INTERNATIONAL STANDARD

ISO/IEC 23003-2

Second edition
2018-12

Information technology — MPEG audio technologies —

Part 2: Spatial Audio Object Coding (SAOC)

Technologies de l'information — Technologies audio MPEG —

Partie 2: Codage d'objet audio spatial (SAOC)

Reference number
ISO/IEC 23003-2:2018(E)

© ISO/IEC 2018



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2018

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
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword v

Introduction vi

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Notations and abbreviated terms 3
4.1 Notation 3
4.2 Operations 3
4.3 Constants 3
4.4 Variables 3
4.5 Abbreviated terms 6

5 SAOC overview 7
5.1 General 7
5.2 Basic structure of the SAOC transcoder/decoder 8
5.3 Tools and functionality 10
5.4 Delay and synchronization 11
5.5 SAOC Profiles and levels 17

6 Syntax 20
6.1 Payloads for SAOC 20
6.2 Definition 35

7 SAOC processing 43
7.1 Compressed data stream decoding and dequantization of SAOC data 43
7.2 Compressed data stream encoding and quantization of MPS data 46
7.3 Time/frequency transforms 47
7.4 Signals and parameters 47
7.5 SAOC transcoding/decoding modes for baseline and LD profiles 51
7.6 EAO processing for baseline and LD profiles 64
7.7 SAOC-DE profile decoding modes 73
7.8 DCU processing 75
7.9 Modification range control for SAOC-DE processing modes 79
7.10 MBO processing 80
7.11 MCU Combiner 81
7.12 Effects 83
7.13 Low power SAOC processing 86
7.14 Low delay SAOC processing 87

8 Transport of SAOC side information 89
8.1 Overview 89
8.2 Transport and signalling in an MPEG environment 89
8.3 Transport of SAOC data over PCM channels 93

9 Transport of predefined rendering information 94
9.1 General 94
9.2 Rendering information description file format 95

10 Conformance testing 96
10.1 General 96
10.2 Terms and definitions 96
10.3 SAOC conformance testing 96
10.4 Bitstreams 96

© ISO/IEC 2018 – All rights reserved

10.5 SAOC decoder/transcoder **105**

11 Reference software **119**

11.1 Reference software structure **119**

Annex A (normative) **Tables** **121**

Annex B (normative) **Low delay MPEG surround** **150**

Annex C (informative) **Effects processing** **161**

Annex D (informative) **Encoder** **163**

Annex E (informative) **Guidelines for rendering matrix specification** **167**

Annex F (informative) **MCU combiner** **169**

Annex G (informative) **Reference software** **171**

 © ISO/IEC 2018 – 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 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) or the IEC list of patent declarations received (see http://patents.iec.ch).

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.

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 second edition cancels and replaces the first edition (ISO/IEC 23003-2:2010), which has been technically revised. It also incorporates the Amendments ISO/IEC 23003-2:2010/Amd 1:2015, ISO/IEC 23003-2:2010/Amd 2:2015, ISO/IEC 23003-2:2010/Amd 3:2015, ISO/IEC 23003-2:2010/Amd 4:2016 and ISO/IEC 23003-2:2010/Amd 5:2016 and the Technical Corrigenda ISO/IEC 23003-2:2010/Cor 1:2012 and ISO/IEC 23003-2:2010/Cor 2:2014.

The main changes compared to the previous edition are as follows:

— clarifications on SAOC-DE profile description;

— corrections to SAOC-DE profile specification;

— corrections to SAOC-DE profile;

— corrections to MPEG SAOC IS text;

— corrections to the low power mode.

A list of all parts in the ISO/IEC 23003 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.

© ISO/IEC 2018 – All rights reserved

Introduction

In the preferred modes of operating, the SAOC system, the transmitted signal can be either mono, stereo or 3-channel. The audio objects can be represented by a mono, stereo, or 3-channel signal or have the MPEG surround (MPS) multi-channel background object (MBO) format. The additional parametric data exhibits a significantly lower data rate than required for transmitting all objects individually, making the coding very efficient. At the same time, this ensures compatibility of the transmitted signal with legacy devices.

