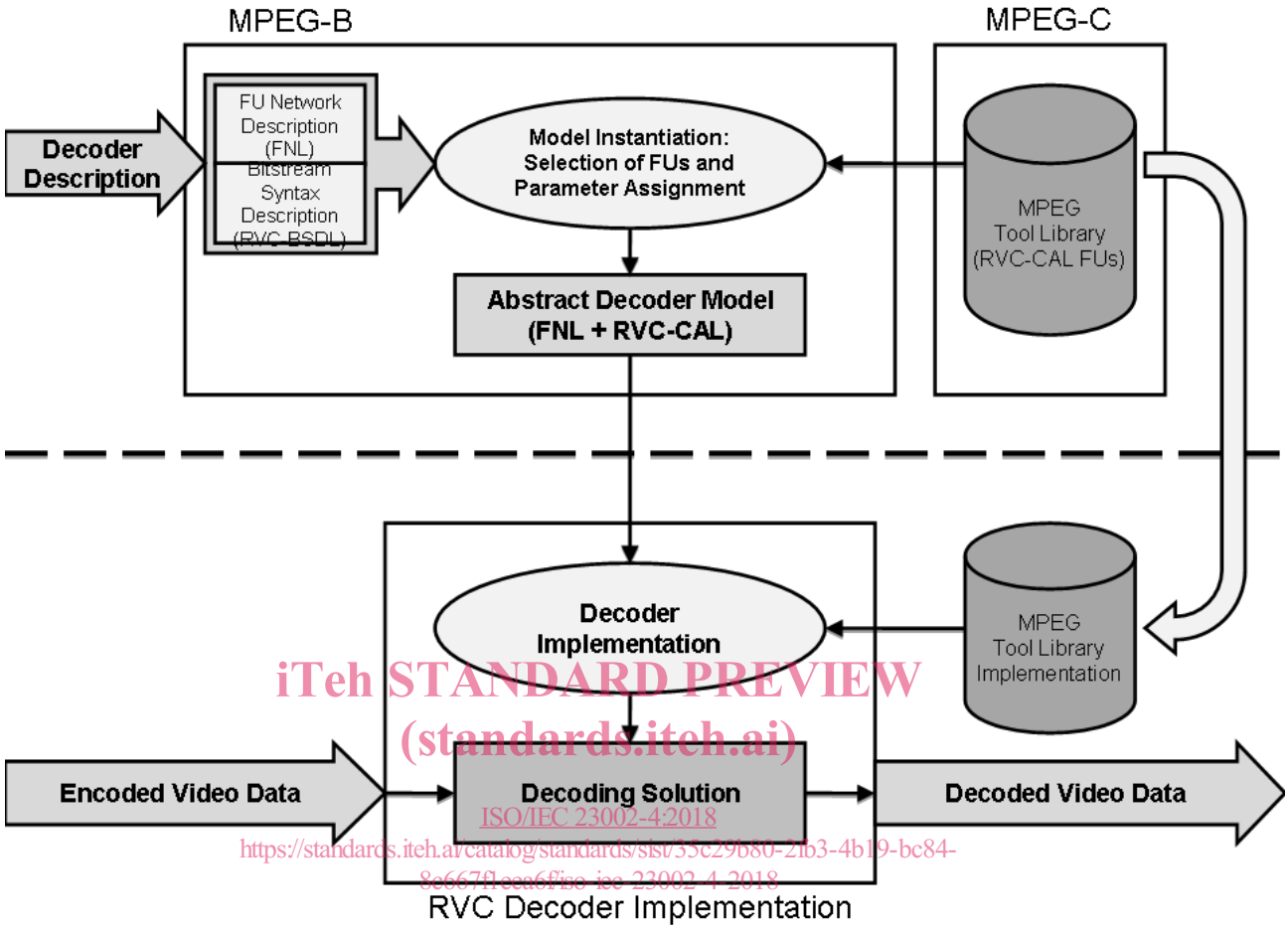

Figure 2 — Graphical representation of the process for setting up a decoding solution under the RVC framework

Within the RVC framework, the decoder description describes a particular decoder configuration and consists of the FND and the BSD. The FND describes the connectivity of the network of FUs used to form a decoder whereas the parsing process for the bitstream syntax is implicitly described by the BSD. These two descriptions are specified using two standard XML-based languages or dialects.

