Standard test materials for three-dimensional multichannel stereophonic sound system – Series B

Manual

On the Publication of Standard test materials for three-dimensional multichannel stereophonic sound system – Series B

The Institute of Image Information and Television Engineers (ITE) has continuously worked to promote the advancement, development, and dissemination of science and technology related to image information media, with the aim of contributing to the progress of such media in Japan. As part of these activities, the Institute has published various standard images, including standard digital images and standard images for high-definition and stereoscopic video, which have played a significant role in the research and development of imaging systems and new services, as well as in maintaining their quality.

In recent years, research and development toward 4K and 8K broadcasting have progressed as media capable of delivering a strong sense of presence through highly detailed images and three-dimensional audio, and practical broadcasting began in 2018. In the field of video, the Institute, in collaboration with the Association of Radio Industries and Businesses (ARIB), produced and published ultra-high-definition, wide color gamut standard still and moving images compliant with ITU-R Recommendation BT.2020, which defines the international standard format for 4K and 8K video. In the field of audio, the Institute also produced and published the Standard test materials for three-dimensional multichannel stereophonic sound system A-Series, a standard reference audio set for 22.2 multichannel sound compliant with ITU-R Recommendation BS.2051 and ARIB STD-B59. These standard images and audio materials have been well received and widely utilized for the development and evaluation of various equipment.

More recently, object-based audio has attracted attention as a technology enabling personalized services by allowing broadcast programs to be customized according to listener preferences and viewing environments. In next-generation broadcasting systems in Europe, the United States, and Brazil, audio coding systems supporting object-based audio have been adopted. In Japan, an audio coding method supporting object-based audio was added in March 2025 to the standard ARIB STD-B32. Furthermore, guidelines for audio metadata, which are essential for content production, were issued as ARIB Technical Report TR-B48. However, there has been no standard test material compliant with object-based audio systems and metadata guidelines for verifying these specifications, and requests for such materials have been received from research institutions, broadcasters, manufacturers, and companies both in Japan and overseas.

In response to this background, the Institute and ARIB have collaborated on the development of standard test materials for three-dimensional multichannel stereophonic sound system compatible with object-based audio. As a result of these efforts, we are pleased to publish the second set of standard test materials for object-based audio, the Standard test materials for three-dimensional multichannel stereophonic sound system – Series B. We would like to express our

sincere appreciation for the tremendous contributions and efforts of all those involved. We strongly hope that this publication will further stimulate research and development in next-generation image information media and promote the smooth and widespread dissemination of stable broadcasting services that offer immersion and dynamism.

July 2026

Kenji Terada, President

The Institute of Image Information and Television Engineers

On the Selection and Production of the Standard test materials for three-dimensional multichannel stereophonic sound system – Series B

The Association of Radio Industries and Businesses (ARIB), R&D Group on Program Production Systems, has been promoting research, development, and standardization of systems and quality evaluation methods related to program production, editing, and signal transmission within broadcasting stations (excluding contribution links).

To date, the Committee has undertaken, at the request of the Institute of Image Information and Television Engineers (ITE), the selection and production of standard moving and still images distributed by the Institute, including those for high-definition television and ultra-high-definition, wide color gamut content (including HDR versions). These materials have been incorporated into ITU-R Report BT.2245, “HDTV and UHDTV including HDR-TV test materials for assessment of picture quality,” and are now widely utilized both domestically and internationally.

In the field of audio, the Committee conducted the selection and production of the Standard test materials for three-dimensional multichannel stereophonic sound system A-Series in 2017, assuming 4K and 8K broadcasting services. These materials were included in ITU-R Report BS.2494, “Sound test materials for advanced sound systems,” and have since been widely used worldwide. Subsequently, research and development in object-based audio has accelerated both in Japan and internationally, leading to the establishment of various international recommendations, including the Audio Definition Model (ADM) for defining common audio-related metadata for program production, renderers, a serial representation of ADM (S-ADM), file formats for program exchange, and profiles for broadcasting and other services. In response to these developments, ARIB has also established the technical report ARIB TR-B48, “Operational Guidelines for Audio Metadata Used in Object-Based Audio,” thereby supporting production and implementation environments.

In light of these global and domestic trends, as well as the increasing demand for object-based audio content, ARIB has newly selected and produced the “Standard test materials for three-dimensional multichannel stereophonic sound system - Series B” under commission from ITE. This series is designed to comprehensively cover audio-related metadata required for anticipated future services, such as multilingual broadcasting with narration, while utilizing the A-Series as background audio. It complies not only with ARIB TR-B48 but also with relevant ITU-R recommendations, including broadcast emission profiles, and is expected to be applicable to a wide range of use cases beyond broadcasting.

On the release of this standard test material, we would like to express our sincere appreciation to the Institute of Image Information and Television Engineers for its supervision and distribution.

We hope that this standard test material, together with conventional standard images, will be widely utilized and will contribute to the further development of image and audio media.

July 2026

Tsukuru Kai, Chairman

R&D Group on Program Production Systems

Association of Radio Industries and Businesses

※ The copyrights for these standard test materials, this manual, and appendices belong to the Institute of Image Information and Television Engineers and the Association of Radio Industries and Businesses.

※ Unauthorized copying of these standard test materials, this manual and appendices is a violation of copyright and is strictly prohibited. However, appearance in academic publications and scientific journals will not be contested.

※ Use of these standard test materials is limited to the following purposes:

(1) Technical evaluation
- Device and system R&D
- Testing and inspection in device manufacturing processes
- Transmission path evaluation for broadcasting and telecommunications
- Equipment maintenance

(2) Exhibition use
- Presentation/Exhibition at Academic/Research conferences
- Exhibition of equipment performance and functionality at trade shows[1]
  (Use in promotion of sales is excluded)

[1] Please contact the Institute of Image Information and Television Engineers before using these materials for this purpose (Exhibiting equipment at a trade show).

The following people and organizations were involved in production of these standard test materials[2].

The Institute of Image Information and Television Engineers, Test Chart Committee

Committee Chair:	Yuichi Kusakabe (NHK)
Secretary:	Satoshi Morimoto (Fuji Television)
Members:	Toshiharu Horiuchi (KDDI Research), Toru Yokoyama (Hitachi), Kouichi Hamada (Teikyo University)

Institute of Image Information and Television Engineers, 3D Multichannel Standard Test Materials Subcommittee

Chief Investigator:	Yasushige Nakayama (NHK)
*Ex-chief Investigator:	Kunimoto Hamanaka (NHK)
Members:	Takashi Taniguchi (Osaka Gakuin University), Hiroshi Furukawa (Sony)
Executive Office:	Shinji Tamanoi
Ex-executive Office:	Yukio Iwahana

Association of Radio Industries and Businesses, R&D Group on Program Production Systems, Sound Quality Evaluation WG

Chief Investigator:	Satoshi Oode (NHK)
Members:	Hiroyasu Miyashita (NHK), Yasuhiro Nishimura (NHK), Noriyuki Koga (NHK), Hiroki Kubo (NHK), Taishi IWASAKI (NHK), Masaki Nakamura (TBS Television), Kenichi Shikamata (Fuji Television), Shoma Tsuboi (Nippon TV), Seiji Tanaka (Mainichi Broadcasting System), Yutaka Kamamoto (NTT), Satoshi Teshigawara (Dolby Japan), Taro Hagiya (Dolby Japan)
*Ex-members	Osamu Nakagawara (NHK), Tsunenori Nohara (NHK), Shogo Takahashi (NHK), Yuki Chubachi (NHK), Akinori Yamamura (NHK), Yoko Hirose (TV Asahi), Risa Sugimura (Nippon TV), Yuichiro Takamizawa (Dolby Japan)

[2] The affiliations listed are those at the time of joining each committee.

