DRAFT International Standard

ISO/IEC DIS 23888-2

Information technology — Artificial intelligence for multimedia —

Part 2: Video coding for machines

ICS: 35.040.40; 35.240.01

ISO/IEC JTC **1**/SC **29**

Secretariat: **JISC**

Voting begins on:
2026-04-24

Voting terminates on:
2026-07-17

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENTS AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

This document is circulated as received from the committee secretariat.

Reference number
ISO/IEC DIS 23888-2:2026(en)

© ISO/IEC 2026



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2026

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

Published in Switzerland

Contents

Page

Foreword ..... **v**

1 Scope **1**

2 Normative references **1**

3 Terms and definitions **1**

4 Abbreviations **3**

5 Conventions **3**
5.1 General 3
5.2 Arithmetic operators 3
5.3 Logical operators 4
5.4 Relational operators 4
5.5 Bit-wise operators 4
5.6 Assignment operators 5
5.7 Range notation 5
5.8 Mathematical functions 5
5.9 Order of operation precedence 8
5.10 Variables, syntax elements and tables 8
5.11 Text description of logical operations 9
5.12 Processes 11
5.13 Tensor notation 11
5.14 Picture notation 11
5.15 Dictionary notation 12
5.16 Parameter initialization 12
5.17 Floating-point number operation 12
5.18 Tensor processes 12
5.18.1 Quantized convolution 12
5.18.2 Quantized transposed convolution 13
5.18.3 Parametric rectified linear unit 13
5.18.4 Approximated sigmoid 14
5.18.5 Bilinear interpolation 14
5.18.6 Concatenation 15
5.18.7 Custom bilinear interpolation 16
5.18.8 Quantized convolution with activation 17
5.18.9 Incremental convolution 17
5.18.10 Stacked convolution with activation 17
5.18.11 Shrink transposed convolution 18
5.18.12 SILU 18
5.18.13 Linear space 18
5.18.14 ReLU 18
5.18.15 Approximated Softmax 19
5.18.16 Transpose 19
5.18.17 Flatten 19
5.18.18 LinearNorm 20
5.18.19 MaxPool2D 20
5.18.20 Grid sample 20
5.18.21 Padding 20
5.19 Format conversion processes 21
5.19.1 YCbCr420 to YCbCr400 process 21
5.19.2 YCbCr420 to YCbCr444 process 21
5.19.3 YCbCr420 to RGB444 process 22
5.19.4 RGB444 to YCbCr420 process 23

6 Bitstream and picture formats **24**
6.1 Bitstream formats 24
6.2 Source, decoded and output picture formats 24

© ISO/IEC 2026 – All rights reserved

7 Syntax and semantics **25**
7.1 Method of specifying syntax in tabular form 25
7.2 Specification of syntax functions and descriptors 27
7.3 Syntax in tabular form 27
7.3.1 General 27
7.3.2 VCM NAL unit syntax 28
7.3.3 VCM parameter set NAL unit payload syntax 29
7.3.4 VCM restoration data NAL unit payload syntax 30
7.3.5 Restoration data structure syntax 32
7.3.6 VCM temporal synthesis picture data NAL unit payload syntax 34
7.3.7 VCM end of bitstream NAL unit payload syntax 35
7.3.8 Filler data NAL unit payload syntax 35
7.3.9 Supplemental enhancement information message syntax 35
7.4 Semantics 35
7.4.1 General 35
7.4.2 VCM NAL unit semantics 36
7.4.3 VCM parameter set NAL unit payload semantics 38
7.4.4 VCM restoration data NAL unit payload semantics 40
7.4.5 Restoration data structure semantics 43
7.4.6 VCM temporal synthesis picture data NAL unit payload semantics 46
7.4.7 Filler data payload semantics 47
7.4.8 VCM NAL unit semantics Supplemental enhancement information message semantics 47
7.4.9 Coded video data semantics 48

8 Decoding process **48**
8.1 General 48
8.2 Decoding process of VCM NAL units 49
8.3 Decoding process of restoration and synthesis data 49
8.3.1 General restoration data decoding 49
8.3.2 Decoding process of sequence restoration and synthesis data 51
8.3.3 Decoding process of picture restoration and synthesis data 53
8.4 Decoding process of coded video data 58
8.5 Resampling, restoration, and synthesis processes 59
8.5.1 Neural network bit depth restoration 59
8.5.2 Spatial resampling 60
8.5.3 Spatial restoration 66
8.5.4 Temporal synthesis 72
8.5.5 Bit depth restoration 78
8.5.6 Luminance shifting and chroma synthesis 84
8.6 Format adapter process 86
8.6.1 General 86