— Functional unit network language (FNL) is a language that describes the FND, known also as "network of FUs". The FNL specified normatively within the scope of the RVC framework is provided in ISO/IEC 23001-4.

— Bitstream syntax description language (BSDL), standardized in ISO/IEC 23001-5 (MPEG-B Part 5), describes the bitstream syntax and the parsing rules. A pertinent subset of this BSDL named RVC-BSDL is defined within the scope of the current RVC framework. This RVC-BSDL also includes possibilities for further extensions, which are necessary to provide complete description of video bitstreams. RVC-BSDL specified normatively within the scope of the RVC framework is provided in ISO/IEC 23001-4.

The decoder configuration specified using FNL, together with the specification of the bitstream syntax using RVC-BSDL fully specifies the ADM and provides an “executable” model of the RVC decoder description.

The instantiated ADM includes the information about the selected FUs and how they should be connected. As already mentioned, the FND with the network connection information is expressed by using FNL. Furthermore, the RVC framework specifies and uses a dataflow-oriented language called RVC-CAL for describing FUs' behavior. The normative specification of RVC-CAL is provided in ISO/IEC 23001-4. The ADM is the behavioural model that should be referred to in order to implement any RVC conformant decoder. Any RVC compliant decoding solution/implementation can be achieved by using proprietary non-normative tools and mechanisms that yield decoders that behave equivalent to the RVC ADM.

The decoder description, the MPEG tool library, and the associated instantiation of an ADM are normative. More precisely, the ADM is intended to be normative in terms of a behavioural model. In other words, what is normative is the input/output behaviour of the complete ADM, as well as the input/output behaviour of all the FUs that are included in the ADM.

A statement concerning patents is included in Annex H of this document.

Information technology — MPEG video technologies —

Part 4:
Video tool library

1 Scope

This document defines the description of the MPEG video tool library (VTL) based on the decoder description specified in ISO/IEC 23001-4. This tool library defines the specification of FUs, which are sufficient to build complete decoding solutions according to the following coding standards:

— ISO/IEC 14496-2 (MPEG-4 Simple Profile),

— ISO/IEC 14496-10 (MPEG-4 AVC Constrained Baseline Profile and Progressive High Profile),

— ISO/IEC 14496-16 (MPEG-4 SC3DMC), and

— ISO/IEC 23008-2 (HEVC Main Profile).

The objective of ISO/IEC 23001-4 is to define the general framework principles, and this document defines the MPEG VTL that includes relevant tools (or FUs) from the existing MPEG coding standards. Each FU is defined in the form of a textual description, which can be found in 4.1. The input and output behaviour follows the conventions described in Clause 5, Clause 6, and Clause 7.

This document compliant implementations can be designed using any software or hardware language and components. The reference software for the textual specification of FUs is written in RVC-CAL language of which a formal syntax is provided in ISO/IEC 23001-4.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 14496-2, *Information technology — Coding of audio-visual objects — Part 2: Visual*

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

ISO/IEC 14496-16, *Information technology — Coding of audio-visual objects — Part 16: Animation Framework eXtension (AFX)*

ISO/IEC 23001-4, *Information technology — MPEG systems technologies — Part 4: Codec configuration representation*

ISO/IEC 23008-2, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 23001-4 apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

4 FU description convention

4.1 FU interfaces

As shown in Table 1, each FU is described with the following elements.

— **FU Name**: Name to represent the functional unit in this specification. The name of the FU is normative and follows the naming convention described in Annex A.

— **Description**: Textual explanation to describe the functionality of the FU. The description should be concise. The precise normative behaviour of the algorithm (input/output, timing etc.) is specified by the RVC-CAL reference code in ISO/IEC 23002-4:2014/Amd 1:2014.

