ETSI TS 103 584 V1.1.1 (2018-01)

DTS-UHD Point Source Renderer

EBU
OPERATING EUROVISION

Reference

DTS/JTC-DTS-PSR

Keywords

audio, broadcast, codec, digital, object audio

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

Contents

List of Tables

List of Figures

Intellectual Property Rights

Essential patents

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.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

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.

1 Scope

The present document defines an audio renderer associated with the DTS-UHD codec defined in ETSI TS 103 491 [1]. The inputs to the renderer are one or more sets of audio waveforms along with mixing instructions, and the output is a single set of waveforms mapped to a defined speaker configuration. Each set of waveforms may represent either audio channels or audio objects. The mixing instructions may vary with time and they come from metadata carried in the DTS-UHD bitstream and from other application interfaces described herein.

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 referenced document (including any amendments) applies.

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

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 necessary for the application of the present document.

[1] ETSI TS 103 491: "DTS-UHD Audio Format; Delivery of Channels, Objects and Ambisonic Sound Fields", version 1.1.1.

[2] Pulkki, Ville: "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", JAES Volume 45 Issue 6 pp. 456-466; June 1997.

2.2 Informative 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 referenced 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 and abbreviations

3.1 Definitions

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

audio object: waveform or set of waveforms that have been assigned a location in space

hull: collection of points forming a surface mesh to be used as a basis for panning

output feeds: set of audio waveforms mapping one to one to the physical speakers

panning: distributing a point source location into a speaker layout

physical speaker: speaker that physically exists in the listening space

point: location on the surface of the unit sphere

point source: audio waveform with defined spatial coordinates within the listening room

NOTE: A point source is treated (mathematically) as a uniform spherical radiation pattern, though in reality most speakers have a pattern of less than 180 degrees on the face of the speaker. This reality is accounted for by limiting the panner range.

render: combine waveforms with their gain contributions to produce audio waveforms in a specified speaker configuration

speaker: transducer that converts an electrical signal into soundwaves

speaker configuration: defined list of physical speakers in the listening space

triplet: set of three points

virtual speaker: speaker that does not physically exist in the listening space

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

∪	Union operator
3D	three dimensional
C	Centre speaker
Ch	Centre high speaker
Chr	Centre high rear speaker
Clf	Centre low front speaker
Cs	Centre surround speaker
dB	decibel
EMA	Exponential Moving Average
L	Left speaker
Lc	Left centre speaker
LFE	Low Frequency Effects
LFE1	Low Frequency Effects - 1 speaker
LFE2	Low Frequency Effects - 2 speaker
Lh	Left high speaker
Lhr	Left high rear speaker
Lhs	Left high side speaker
Llf	Left low front speaker
LR	Left Right
Ls	Left surround speaker
Lsr	Left surround rear speaker
Lss	Left side surround speaker
Ltf	Left top front speaker
Ltr	Left top rear speaker
Lw	Left wide speaker
N/A	Not Applicable
Oh	Overhead speaker
R	Right speaker
Rc	Right centre speaker
Rh	Right high speaker
Rhr	Right high rear speaker
Rhs	Right high side speaker
Rlf	Right low front speaker
Rs	Right surround speaker
Rsr	Right surround rear speaker
Rss	Right side surround speaker

Rtf	Right top front speaker
Rtr	Right top rear speaker
Rw	Right wide speaker

4 Renderer Metadata

4.1 Overview of the Renderer Inputs

The DTS-UHD point source object renderer is a system that computes the waveform associated with a specified speaker configuration, given a collection of audio objects as input, as shown in Figure 1. The renderer metadata assumes the model of a sphere with the listener's head at the centre (an ego-centric model). All speakers are treated as point sources and are assumed to be on the surface of a unit sphere, and thus have a radius of 1.