Guest members:	Toru Kamekawa (Tokyo University of the Arts), Atsushi Marui (Tokyo University of the Arts)
Executive Office:	Ryosuke Amano, Yuko Yamanouchi
Ex-Executive Office:	Naoki Okano, Toshiaki Otsuka, Hideki Kouno, Ryoichi Nakai, Hideho Yamada

Table of Content

1. Overview of the Standard Test Materials for Three-Dimensional Multichannel Stereophonic Sound System – Series B

In the past, the Institute of Image Information and Television Engineers (ITE) has published various standard test materials to facilitate accurate evaluation of performance and image quality in video equipment, systems and processing methods. For the 4K and 8K television broadcasting launched in December 2018, ITE also published a series of standard test materials. These include Ultra-high definition/wide-color-gamut standard test sequences - Series A[3] (December 2015), Ultra-high definition/wide-color-gamut standard test sequences - Series B[4] (November 2017) and Ultra-high definition/wide-color-gamut HDR standard test sequences[5] (March 2019). In addition, for the three-dimensional multichannel sound formats used in 4K and 8K broadcasting, ITE published Standard test materials for three-dimensional multichannel stereophonic sound systems Series A[6] in July 2017. The purpose of these materials was to promote research, development and deployment of 22.2 multichannel sound systems.

In recent years, research and development on advanced terrestrial broadcasting has progressed, and metadata-guided audio (MGA) has attracted increasing attention as its candidate audio systems. In Japan, ARIB STD-B32[7] was revised in March 2025 to incorporate audio coding schemes compatible with MGA. However, there have been no standard test materials for the MGA. We have received many requests for such sound test materials from research facilities, broadcasters, and manufacturers both within and outside Japan. To address this demand and promote research, development and deployment of MGA, we have published Standard test materials for three-dimensional multichannel stereophonic sound system - Series B.

MGA delivers personalized audio experiences based on the listener's preferences and listening environment. This is enabled by the rendering processes that convert a set of

---

[3] Ultra-high definition/wide-color-gamut standard test sequences - Series A Manual, Institute of Image Information and Television Engineers (ITE), 2015.
[4] Ultra-high definition/wide-color-gamut standard test sequences - Series B Manual, Institute of Image Information and Television Engineers (ITE), 2017.
[5] Ultra-high definition/wide-color-gamut HDR standard test sequences Manual, Institute of Image Information and Television Engineers (ITE), 2019.
[6] Standard test materials for three-dimensional multichannel stereophonic sound systems Series A Manual, Institute of Image Information and Television Engineers (ITE), 2017.
[7] ARIB STD-B32, VIDEO CODING, AUDIO CODING AND MULTIPLEXING SPECIFICATIONSFOR DIGITAL BROADCASTING, Association of Radio Industries and Businesses (ARIB), 2025.

audio signals with audio-related metadata[8] into a different set of audio signals fed to loudspeakers. The transmitted audio signals contain one or more audio objects, such as dialog, background ambience, and sound effects. As an international standard of audio-related metadata, the Audio Definition Model (ADM) is specified in International Telecommunication Union-Radiocommunication (ITU-R) Recommendation (Rec.) ITU-R BS.2076[9]. In practical broadcasting cases, audio–related metadata is expected to conform to a subset of ADM, emission profile. This profile is specified in Rec. ITU-R BS.2168[10] or, domestically in Japan, Technical Report ARIB TR-B48[11]. These test materials have been produced to provide, at an early stage, various sets of test materials and audio-related metadata that are expected to be typically used for MGA services in Japan, namely compliant with ARIB TR-B48. The primary purpose of the test materials is to support testing and monitoring channel identification or objects' motions and functionalities caused by audio-related metadata. The materials were produced by the ARIB, R&D Group on Program Production Systems, Sound Quality Evaluation Working Group.

In MGA, sound sources often consist of combinations of multiple audio objects, including mono audio objects and channel-based audio objects compliant with Recommendation ITU-R BS.2051[12]. Accordingly, the test materials include newly recorded mono audio signals, such as noise and dialog, and 22.2 multichannel audio signals previously recorded for Standard test materials for three-dimensional multichannel stereophonic sound systems - Series A. In these test materials, two file formats are provided WAV and BW64[13] to accommodate a wide range of rendering devices and software. Each file contains multi-track audio signals (one or more tracks).

---

[8] Metadata describing rendering conditions of audio objects, such as position and gain. The Audio Definition Model (ADM), a type of audio-related metadata, is represented in XML (eXtensible Markup Language) format.

[9] Recommendation ITU-R BS.2076-3, Audio definition model, 2025.

[10] Recommendation ITU-R BS.2168-0, Audio definition model and serial representation of audio definition model profile for advanced sound systems emission, 2025.

[11] ARIB Technical Document TR-B48, OPERATIONAL GUIDELINES FOR AUDIO METADATAUSED IN OBJECT-BASED AUDIO, 2025.

[12] Recommendation ITU-R BS.2051-3, Advanced sound system for programme production, 2022.

[13] An audio-file format specified in Recommendation ITU-R BS.2088-2, "Long-form file format for the international exchange of audio programme materials with metadata," capable of storing data exceeding the 4 GB limit and audio-related metadata.

In addition to the audio signals, BW64 files contain audio-related metadata[14], and audio–track association information[15].

[14] Stored in the <axml> chunk in a BW64 file.
[15] Information used to associate audio-tracks referenced in audio-related metadata with the actual audio-track channels in the audio data. Stored in the <chna> chunk container in a BW64 file.

2. Recording Format (English Translation)

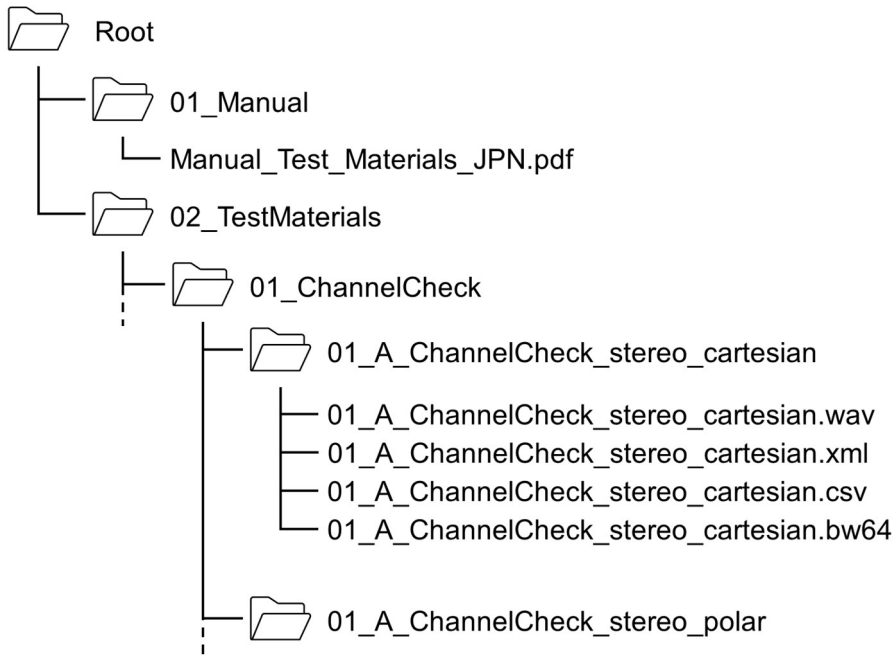
This chapter describes the directory structure, file formats, and metadata used in the Standard test materials for three-dimensional multichannel stereophonic sound system - Series B.