— **Profiles@levels supported**: The profiles@level supported for this functional unit. It may append that a given range of values makes the FU behave for a given profile@level and another range of values makes the FU behave for another profile@level.

— **Input**: A token that is entering the FU through the designated input port. The token type refers to the token pool described in 4.3. The “name” field indicates the input port.

— **Output**: A token that is coming out of the FU through the designated output port. The “name” field indicates the output port.

— **Parameter (*optional*)**: Parameters are optionally described to adjust the behaviour of the FU. All the parameters shall be specified with name, description and range.

— In several FU diagrams, the ports are named with a trailing “_i” for the input port type and with a trailing “_o” for the output port type.

— Some FU diagrams contain as well the Finite State Machine diagram. The following conventions apply: INPUT - the action of reading a token or a set of tokens from the input port, OUTPUT - the action of writing the token or a set of tokens to an output port.

— “Parameter” is set at network configuration stage (cannot be changed during the process) and it is characteristic for each FU.

— Token RANGE: describes the mathematical interval for the token value

EXAMPLE

Token RANGE: { 0, 1} – binary value.

Token RANGE: [0 .. N], value ∈ [0, N] – real values, closed interval.

— All the FUs require the data to be in little-endian format.

Table 1 — Template of description of an FU (example)

FU Name	e.g. Algo_IDCT2D_ISOIEC_23002_1	
Description	e.g. This module computes the 8×8 Inverse Discrete Cosine Transform (IDCT) defined as $f(x,y)=\frac{2}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}C(u)C(v)F(u,v)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N}$ with u, v, x, y = 0, 1, 2, ..., N-1 where x, y are spatial coordinates in the sample domain u, v are coordinates in the transform domain $C(u),C(v)=\begin{cases}\frac{1}{\sqrt{2}} & \text{for } u,v=0\\ 1 & \text{otherwise}\end{cases}$ It inputs a list of 64 coefficients and outputs a list of 64 decoded coefficients.	
Profiles@levels supported	e.g. MPEG-4 SP	
Input		
Name	**Token**	
e.g. X	e.g. BLOCK token	
Output		
Name	**Token**	
e.g. Y	e.g. BLOCK token	
Parameter		
Name	**Description**	**Range**

4.2 FU IDs

FU of the specific functionality is identified by its unique identification number. Table 2 lists IDs and names of all FUs in VTL. IDs and names are used in FND to select FUs.

Table 2 — List of FUs and their IDs

ID	FU Name
1	org.sc29.wg11.common.Algo_SynP_Generic
2	org.sc29.wg11.mpeg4.part2.sp.parser.Algo_SynP
3	org.sc29.wg11.mpeg4.part2.sp.parser.Mgnt_BlockExpand
4	org.sc29.wg11.mpeg4.part2.sp.parser.Mgnt_Splitter420B
5	org.sc29.wg11.mpeg4.part2.sp.parser.Mgnt_Splitter420MV
6	org.sc29.wg11.mpeg4.part2.sp.parser.Algo_MVR_MedianOfThreeLeftAndTopAndTopRight
7	org.sc29.wg11.mpeg4.part2.sp.parser.Algo_MVSequence_LeftAndTopAndTopRight
9	org.sc29.wg11.mpeg4.part2.sp.parser.Mgnt_Splitter_420_TYPE
10	org.sc29.wg11.mpeg4.part2.sp.parser.vlc.Algo_VLDtableB6_MPEG4Part2
11	org.sc29.wg11.mpeg4.part2.sp.parser.vlc.Algo_VLDtableB7_MPEG4Part2
12	org.sc29.wg11.mpeg4.part2.sp.parser.vlc.Algo_VLDtableB8_MPEG4Part2
13	org.sc29.wg11.mpeg4.part2.sp.parser.vlc.Algo_VLDtableB12_MPEG4Part2
14	org.sc29.wg11.mpeg4.part2.sp.parser.vlc.Algo_VLDtableB13_MPEG4Part2
15	org.sc29.wg11.mpeg4.part2.sp.parser.vlc.Algo_VLDtableB14_MPEG4Part2