ETSI TS 103 448 V1.1.1 (2016-09)

AC-4 Object Audio Renderer for Consumer Use

EBU
OPERATING EUROVISION

Reference
DTS/JTC-038

Keywords
audio, broadcasting, digital

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status.
Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.
The content of the PDF version shall not be modified without the written authorization of ETSI.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2016.
© European Broadcasting Union 2016.
All rights reserved.

DECT™, **PLUGTESTS**™, **UMTS**™ and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ and **LTE**™ are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Contents

Intellectual Property Rights 5
Foreword 5
Modal verbs terminology 5
Introduction 5
1 Scope 7
2 References 7
2.1 Normative References 7
2.2 Informative References 7
3 Definitions, symbols, abbreviations and conventions 7
3.1 Definitions 7
3.2 Symbols 8
3.3 Abbreviations 8
3.4 Conventions 9
4 System overview 10
4.1 Architecture 10
4.1.1 Introduction 10
4.1.2 Requirements 10
4.2 Input 11
4.2.1 Audio 11
4.2.2 Metadata 11
4.2.3 Playback parameters 13
4.3 Output 13
5 Functional overview 13
5.1 Metadata pre-processing 13
5.1.1 Processing order 13
5.1.1.1 Introduction 13
5.1.1.2 Requirements 14
5.1.2 Screen-anchored coordinates 14
5.1.2.1 Introduction 14
5.1.2.2 Recommendations 15
5.1.2.3 Algorithmic details 16
5.1.3 Zone constraints 16
5.1.3.1 Introduction 16
5.1.3.2 Requirements 16
5.1.4 Snap 18
5.1.4.1 Introduction 18
5.1.4.2 Requirements 18
5.1.4.3 Algorithmic details 19
5.1.5 Metadata for future use cases 19
5.2 Source panner 19
5.2.1 Source panner architecture 19
5.2.1.1 Introduction 19
5.2.1.2 Requirements 20
5.2.2 Rendering point sources 20
5.2.2.1 Introduction 20
5.2.2.2 Requirements 21
5.2.2.3 Algorithmic details 21
5.2.3 Rendering objects with size 23
5.2.3.1 Introduction 23
5.2.3.2 Requirements 23
5.2.3.3 Algorithmic details 25
5.2.4 Rendering objects with divergence 28
5.2.4.1 Introduction 28

5.2.4.2 Requirements 28
5.2.4.3 Algorithmic details 29
5.2.5 Rendering objects with speaker-anchored coordinates 29
5.2.5.1 Introduction 29
5.2.5.2 Requirements 29
5.3 Trim 30
5.3.1 Introduction 30
5.3.2 Requirements 30
5.3.3 Algorithmic details 31
5.4 Gain mixer 32
5.4.1 Introduction 32
5.4.2 Requirements 32
5.4.3 Algorithmic details 32
5.5 Ramp mixer 32
5.5.1 Introduction 32
5.5.2 Requirements 32
5.5.3 Algorithmic details 32

Annex A (normative): Loudspeaker configurations and source panner coordinates 34

A.1 Introduction 34

A.2 Requirements 34

A.3 Tables 34

History 39

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by Joint Technical Committee (JTC) Broadcast of the European Broadcasting Union (EBU), Comité Européen de Normalisation ELECtrotechnique (CENELEC) and the European Telecommunications Standards Institute (ETSI).

NOTE: The EBU/ETSI JTC Broadcast was established in 1990 to co-ordinate the drafting of standards in the specific field of broadcasting and related fields. Since 1995 the JTC Broadcast became a tripartite body by including in the Memorandum of Understanding also CENELEC, which is responsible for the standardization of radio and television receivers. The EBU is a professional association of broadcasting organizations whose work includes the co-ordination of its members' activities in the technical, legal, programme-making and programme-exchange domains. The EBU has active members in about 60 countries in the European broadcasting area; its headquarters is in Geneva.

European Broadcasting Union
CH-1218 GRAND SACONNEX (Geneva)
Switzerland
Tel: +41 22 717 21 11
Fax: +41 22 717 24 81

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

Motivation

Current industry trends for authoring and reproduction of audio content include immersive audio and support for personalization of the audio, as well as many different speaker setups and layouts.

Different means of immersive audio and personalization are provided in ETSI TS 103 190-2 [1]. Object-based audio is one of the means for supporting these trends.

Objects can be thought of as the input tracks to a mixing console, the mixing console being the renderer. But objects are more than audio tracks. They carry metadata that is authored with the tracks. Contemporary mixing consoles have automated gains. For a renderer accepting object-based audio, those gains are driven by the object's own metadata. Metadata is also used to define object location and size, as well as many other ancillary parameters that control the object presentation.

The final mix output by the contemporary mixing console is targeted at a specific playback system. Other channel configurations can be derived from the mix, but they are not necessarily what is monitored. A renderer, located in a playback device in a consumer's home, acts as the mixing console for that device, with the advantage that the speaker setup is known to the renderer. The renderer can use the location and size metadata defined for each object to produce the playback experience that best matches the content creator's intention, within the possibilities and constraints of the available speaker setup.

The present document specifies an object audio renderer for use with ETSI TS 103 190-2 [1], using the metadata as specified therein.

Structure of the document

The present document is structured as follows.

- Clause 4 specifies the input and output interfaces, and the architecture of the renderer.
- Clause 5 specifies the processing blocks of the renderer. These are:
  - Metadata preprocessing, specified in clause 5.1
  - Source panners, specified in clause 5.2
  - Trim processing, specified in clause 5.3
  - Gain mixing, specified in clause 5.4
  - Ramp mixing, specified in clause 5.5
- Annex A lists the supported loudspeaker configurations and associated parameters that are utilized by the processing blocks of the renderer.

An overview of the incoming metadata and the result of the rendering process is presented in clause 4, which makes it a proper starting point when reading the document.

1 Scope

The present document defines an extension to the AC-4 codec.

The present document specifies a consumer object-based audio renderer for use with the AC-4 codec as specified in ETSI TS 103 190-2 [1], and the object-based audio metadata specified therein. The renderer takes the object audio essence and the corresponding metadata defined in ETSI TS 103 190-2 [1] as inputs, and produces loudspeaker feeds for consumer loudspeaker layouts.

2 References

2.1 Normative References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

Referenced documents that are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: Although any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] ETSI TS 103 190-2: "Digital Audio Compression (AC-4) Standard; Part 2: Immersive and personalized audio".

[2] Recommendation ITU-R BS.2051-0: "Advanced sound system for programme production".

2.2 Informative References

References are either specific (identified by date of publication and/or edition number or version number) or nonspecific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document, but they assist the user with regard to a particular subject area.

Not applicable.

3 Definitions, symbols, abbreviations and conventions

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

bitstream: sequence of bits

channel: audio signal intended for playback by one of a set of dedicated loudspeakers with predetermined locations, e.g. Left, Right, and Centre channels

codec: system that consists of an encoder and a decoder