2.1. Folder structure

The standard test materials consist of three folders located under the root directory: the “01_Manual” folder, which contains this manual, the “02_TestMaterials” folder, which contains the test sounds and audio-related metadata, and the “03_ARIBAudioPlayer” folder, which contains software to play back rendered audio. Within the “02_TestMaterials” folder, subfolders are provided for each evaluation item. Each evaluation item is assigned to an item number from 01 to 06, together with the name of the evaluation item. The folder name is formed by combining these elements. For example, for the evaluation item called Channel Check, the corresponding folder is named “01_ChannelCheck”. Furthermore, within this folder, additional subfolders are corresponding to different sets of sounds and audio-related metadata (e.g., “01_A_ChannelCheck_stereo_cartesian”). Each of these subdivided folders contains:

- audio files (WAV format),
- audio-related metadata (XML format),
- <chna> chunks (CSV format), and
- BW64-format files integrating the above.

All file types required for evaluation are stored within these folders.



2.2. File Format

The file formats used in this set of standard test materials are as follows.

- Audio data
  - Wave format : Linear PCM
  - Sampling frequency : 48 kHz
  - Quantization bit depth : 24 bits
  - Number of tracks : 1-56 (depending on the evaluation item)
  - File formats : WAV, BW64
  - File name:
    Number_Alphabet_EvaluationItem_EvaluationCondition_CoordinateSystem
    - *Number*: ID of the evaluation item
    - *Alphabet*: ID of the evaluation condition (assigned sequentially from “A” for each evaluation item)
    - *Evaluation Item*: Name of the evaluation item
    - *Evaluation Condition*: Name of the evaluation condition
    - *Coordinate System*: Name of the coordinate system used for describing the positional information of audio objects
- Audio-related Metadata
  - Metadata format : Audio Definition Model (ADM)
  - File format : XML
- Audio track information[16]
  - File format : CSV

2.3. Audio Definition Model

The Audio Definition Model (ADM) is specified in Rec. ITU-R BS.2076. This describes various information required to render multi-track audio signals with the same number of samples. In this section, some ADM-related terms that frequently appear in this manual are introduced and then different coordinate systems that are described in section 2.2 are described in contrast with rendering algorithms.

XML elements in ADM

The primary XML elements that ADMs contain in these test materials are listed in Table 1. Each element has child-elements to refer to the other elements. These reference relationships enable specifying program composition and rendering parameters. For

[16] Correspondence between the physical audio tracks and virtual audio tracks, audio channels, and audio packs of the audio signals in ADM. This information enables mapping between each track in the audio file and the corresponding ADM description.

details on the attributes and child-elements of each element, please refer to Rec. ITU-R BS.2076, Technical Report ARIB TR-B48, and other relevant documents.

Table 1 Elements mainly used in ADM

Element[17]	Description
audioProgramme (audio program)	It defines the program composition by combining audio content.
audioContent (audio content)	It defines semantic units that constitute an audio program, such as background ambience and dialog.
audioObject (audio object)	It defines the processing units in the renderer by referencing audio packs (channel configurations) and audio track UIDs (physical tracks). These processing units include channel-based audio objects, such as stereo or 5.1, and mono audio objects that can be panned based on positional information. In the ADM, these are distinguished using the attributes typeDefinition and type. For channel-based audio objects, typeDefinition = DirectSpeakers and type = 0001 are specified, while for mono audio objects, typeDefinition = Objects and type = 0003 are specified.
audioPackFormat (audio pack)	It defines groups of audio channels (channel configurations). These configurations include loudspeaker layouts for channel-based audio, such as stereo and 5.1.
audioChannelFormat (audio channel)	It defines changes in the position, gain, and other parameters of audio objects. Audio channels are divided into one or more audio blocks. By dividing them into multiple audio blocks, temporal variations in position and gain can be represented.
audioBlockFormat (audio block)	It defines the position, gain, and other parameters of an audio object at a specific point in time. Positional information is specified using either polar coordinates or Cartesian coordinates.

[17] Items in parentheses denote the names of each element as used in this document.

audioTrackUID (audio track)	It defines the correspondence between physical tracks and audio channels within the ADM.

Rendering methods and coordinate systems

In Rec. ITU-R BS.2127 which specifies the standard renderer for MGA, two types of rendering methods are defined: egocentric and allocentric. The egocentric renderer generally receives audio signals and audio-related metadata containing position information described in a polar coordinate system (φ, θ, d), i.e., trios of azimuth φ $(-180 \leq \varphi \leq 180)$, elevation θ $(-90 \leq \theta \leq 90)$, and distance d $(d = 1)$. It subsequently generates reproduced signals that maintain the directions of the audio objects relative to the position of the listener for any loudspeaker layout. Therefore, in the case that the loudspeaker layout differs between the production and listening environments, an audio object reproduced by a physical loudspeaker in the production environment may be reproduced as a phantom image in the listening environment to maintain the perceived direction of the audio object. Conversely, the allocentric renderer generally receives audio signals and audio-related metadata containing position information described in a Cartesian coordinate system (x, y, z), i.e., trios of left-right, front-back, and up-down displacements: X $(-1 \leq X \leq 1)$, Y $(-1 \leq Y \leq 1)$, and Z $(-1 \leq Z \leq 1)$. It subsequently generates reproduced signals that maintain locations of the audio objects relative to the positions of loudspeakers depending on the listening environment. Therefore, in the case that the loudspeaker layout differs between the production and listening environments, the perceived direction of an audio object may change, but its clarity is preserved.

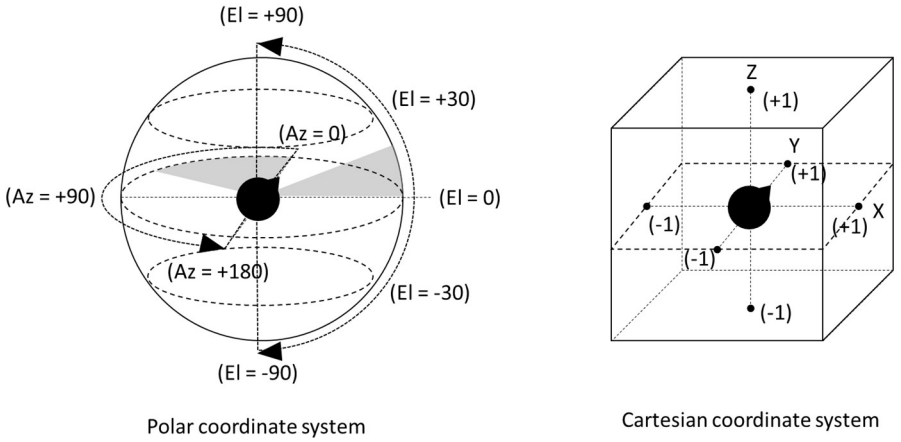


Figure 1 Polar (left) and Cartesian coordinate system (right) in ADM.