9 Parsing process **88**
9.1 General 88
9.2 Parsing process for k-th order Exp-Golomb codes 88
9.2.1 General 88

Annex A (Normative) **Profiles, tiers and levels** **90**

Annex B (Normative) **VCM sample stream format** **93**

Annex C (Normative) **Hypothetical reference decoder** **95**

Annex D (Normative) **Supplemental enhancement information** **105**

Annex E (Normative) **Neural network based processes and decoding information** **109**

Bibliography **131**

© ISO/IEC 2026 – All rights reserved

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.

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 document 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 or www.iec.ch/members_experts/refdocs).

ISO and IEC draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC *[had/had not]* received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of 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 www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

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

A list of all parts in the ISO/IEC 23888 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

© ISO/IEC 2026 – All rights reserved

Information technology — Artificial intelligence for multimedia —

Part 2: Video coding for machines

1 Scope

This document specifies video coding technologies for machine consumption.

Conventional video coding technologies aim for the best video under certain bit-rate constraints for human consumption. However, with the emergence of artificial intelligence systems, along with the abundance of sensors, many artificial intelligence platforms have been implemented with massive data requirements including scenarios such as connected vehicles, video surveillance, and smart city. The sheer quantity of data being produced constantly leads previous methods with a human in the pipeline to be inefficient and unrealistic in terms of latency and scale. There are additional concerns in transmission and archive systems which require a more compact data representation and low latency solution. This led to the introduction of the Video Coding for Machines standard specified in this document.

In some cases, machines will communicate amongst themselves to perform tasks without a human in the mix, while in others there will be a need for additional human consumption of the specific decompressed stream. This specific scenario is possible in surveillance use cases, where a human “supervisor” occasionally searches for a specific person, or scene in video. In other cases, the corresponding bitstream is used for both human and machine consumption.

The syntax, semantics, and decoding processes that are required to decode a bitstream are described in this document.

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.

Rec. ITU-T H.266 | ISO/IEC 23090-3, *Versatile video coding*

Rec. ITU-T H.265 | ISO/IEC 23008-2, *High efficiency video coding*

Rec. ITU-T H.264 | ISO/IEC 14496-10, *Advanced video coding*

ISO/IEC 60559IEEE 754, *IEEE Standard for Floating-Point*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in Rec. ITU-T H.266 | ISO/IEC 23090-3, Rec. ITU-T H.265 | ISO/IEC 23008-2, Rec. ITU-T H.264 | ISO/IEC 14496-10, ISO/IEC 60559 | IEEE 754 and the following apply.

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

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

© ISO/IEC 2026 – All rights reserved

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

3.1
inner decoder
module conforming to a video compression standard which provides functionality to convert a compressed bitstream into a sequence or one or more pictures

3.2
inner decoded picture
picture obtained from the inner decoder

3.3
tensor
multidimensional array of scalar values grouping related model parameters

[SOURCE: ISO/IEC 15938-17:2024]

3.4
machine
fit purpose process performed by an artificial intelligence algorithm

3.5
machine consumption
applying a machine analysis task such as object detection, segmentation, or object tracking

3.6
artificial intelligence
AI
<discipline> research and development of mechanisms and applications of AI systems

[SOURCE: ISO/IEC 22989:2022, 3.1.3, modified — Note 1 to entry has been removed.]

3.7
artificial intelligence system
AI system
engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives

[SOURCE: ISO/IEC 22989:2022, 3.1.4, modified — Note 1 to entry has been removed.]

3.8
region of interest
specified boundary as defined in the image

[SOURCE: ISO/IEC 3532-2:2024]

3.9
spatial restoration
spatial dilation of a rectangular region of interest capable of restoring a corresponding spatial contraction

3.10
chroma synthesis
synthesis of chroma colour planes for an image from a luma colour plane of the image

3.11
synthesis
<artificial intelligence> generation, by a functional unit, of artificial voice, text, music, and images.

[SOURCE: ISO/IEC 2382:2015, notes removed]

© ISO/IEC 2026 – All rights reserved