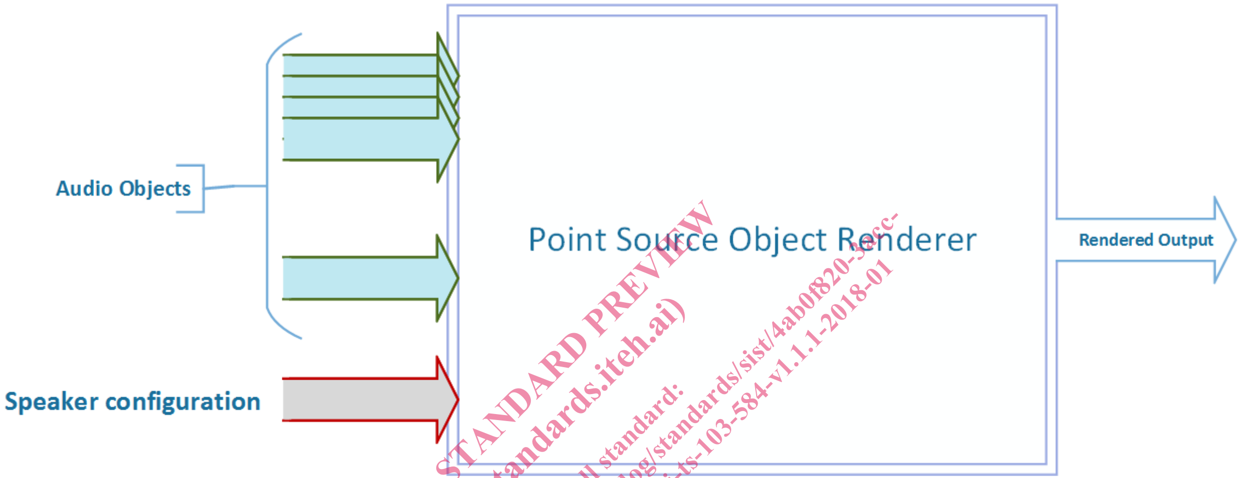

Figure 1: Point Source Object Renderer System

DTS-UHD audio objects consist of metadata plus one or more waveforms. The metadata always includes spatial location information, and may include additional parameters, such as gain information.

Audio objects can be manipulated by changing the gain and position metadata. For example, when the renderer metadata is coming from a DTS-UHD bitstream, the metadata can be updated as often as every 512 clock periods, where the clock shall be defined by `m_unClockRateInHz` in ETSI TS 103 491 [1]. If the audio within the DTS-UHD stream is carrying audio compressed using ACE, each ACE frame has a duration of 1 024 samples, therefore the mixing metadata can be updated twice per ACE frame. For audio with a sampling frequency of 48 KHz, this translates to a minimum update interval of 10,67 ms.

It is also possible for object metadata to be modified by user intervention. In this case, the bitstream metadata is being overridden until control is given back to the bitstream metadata.

The present document describes these in details to provide a comprehensive understanding of the DTS-UHD Point Source Object Renderer.

4.2 DTS-UHD Bitstream Metadata

DTS-UHD metadata is defined in ETSI TS 103 491 [1]. Since the renderer described in the present document is intended to be used in conjunction with the DTS-UHD bitstream, the DTS-UHD metadata parameters relating to the renderer are shown in Table 1.

Note that Table 1 accommodates restrictions inherent in DTS-UHD metadata, but this does not imply such restrictions are imposed on the renderer capabilities.

Table 1: DTS-UHD Metadata Parameters used for Rendering

Metadata Parameter	Reference in ETSI TS 103 491 [1]	Description	Object Renderer Parameter
m_unObjectID	7.8.7	Identifier to access and modify an object state.	`AudioObject::object_id`
m_bMonoObjWithMultipleSourcesFlag (or bObjMS for short)	7.8.11.2	Indicates whether the AudioObject with one MonoObject is associated with multiple 3D points.	used with m_ucNum3DSourcesInObj (below)
m_ucNumWaveFormsInObj	7.8.11.17	Specifies the number of waveforms (MonoObjects) in the AudioObject.	`AudioObject:: mono_objects[].size()`
m_bPerSamplPeriodObjMDUpdFlag	7.8.11.22.2	Flag indicates that there are metadata updates per processing block within a larger audio block.	N/A
m_ppObjGain	7.8.11.22.3	Specifies the AudioObject gain.	`AudioObject:: object_gain`
m_ucNum3DSourcesInObj	7.8.11.22.4	Specifies the total number of point source locations in the AudioObject. Note that DTS-UHD metadata restricts an AudioObject to have either multiple MonoObjects each with a single point source location, or a single MonoObject with multiple point source locations.	`if (bObjMS) { MonoObject:: positions[].size() AudioObject:: mono_objects[].size() = 1 } else { AudioObject:: mono_objects[].size() MonoObject:: positions[].size() = 1 }`
unsrc_index	7.8.11.22.5	Identifies the source index within the AudioObject.	N/A
m_rPerObjExpWinLambda	7.8.11.23	Specifies the smoothing factor for the AudioObject.	`AudioObject:: obj_lambda`
m_3DSrcRadius	7.8.11.28.2	Specifies radius of the point source location.	`AudioObject:: mono_objects[]. positions[].radius`
m_3DSrcAzimuth	7.8.11.28.3	Specifies azimuth of the point source location. Range is $[-180°, 180°]$.	`AudioObject:: mono_objects[]. positions[].azimuth`
m_3DSrcElevation	7.8.11.28.4	Specifies elevation of the point source location. Range is $[-90°, 90°]$.	`AudioObject:: mono_objects[]. positions[].elevation`

4.3 DTS-UHD Object Interactivity Manager

Renderer APIs may allow the object metadata from the DTS-UHD bitstream to be overridden by the user during playback. Metadata in the bitstream may also be set to limit or disable a user interaction. The object interactivity manager enforces these rules and applies any user changes to the metadata before calling the renderer. Figure 2 shows that the object interactivity manager sits just before the renderer, where it handles the user input and the limit rules specified by the bitstream creator.