Note that, in both coordinate systems, the coordinates represent not physical positions in the listening environment, but virtual positions assumed in the audio-related metadata. In the polar coordinate system, a normalized spherical surround space is assumed, where the distance from the listener to the loudspeakers arranged on a sphere is normalized to 1. In contrast, in the Cartesian coordinate system, a normalized cubic room is assumed. The front-left, front-right, back-left, and back-right loudspeakers are assigned $(X, Y) = (-1, +1)$, $(+1, +1)$, $(-1, -1)$, and $(+1, -1)$, respectively, and the upper, middle, and lower layers are assigned $Z = +1, 0, -1$, respectively. Due to the differences in normalization concepts for position information, a general geometric conversion between polar and Cartesian coordinate is not applicable for MGA system. As a result, the correspondence between polar and Cartesian coordinate values depends on the listening environment and is not uniquely determined. For example, the Cartesian coordinate $(-1, -1, 0)$, corresponding to the “back-left” position in the normalized cubic room, maps to polar coordinates $(+110,\ 0, 1)$ in a 5.1.4 loudspeaker configuration. However, in 7.1.4 or 22.2 configurations, it corresponds to $(+135,\ 0, 1)$. This is because the “back-left” position in Cartesian coordinates is a virtual normalized position, while the directions to the actual physical loudspeakers differ across loudspeaker configurations.

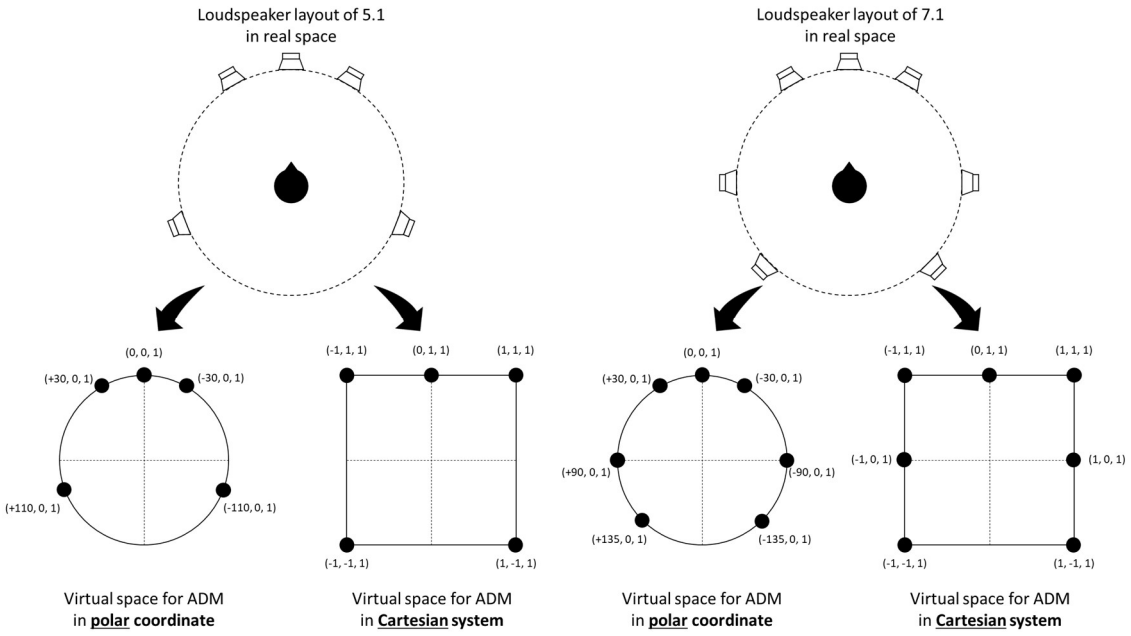


Figure 2 Example of correspondence between loudspeaker layouts in real space and those in virtual space for ADM (Left: 5.1, right: 7.1)

Rendering

The standard test materials are intended to be reproduced using playback software (e.g., the attachment “03_ARIBAudioPlayer”) after these materials have been rendered, based on audio-related metadata, into signals depending on the loudspeaker configuration of the listening environment. There are rendering algorithms capable of rendering the Audio Definition Model (ADM) defined in Recommendation ITU-R BS.2076, such as those specified in Recommendation ITU-R BS.2127 (https://www.itu.int/rec/R-REC-BS.2127/en). Renderers whose operation has been confirmed with this standard sound source (based on Recommendation ITU-R BS.2076-3) include MIST Studio by Marquise Technologies (https://www.marquise-tech.com/mist/), among others.

3. Structure of the Standard Test Materials

In a MGA system, the audio signals consisting of one or more audio objects and their associated audio-related metadata, are transmitted. The audio renderer in the receiver renders signals for reproduction via loudspeakers based on the audio-related metadata, the listening environment, and the user's control parameters. As a result, MGA services can provide listening-environment adaptation (providing sounds and functionality adapted to the user's listening environment), accessibility enhancement (providing accessibility-related sounds and functionality, such as multilingual broadcasting or enhancing the intelligibility of specific dialog), and personalization (providing functionality enabling users to select audio presets according to their preferences). The expected services and the corresponding functions of the audio-related metadata are shown in Table 2. In the Standard test materials for three-dimensional multichannel stereophonic sound system - Series B, the following six evaluation items have been selected so that the functions of the audio-related metadata required to realize these expected services can be evaluated.

No.1 "Channel Check"	: Evaluation material primarily intended to confirm that, for channel-based audio objects, the reproduction chain from the playback device to the loudspeakers is correctly connected and functioning.
No.2 "Localization"	: Evaluation material designed to verify the localization accuracy of panning for mono audio objects.
No.3 "Functional Check"	: Evaluation material primarily intended to verify the functions of audio-related metadata, such as accessibility enhancement and personalization.
No.4 "Simulcast"	: Evaluation material that simulates simultaneous delivery with different loudspeaker layouts in current 4K/8K television broadcasting. It is used to confirm correct switching between channel‐based audio objects encoded for multiple loudspeaker configurations.
No.5 "Downmix"	: Evaluation material designed to confirm the application of custom downmix coefficients specified within the audio‐related metadata when converting channel‐based audio objects to different loudspeaker configurations.

No.6 “Loudness”	: Evaluation material intended to verify loudness correction performed on the basis of the loudness‐related metadata defined within the audio‐related metadata.

The correspondence between the main evaluation items and the selected test materials is shown in Table 2. Multiple variants of each evaluation item are provided due to differences in evaluation conditions and differences in the coordinate systems used to describe the positional information of audio objects in the audio-related metadata. For example, the No.1 “Channel Check” material is intended to confirm whether the loudspeakers are correctly connected and whether each audio object is reproduced correctly at the intended loudspeaker position. However, due to variations in loudspeaker configuration (stereo, 5.1, 5.1.2, 5.1.4, 7.1, 7.1.4, 22.2) and variations in coordinate systems (polar, Cartesian), a total of 18 variants of this evaluation item are provided. Unless otherwise noted, the audio-related metadata in these test materials has been produced in conformity with the “ARIB ADM Completed Program Profile” specified in ARIB TR-B48. A complete list of these test materials is presented in Table 3.

