STANDARDS AND INFORMATION DOCUMENTS

AES41-5-2012
reaffirmed 2022

STANDARDS

AES standard for digital audio - Audio-embedded metadata - Part 5: EBU loudness, true-peak, and downmix

AUDIO ENGINEERING SOCIETY, INC.
132 East 43rd St., Suite 405, New York, NY 10017, US. www.aes.org/standards

STANDARDS

The AES Standards Committee is the organization responsible for the standards program of the Audio Engineering Society. It publishes technical standards, information documents and technical reports. Working groups and task groups with a fully international membership are engaged in writing standards covering fields that include topics of specific relevance to professional audio. Membership of any AES standards working group is open to all individuals who are materially and directly affected by the documents that may be issued under the scope of that working group.

Complete information, including working group scopes and project status is available at http://www.aes.org/standards. Enquiries may be addressed to standards@aes.org

The AES Standards Committee is supported in part by those listed below who, as Standards Sustainers, make significant financial contribution to its operation.

Standards Sustainer Gold

NETFLIX

Focusrite

Standards Sustainer Silver

LECTROSONICS

21ST CENTURY FOX

Audio precision

WAVES

ADAMSON

fulcrum ACOUSTIC

DOLBY

Eventide

PreSonus

COVELOZ

SENNHEISER

THE TELOS ALLIANCE

MUSIC TRIBE

BOSE

Standards Sustainer

AUDYSSEY

NTi AUDIO

MQA

audio-technica

LAWO

CLAIR

WEISS

METRIC HALO

HEA

U SOUND

THAT Corporation

NBCUniversal

Control 4

WorldCast Systems

Fraunhofer IIS

PRO AUDIO TECHNOLOGY

MERGING

AUDIO FOR THE NETWORKING AGE

RAVENNA

steinberg

ARCHWAVE

audio networking & streaming technologies

ATTERO TECH

HARMAN

A SAMSUNG COMPANY

KLIPPEL

NEUTRIK

dbx-tv

XPERI

This list is current as of 2018/9/01

AES41-5-2012
reaffirmed 2017

AES standard for digital audio - Audio-embedded metadata - Part 5: EBU loudness, true-peak, and downmix

Published by
Audio Engineering Society, Inc.

Abstract
AES41 provides for the carriage of audio metadata by embedding it in the audio samples themselves. This tightly associates the meta-data with the audio, yet makes it fragile so that changes to the audio will invalidate the meta-data. Several metadata sets have been defined, covering applications such as cascaded compression (bit rate reduction), and loudness control.

This part describes the format for the data to be transmitted with audio to signal loudness and true peak meta-data as used in EBU Technical Recommendation R128. A method of carrying this data is described in Part 1 of this Standard.

2017-12-31 printing

Contents

Foreword

This foreword is not part of AES41-5-2012 *AES standard for digital audio - Audio-embedded metadata - Part 5: EBU loudness, true-peak, and downmix*

This document describes a set of data that may be conveyed according to the method described in Part 1 of this Standard.

As predicted in the foreword to AES41-2000, digital compression techniques now dominate the broadcast television environment. In addition to the problems foreseen relating to cascaded compression, new problems have arisen because of the use of loudness control and surround sound with those digital compression techniques.

Metadata within the compressed audio bit-stream is used to control loudness and the mixing down of multi-channel surround sound to two-channel stereo. These metadata are usually known by terms such as "dialnorm", "prog_ref_level", and "down-mix coefficients".

Whilst this might seem unrelated to the original scope of AES41, dealing with bit allocations and scale factors, it is simply another form of data that can affect a later encoding of the audio: this time is it is more macroscopic than microscopic.

The metadata is lost when the bit-stream is uncompressed unless provision is made to transport it or store it somewhere. Existing methods rely on non-audio mechanisms to convey the metadata alongside the audio, for example a serial data link like RS-422 and serial digital video SMPTE 259M, or a "chunk" in an audio file (for metadata that does not change).

This part extends AES41 to include a data format for carrying loudness metadata as defined by the EBU following the work of the "PLOUD" group, as well as downmix metadata, with the uncompressed PCM using the same transport mechanism as before. The metadata can therefore be carried in the audio to which it relates.

The draft of this document was developed by a writing group whose primary author was Andrew Mason.

John Grant
Chair, working group SC-02-02
2012-03

Note on normative language

In AES standards documents, sentences containing the word “shall” are requirements for compliance with the document. Sentences containing the verb “should” are strong suggestions (recommendations). Sentences giving permission use the verb “may”. Sentences expressing a possibility use the verb “can”.

AES standard for digital audio - Audio-embedded metadata - Part 5: EBU loudness, true-peak, and downmix

0 Introduction

AES41 provides for the carriage of audio meta-data by embedding it in the audio samples themselves. This tightly associates the meta-data with the audio, yet makes it fragile so that changes to the audio will invalidate the meta-data. Several meta-data sets have been defined, covering applications such as cascaded compression (bit rate reduction), and loudness control.

This part describes the format for the data to be transmitted with audio to signal loudness and true peak meta-data as used in EBU Technical Recommendation R128. This part should be read in conjunction with part 1.

0.1 Rationale for part 5 of this standard

Unwanted loudness variations in broadcast audio have historically been the source of many audience complaints. Broadcasters have adopted numerous techniques to address the problem - one of which involves indicating the long-term average loudness of the audio.

The EBU studied the needs of audio signal levels in production, distribution and transmission of broadcast programmes. It formed the opinion that an audio-leveling paradigm was needed based on loudness measurement according to ITU-R BS.1770. This is described in EBU Technical Recommendation R 128, and supporting Technical Documents.

When the PCM signals form part of a 5-channel surround sound signal it may be necessary to include down-mix coefficients to allow later mixing of the center and surround channels into the front left and front right to form a two-channel stereo signal.

1 Scope

This document describes a format for the data to be transmitted to convey loudness and true peak metadata as defined by the EBU following the work of the "PLOUD" group. This Part assumes that the transmission mechanism according to Part 1 of this Standard is used.

2 Normative references

The following referenced documents are indispensable for the application 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.

AES41-1-2012, *AES standard for digital audio - Audio-embedded metadata - Part 1: General*, Audio Engineering Society, New York, NY., US.