When a multi-channel rendering setup (e.g. a 5.1 loudspeaker setup) is required, the SAOC system acts as a transcoder, converting the additional parametric data to MPS parameters, and interfaces to the MPS decoder that acts as rendering device. For certain rendering setups (e.g. a binaural or plain stereo setup), the SAOC system behaves as a decoder, using its own rendering engine. Another key feature is that the SAOC parametric data from different streams can be merged at parameter level to allow for the combination of SAOC streams, similar to the functionality of a multi-point control unit (MCU).

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right. The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ISO and IEC. Information may be obtained from:

Qualcomm Incorporated
6455 Lusk Blvd
US-San Diego, CA 92121-2779

Fraunhofer Institute for Integrated Circuits IIS
Leonrodstrasse 68
DE-80636 München

LG Electronics
16 Woomyeon-Dong Seocho-Gu
KR-Seoul 137-724

Koninklijke Philips Electronics N.V.
High Tech Campus 44
NL-5656 AE , Eindhoven

Electronics and Telecommunications Research Institute
161 Gajeong-dong Yuseong-gu
KR-Daejeon 305-350

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those identified above. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

© ISO/IEC 2018 – All rights reserved

Information technology — MPEG audio technologies — Part 2: Spatial Audio Object Coding (SAOC)

1 Scope

This document specifies the reference model of the spatial audio object coding (SAOC) technology that is capable of recreating, modifying and rendering a number of audio objects based on a smaller number of transmitted channels and additional parametric data.

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 23003-1:2007, *Information technology — MPEG audio technologies — Part 1: MPEG Surround*

3 Terms and definitions

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

3.1
audio object
input audio signal consisting of one, two or multiple channels, including multi-channel background object (MBO)

3.2
frame
time segment (3.15) to which SAOC processing is applied according to the data conveyed in the corresponding SAOCFrame() or SAOCDEFrame() syntax elements

3.3
hybrid filterbank
structure, consisting of a quadrature mirror filter (QMF) bank and oddly modulated Nyquist filter banks, used to transform time domain signals into *hybrid subband* (3.5) samples

3.4
hybrid filtering
filtering step on a quadrature mirror filter (QMF) subband signal resulting in multiple *hybrid subbands* (3.5)

Note 1 to entry: The resulting hybrid subbands can be non-consecutive in frequency.

3.5
hybrid subband
subband obtained after *hybrid filtering* (3.4) of a quadrature mirror filter (QMF) subband

Note 1 to entry: The hybrid subband can have the same time/frequency resolution as a QMF subband.

© ISO/IEC 2018 – All rights reserve

3.6
input channel
input audio channel corresponding to the channels of an *audio object* (3.1)

3.7
output channel
audio channel corresponding to a specific speaker

Note 1 to entry: Channel abbreviations and loudspeaker positions are given in Table 1.

3.8
parameter band
one or more *hybrid subbands* (3.5) applicable to one parameter

3.9
parameter time slot
specific *time slot* (3.16) for which the parameter is defined

3.10
parameter set
parameters associated with a specific *parameter time slot* (3.9)

3.11
parameter subset
parameters associated with a specific *parameter time slot* (3.9) and a specific one-to-two (OTT) box or two-to-three (TTT) box

3.12
processing band
one or more *hybrid subbands* (3.5) defining the finest frequency resolution that could be controlled by the parameters

3.13
QMF bank
bank of complex exponentially modulated filters

3.14
QMF subband
subband obtained after QMF filtering of a time-domain signal, without any additional hybrid filtering stage

3.15
time segment
group of consecutive *time slots* (3.16)

3.16
time slot
finest resolution in time for spatial *audio object* (3.1) coding (SAOC) time borders

Note 1 to entry: One time slot equals one subsample in the hybrid quadrature mirror filter (QMF) domain.

© ISO/IEC 2018 – All rights reserved