Table 2 Expected services enabled by MGA and the corresponding functions of audio-related metadata

Purpose	Intended Service	Functions of Audio-Related Metadata (Descriptors in italics refer to those defined in ADM)
Listening-environment adaptation	Reproduction adapted to the user's home loudspeaker configuration	・ Conversion of channel-based audio objects using downmix coefficients independently defined via *Matrix*
	Loudness correction between programs	・ Correction of reproduction level based on the loudness value described in *loudnessMetadata*
	Services equivalent to current broadcast practices	・ Reproduction according to the loudspeaker configuration of channel-based audio objects (*type* = *0001*) ・ Panning of mono audio objects (*type* = *0003*) ・ Substitution of channel-based audio objects (*type* = *0001*) corresponding to simulcast-type delivery
Accessibility enhancement	Dialog enhancement[18]	・ Specification of static level adjustment values for dialog enhancement using *alternativeValueSet.* ・ Specification of user-adjustable level ranges using *gainInteractionRange.*
	Emergency broadcasting[19]	・ Restriction of user control (e.g., level adjustment) using parameters defined in *audioObjectInteraction*

[18] Dialogue enhancement: A service enabling either raising the level of specific dialogue only or lowering the levels of all components other than the target dialogue.
[19] Emergency broadcasting: A service related to information requiring immediate distribution, such as natural disasters (e.g., earthquakes and typhoons) and urgent evacuation information, or sound effects accompanying such information.

	Multilingual broadcasting	・ Substitution of mono audio objects (*type* = *0003*) for dialog by switching audio objects (*audioObject*) defined in *audioProgramme* or *audioComplementaryObject* ・ Specification of dynamic background-sound level control (ducking) using *gainControl*
	Audio guides, audio descriptions[20], and narration of captions	・ Simultaneous reproduction of multiple mono audio objects (*type* = *0003*) or channel-based audio objects (*type* = *0001*)
Personalization	Alternative commentary ("back-stage talk")[21], lifeline broadcasting for users with sensory disabilities[22], detailed commentary[23], commentary-off mode[24],	・ Substitution of mono audio objects (*type* = *0003*, dialog) by switching audio objects defined in *audioProgramme* or *audioComplementaryObject*

[20] Audio description: A service providing explanations of visual scenes or the movements of people, intended for users with visual impairments.
[21] Alternative commentary ("back-stage talk"): A service providing a secondary commentary track where performers, guests, or production staff speak freely in a more relaxed style, separate from the main program audio.
[22] Lifeline broadcasting for users with sensory disabilities: A service intended to supplement audio for users with visual or hearing impairments by providing accessible delivery of essential information, comparable to emergency broadcasting but also including daily-life information.
[23] Detailed commentary: A service intended for knowledgeable users, such as experienced participants of a sport, providing advanced expert commentary during sports broadcasts.
[24] Commentary-off mode: A service intended for users who prefer to focus on the environmental ambience of the venue (e.g., stadium sounds) by muting commentary and delivering program audio that emphasizes the sense of presence.

	karaoke mode[25], and "favorite-voice" mode[26]	
	Multi-angle audio[27], home-and-away[28], and alternative musical arrangements	・ Substitution of channel-based audio objects (*type = 0001*, background sound) by switching audio objects defined in *audioProgramme* or *audioComplementaryObject*
	User-dependent audio[29]	・ Specification of supplementary information for audio objects using *tagList*
	Simple user interface[30]	・ Specification of user-selectable presets using *tagList*

[25] Karaoke mode: A service for users who wish to sing or play instruments themselves, in which vocals or melodic elements are muted while another program audio is reproduced.
[26] "Favorite-voice" mode: A service intended for users who wish to listen only to a specific performer's voice muting the voices of all other performers in music or entertainment programs.
[27] Multi-angle audio: A service intended for programs that provide simultaneous multi-camera views or allow the user to select one of several camera feeds, offering background sound (e.g., environmental ambience at the camera position) that corresponds to the selected video angle.
[28] Home-and-away: A service intended for sports broadcasts such as football, providing ambience, commentary, and surrounding sound from the stands of each team's supporters ("home" and "away").
[29] User-dependent audio: A service that enables adding or restricting content according to the user's region or viewing context. Based on the additional information defined for audio objects in *tagList*, the system selects or restricts audio objects dynamically.
[30] Simple user interface: In a metadata-guided audio system, combinations of audio-program switching and audio-object switching normally allow the user to select a preferred audio mix. This service defines such combinations in *tagList* as preset audio configurations, enabling users to select their preferred audio simply by switching presets.

Table 3 Correspondence between the main evaluation Items and the test materials

Evaluation item / Test material	Channel−based audio objects			Mono audio objects	Dialog switching (type = 0003)		Background-sound switching (type = 0001)	
	Reproduction according to loudspeaker configuration	Simulcast	Downmix	Panning	APR switching	AO switching	APR switching	AO switching
No.1 “Channel Check”	◎		○	○				
No.2 “Localization”				◎				
No.3 “Functional Check”					○	○	○	○
A) Amusement park						◎		◎
B) Volleyball					◎	◎	◎	
C) Train								
D) Octet								
No.4 “Simulcast”		◎						
No.5 “Downmix”			◎					
No.6 “Loudness”								

◎: Highly suitable for evaluation

○: Suitable for evaluation

(Table 3 continued)

Evaluation item / Test material	Simultaneous reproduction of audio objects		Specification of the adjustable dialog-level range	Specification of dialog-level restrictions (non-adjustable)	Specification of adjustment values for dialog enhancement	Ducking (dynamic background-sound level control)	Supplementary control using *tagList*	Loudness correction
	Dialog (type=0003)	Background sound (type=0001)						
No.1 “Channel Check”								
No.2 “Localization”								
No.3 “Functional Check”	○	○	○	○	○	○	○	
A) Amusement park		◎		◎			◎	
B) Volleyball				◎			◎	
C) Train	◎			◎	◎		◎	
D) Octet			◎	◎		◎	◎	
No.4 “Simulcast”								
No.5 “Downmix”								
No.6 “Loudness”							○	◎

◎: Highly suitable for evaluation

○: Suitable for evaluation

Table 4 List of the test materials for MGA

No.	Test material		[sec]	Folder
1	Channel Check			01_ChannelCheck
	A) stereo	C	24.0	├ 01_A_ChannelCheck_stereo_cartesian
		P	20.0	├ 01_A_ChannelCheck_stereo_polar
	B) 5.1	C	49.0	├ 01_B_ChannelCheck_5.1_cartesian
		P	43.0	├ 01_B_ChannelCheck_5.1_polar
	C) 5.1.2	C	63.0	├ 01_C_ChannelCheck_5.1.2_cartesian
		P	55.0	├ 01_C_ChannelCheck_5.1.2_polar
	D) 5.1.4	C	76.0	├ 01_D_ChannelCheck_5.1.4_cartesian
		P	69.0	├ 01_D_ChannelCheck_5.1.4_polar
	E) 7.1	C	56.0	├ 01_E_ChannelCheck_7.1_cartesian
		P	52.0	├ 01_E_ChannelCheck_7.1_polar
	F) 7.1.4	C	80.0	├ 01_F_ChannelCheck_7.1.4_cartesian
		P	76.0	├ 01_F_ChannelCheck_7.1.4_polar
	G) 22.2-1	C	159	├ 01_G_ChannelCheck_22.2-1_cartesian
		P	152	├ 01_G_ChannelCheck_22.2-1_polar
	H) 22.2-2	C	159	├ 01_H_ChannelCheck_22.2-2_cartesian
		P	147	├ 01_H_ChannelCheck_22.2-2_polar
	I) 22.2-3	C	161	├ 01_I_ChannelCheck_22.2-3_cartesian
		P	151	└ 01_I_ChannelCheck_22.2-3_polar
2	Localization			02_Localization
	A) Left to right（30）	P	96.0	├ 02_A_Localization_LtoR_30_polar_detailed
		P	96.0	├ 02_A_Localization_LtoR_30_polar_simple
	Left to right（45）	P	96.0	├ 02_A_Localization_LtoR_45_polar_detailed
		P	96.0	├ 02_A_Localization_LtoR_45_polar_simple
	Left to right	C	96.0	├ 02_A_Localization_LtoR_cartesian_detailed
		C	96.0	├ 02_A_Localization_LtoR_cartesian_simple
	B) Up to down（30）	P	128	├ 02_B_Localization_UtoD_30_polar_detailed
		P	128	├ 02_B_Localization_UtoD_30_polar_simple
	Up to down（45）	P	128	├ 02_B_Localization_UtoD_45_polar_detailed
		P	128	├ 02_B_Localization_UtoD_45_polar_simple
	Up to down	C	128	├ 02_B_Localization_UtoD_cartesian_detailed
		C	128	├ 02_B_Localization_UtoD_cartesian_simple
	C) Rotation	C	39.0	├ 02_C_Localization_Rot_cartesian_detailed

		C	39.0	├ 02_C_Localization_Rot_cartesian_simple
		P	39.0	├ 02_C_Localization_Rot_polar_detailed
		P	39.0	├ 02_C_Localization_Rot_polar_simple
	D) Over	C	127	├ 02_D_Localization_Over_cartesian
		P	127	├ 02_D_Localization_Over_polar
	E) Helix	C	39.0	├ 02_E_Localization_Helix_cartesian
		P	39.0	├ 02_E_Localization_Helix_polar
	F) Static	C	53.0	├ 02_F_Localization_Static_cartesian
		P	53.0	└ 02_F_Localization_Static_polar
3	Function check			03_Function
	A) Amusement park	C	41.7	├ 03_A_Function_AmusementPark_cartesian
		P	41.7	├ 03_A_Function_AmusementPark_polar
	B) Volleyball	C	31.6	├ 03_B_Function_Volleyball_cartesian
		P	31.6	├ 03_B_Function_Volleyball_polar
	C) Train	C	54.5	├ 03_C_Function_Train_cartesian
		P	54.5	├ 03_C_Function_Train_polar
	D) Octet	C	80.3	├ 03_D_Function_Octet_cartesian
		P	80.3	└ 03_D_Function_Octet_polar
4	Simulcast			04_Simul
	A) Drama	C	23.2	├ 04_A_Simul_Drama_cartesian
		P	23.2	└ 04_A_Simul_Drama_polar
5	Downmix			05_Downmix
	A) Coefficient set 1	C	161	├ 05_A_Downmix_ChannelCheck_22.2_coefficient1_cartesian
		P	152	├ 05_A_Downmix_ChannelCheck_22.2_coefficient1_polar
	B) Coefficient set 2	C	161	├ 05_B_Downmix_ChannelCheck_22.2_coefficient2_cartesian
		P	152	└ 05_B_Downmix_ChannelCheck_22.2_coefficient2_polar
6	Loudness			06_Loudness
	A) measurement (+3 dB)	C	54.5	├ 06_A_Loudness_Train_measure_+3dB_cartesian
		P	54.5	├ 06_A_Loudness_Train_measure_+3dB_polar
	B) measurement (+0 dB)	C	54.5	├ 06_B_Loudness_Train_measure_+0dB_cartesian
		P	54.5	├ 06_B_Loudness_Train_measure_+0dB_polar
	C) measurement (-3 dB)	C	54.5	├ 06_C_Loudness_Train_measure_-3dB_cartesian
		P	54.5	├ 06_C_Loudness_Train_measure_-3dB_polar
	D) Creator's intent (+3 dB)	C	54.5	├ 06_D_Loudness_Train_intention_+3dB_cartesian
		P	54.5	├ 06_D_Loudness_Train_intention_+3dB_polar

	E) Creator’s intent （+0 dB）	C	54.5	├ 06_E_Loudness_Train_intention_+0dB_cartesian
		P	54.5	├ 06_E_Loudness_Train_intention_+0dB_polar
	F) Creator’s intent （-3 dB）	C	54.5	├ 06_F_Loudness_Train_intention_-3dB_cartesian
		P	54.5	└ 06_F_Loudness_Train_intention_-3dB_polar

3.1. Test Materials for MGA

For each test material, the following information is provided the file names contained in each folder, the duration of the material, the program structure, which shows the combinations of audio objects simultaneously reproduced, as defined in the audio-related metadata, the audio track configuration, which shows the correspondence between the audio tracks stored in WAV/BW64, the audio objects, and their audio packs defined in the audio-related metadata, and an description of aspects relevant to evaluating the material and the associated audio-related metadata.

In the following descriptions, each test material may contain multiple audio programs and audio objects. When referring to a specific audio program or audio object, the ID assigned within the audio-related metadata is used to distinguish them. The IDs for audio programs and audio objects are defined respectively as “APR_XXXX” for audio programs and “AO_XXXX” for audio objects where XXXX represents a four-digit hexadecimal value.

No.1 Channel check （01_ChannelCheck/）

Description:

Test material No.1, “Channel Check,” consists of an English female voice reading the loudspeaker position names and their corresponding coordinates value (in either polar or Cartesian coordinates) for the following loudspeaker configurations: stereo, 5.1, 5.1.2, 5.1.4, 7.1, 7.1.4, and 22.2. The channel order for each configuration is described in a later subsection. This test material contains two switchable audio programs: APR_1001 and APR_1002. APR_1001 consists of a mono audio object (Objects, type=0003) assigned to the positions of the loudspeakers and allows evaluation of the localization accuracy of mono audio objects at each loudspeaker position. Note that an audio object related to LFE (Low Frequency Effect) is defined as a channel-based audio object so that it can be reproduced from the LFE loudspeaker[31]. In contrast, APR_1002 consists of a channel-based audio object (DirectSpeakers, type=0001) corresponding to each loudspeaker configuration. This allows confirmation that the reproduction chain is correctly connected from the playback device to the loudspeakers and also enables evaluation of localization accuracy when format conversion is performed to different loudspeaker configurations. In particular, when converting formats between configurations with different rear-speaker azimuths in polar coordinates: 135° (7.1, 7.1.4, 22.2) versus 110° (5.1, 5.1.2, 5.1.4), this material enables comparison between localization obtained via panning of mono audio objects, and localization after format conversion of channel‐based audio objects, thus enabling assessment of perceptual differences. It is also suitable for assessing audio quality (e.g., whether the female voice is reproduced naturally).

The audio track configuration consists of multiple tracks, with Track 1 containing the audio signal for the mono audio object, and Track 2 and the subsequent tracks containing the audio signal for the channel-based audio objects. For “Channel Check” materials except for stereo contain two channel-based audio objects. The first one is an audio object that includes only signal for LFE in APR_1001, and the second one is an audio object that includes signals for all loudspeakers in APR_1002. Figure 3 shows an example of the audio track configuration for channel check material of 5.1ch.

[31] The panning of mono audio object ignores the LFE loudspeakers, thus these loudspeakers are not used for mono audio objects even in the case that their positions are set as the coordinate value of the LFE.

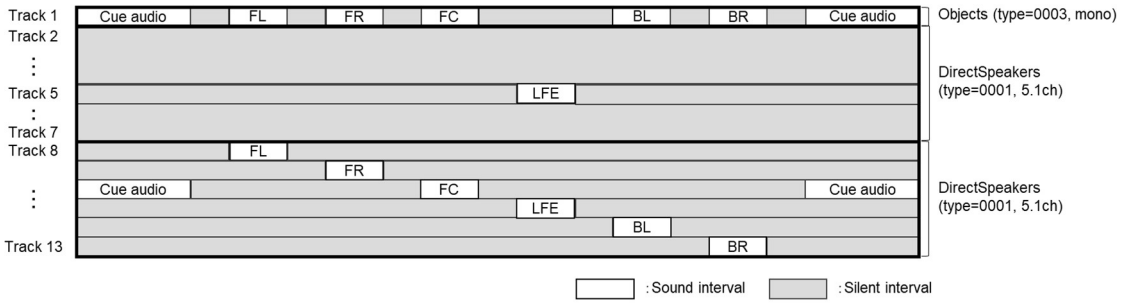


Figure 3 Audio Track Configuration (Channel Check – 5.1)

A) stereo

Files:

```
01_ChannelCheck
├── 01_A_ChannelCheck_stereo_cartesian
│   ├── 01_A_ChannelCheck_stereo_cartesian.bw64
│   ├── 01_A_ChannelCheck_stereo_cartesian.csv
│   ├── 01_A_ChannelCheck_stereo_cartesian.wav
│   └── 01_A_ChannelCheck_stereo_cartesian.xml
└── 01_A_ChannelCheck_stereo_polar
    ├── 01_A_ChannelCheck_stereo_polar.bw64
    ├── 01_A_ChannelCheck_stereo_polar.csv
    ├── 01_A_ChannelCheck_stereo_polar.wav
    └── 01_A_ChannelCheck_stereo_polar.xml
```

Duration: (Cartesian) 24.0 [s] / (polar) 20.0 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001	
APR_1002: channel check	AO_1002	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-3	AO_1002	DirectSpeakers (stereo)	AP_00010002 (polar) or AP_00010802 (Cartesian)

Channel order (stereo):

No.	**Loudspeaker Name**	**Position (in polar)**	**Position (in Cartesian)**
1	Front Left	(30, 0)	(-1, 1, 0)
2	Front Right	(-30, 0)	(1, 1, 0)

B) 5.1

Files:

```
01_ChannelCheck
├── 01_B_ChannelCheck_5.1_cartesian
│   ├── 01_B_ChannelCheck_5.1_cartesian.bw64
│   ├── 01_B_ChannelCheck_5.1_cartesian.csv
│   ├── 01_B_ChannelCheck_5.1_cartesian.wav
│   └── 01_B_ChannelCheck_5.1_cartesian.xml
└── 01_B_ChannelCheck_5.1_polar
    ├── 01_B_ChannelCheck_5.1_polar.bw64
    ├── 01_B_ChannelCheck_5.1_polar.csv
    ├── 01_B_ChannelCheck_5.1_polar.wav
    └── 01_B_ChannelCheck_5.1_polar.xml
```

Duration: (Cartesian) 49.0 [s] / (polar) 43.0 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-7	AO_1002※	DirectSpeakers (5.1)	AP_00010003 (polar) or AP_00010803 (Cartesian)
8-13	AO_1003	DirectSpeakers (5.1)	AP_00010003 (polar) or AP_00010803 (Cartesian)

※ All tracks except for the LFE track (Tracks 2–4 and 6–7) are silent.

Channel order（5.1）

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(30, 0)	(-1, 1, 0)
2	Front Right	(-30, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE	(0, -30)	(0, 1, -1)
5	Back Left	(110, 0)	(-1, -1, 0)
6	Back Right	(-110, 0)	(1, -1, 0)

C) 5.1.2

Files:

```
01_ChannelCheck
├── 01_C_ChannelCheck_5.1.2_cartesian
│   ├── 01_C_ChannelCheck_5.1.2_cartesian.bw64
│   ├── 01_C_ChannelCheck_5.1.2_cartesian.csv
│   ├── 01_C_ChannelCheck_5.1.2_cartesian.wav
│   └── 01_C_ChannelCheck_5.1.2_cartesian.xml
└── 01_C_ChannelCheck_5.1.2_polar
    ├── 01_C_ChannelCheck_5.1.2_polar.bw64
    ├── 01_C_ChannelCheck_5.1.2_polar.csv
    ├── 01_C_ChannelCheck_5.1.2_polar.wav
    └── 01_C_ChannelCheck_5.1.2_polar.xml
```

Duration: (Cartesian) 63.0 [s] / (polar) 55.0 [s]

Program structure: :

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: chennel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-9	AO_1002※	DirectSpeakers (5.1.2)	AP_00010004 (polar) or AP_00010804 (Cartesian)
10-17	AO_1003	DirectSpeakers (5.1.2)	AP_00010004(polar) or AP_00010804 (Cartesian)

※ All tracks except for the LFE track (Tracks 2–4 and 6–9) are silent.

Channel order (5.1.2)

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(30, 0)	(-1, 1, 0)
2	Front Right	(-30, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE	(0, -30)	(0, 1, -1)
5	Back Left	(110, 0)	(-1, -1, 0)
6	Back Right	(-110, 0)	(1, -1, 0)
7	Top Front Left	(30, 30)	(-1, 1, 1)
8	Top Front Right	(-30, 30)	(1, 1, 1)

D) 5.1.4

Files:

```
01_ChannelCheck
├── 01_D_ChannelCheck_5.1.4_cartesian
│   ├── 01_D_ChannelCheck_5.1.4_cartesian.bw64
│   ├── 01_D_ChannelCheck_5.1.4_cartesian.csv
│   ├── 01_D_ChannelCheck_5.1.4_cartesian.wav
│   └── 01_D_ChannelCheck_5.1.4_cartesian.xml
└── 01_D_ChannelCheck_5.1.4_polar
    ├── 01_D_ChannelCheck_5.1.4_polar.bw64
    ├── 01_D_ChannelCheck_5.1.4_polar.csv
    ├── 01_D_ChannelCheck_5.1.4_polar.wav
    └── 01_D_ChannelCheck_5.1.4_polar.xml
```

Duration: (Cartesian) 76.0 [s] / (polar) 69.0 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	・
APR_1002: channel check	AO_1003	・

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-11	AO_1002※	DirectSpeakers (5.1.4)	AP_00010005 (polar) or AP_00010805 (Cartesian)
12-21	AO_1003	DirectSpeakers (5.1.4)	AP_00010005 (polar) or AP_00010805 (Cartesian)

※ All tracks except for the LFE track (Tracks 2–4 and 6–11) are silent.

Channel order (5.1.4) :

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(30, 0)	(-1, 1, 0)
2	Front Right	(-30, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE	(0, -30)	(0, 1, -1)
5	Back Left	(110, 0)	(-1, -1, 0)
6	Back Right	(-110, 0)	(1, -1, 0)
7	Top Front Left	(30, 30)	(-1, 1, 1)
8	Top Front Right	(-30, 30)	(1, 1, 1)
9	Top Back Left	(110, 30)	(-1, -1, 1)
10	Top Back Right	(-110, 30)	(1, -1, 1)

E) 7.1

Files:

```
01_ChannelCheck
├── 01_E_ChannelCheck_7.1_cartesian
│   ├── 01_E_ChannelCheck_7.1_cartesian.bw64
│   ├── 01_E_ChannelCheck_7.1_cartesian.csv
│   ├── 01_E_ChannelCheck_7.1_cartesian.wav
│   └── 01_E_ChannelCheck_7.1_cartesian.xml
└── 01_E_ChannelCheck_7.1_polar
    ├── 01_E_ChannelCheck_7.1_polar.bw64
    ├── 01_E_ChannelCheck_7.1_polar.csv
    ├── 01_E_ChannelCheck_7.1_polar.wav
    └── 01_E_ChannelCheck_7.1_polar.xml
```

Duration: (Cartesian) 56.0 [s] / (polar) 52.0 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-9	AO_1002※	DirectSpeakers (7.1)	AP_0001000f (polar) or AP_0001080f (Cartesian)
10-17	AO_1003	DirectSpeakers (7.1)	AP_0001000f (polar) or AP_0001080f (Cartesian)

※ All tracks except for the LFE track (Tracks 2–4 and 6–9) are silent.

Channel order (7.1):

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(30, 0)	(-1, 1, 0)
2	Front Right	(-30, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE	(0, -30)	(0, 1, -1)
5	Side Left	(90, 0)	(-1, 0, 0)
6	Side Right	(-90, 0)	(1, 0, 0)
7	Back Left	(135, 0)	(-1, -1, 0)
8	Back Right	(-135, 0)	(1, -1, 0)

F) 7.1.4

Files:

```
01_ChannelCheck
├── 01_F_ChannelCheck_7.1.4_cartesian
│   ├── 01_F_ChannelCheck_7.1.4_cartesian.bw64
│   ├── 01_F_ChannelCheck_7.1.4_cartesian.csv
│   ├── 01_F_ChannelCheck_7.1.4_cartesian.wav
│   └── 01_F_ChannelCheck_7.1.4_cartesian.xml
└── 01_F_ChannelCheck_7.1.4_polar
    ├── 01_F_ChannelCheck_7.1.4_polar.bw64
    ├── 01_F_ChannelCheck_7.1.4_polar.csv
    ├── 01_F_ChannelCheck_7.1.4_polar.wav
    └── 01_F_ChannelCheck_7.1.4_polar.xml
```

Duration: (Cartesian) 80.0 [s] / (polar) 76.0 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-13	AO_1002※	DirectSpeakers (7.1.4)	AP_00010017(polar) or AP_00010817(Cartesian)
14-25	AO_1003	DirectSpeakers (7.1.4)	AP_00010017(polar) or AP_00010817(Cartesian)

※ All tracks except for the LFE track (Tracks 2–4 and 6–13) are silent.

Channel order (7.1.4) :

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(30, 0)	(-1, 1, 0)
2	Front Right	(-30, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE	(0, -30)	(0, 1, -1)
5	Side Left	(90, 0)	(-1, 0, 0)
6	Side Right	(-90, 0)	(1, 0, 0)
7	Back Left	(135, 0)	(-1, -1, 0)
8	Back Right	(-135, 0)	(1, -1, 0)
9	Top Front Left	(45, 30)	(-1, 1, 1)
10	Top Front Right	(-45, 30)	(1, 1, 1)
11	Top Back Left	(135, 30)	(-1, -1, 1)
12	Top Back Right	(-135, 30)	(1, -1, 1)

G) 22.2-1

Files:

```
01_ChannelCheck
├── 01_G_ChannelCheck_22.2-1_cartesian
│   ├── 01_G_ChannelCheck_22.2-1_cartesian.bw64
│   ├── 01_G_ChannelCheck_22.2-1_cartesian.csv
│   ├── 01_G_ChannelCheck_22.2-1_cartesian.wav
│   └── 01_G_ChannelCheck_22.2-1_cartesian.xml
└── 01_G_ChannelCheck_22.2-1_polar
    ├── 01_G_ChannelCheck_22.2-1_polar.bw64
    ├── 01_G_ChannelCheck_22.2-1_polar.csv
    ├── 01_G_ChannelCheck_22.2-1_polar.wav
    └── 01_G_ChannelCheck_22.2-1_polar.xml
```

Duration: (Cartesian) 159 [s] / (polar) 152 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-25	AO_1002※	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
26-49	AO_1003	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)

※ All tracks except for the LFE track (Tracks 2–4, 6–9 and 11–25) are silent.

Channel order（22.2-1）:

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(60, 0)	(-1, 1, 0)
2	Front Right	(-60, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE1	(45, -30)	(-1, 1, -1)
5	Back Left	(135, 0)	(-1, -1, 0)
6	Back Right	(-135, 0)	(1, -1, 0)
7	Front Center Left	(30, 0)	(-0.5, 1, 0)
8	Front Center Right	(-30, 0)	(0.5, 1, 0)
9	Back Center	(180, 0)	(0, -1, 0)
10	LFE2	(-45, -30)	(1, 1, -1)
11	Side Left	(90, 0)	(-1, 0, 0)
12	Side Right	(-90, 0)	(1, 0, 0)
13	Top Front Left	(45, 30)	(-1, 1, 1)
14	Top Front Right	(-45, 30)	(1, 1, 1)
15	Top Front Center	(0, 30)	(0, 1, 1)
16	Top Center	(0, 90)	(0, 0, 1)
17	Top Back Left	(135, 30)	(-1, -1, 1)
18	Top Back Right	(-135, 30)	(1, -1, 1)
19	Top Side Left	(90, 30)	(-1, 0, 1)
20	Top Side Right	(-90, 30)	(1, 0, 1)
21	Top Back Center	(180, 30)	(0, -1, 1)
22	Bottom Front Center	(0, -30)	(0, 1, -1)
23	Bottom Front Left	(45, -30)	(-1, 1, -1)
24	Bottom Front Right	(-45, -30)	(1, 1, -1)

H) 22.2-2

Files:

```
01_ChannelCheck
├── 01_H_ChannelCheck_22.2-2_cartesian
│   ├── 01_H_ChannelCheck_22.2-2_cartesian.bw64
│   ├── 01_H_ChannelCheck_22.2-2_cartesian.csv
│   ├── 01_H_ChannelCheck_22.2-2_cartesian.wav
│   └── 01_H_ChannelCheck_22.2-2_cartesian.xml
└── 01_H_ChannelCheck_22.2-2_polar
    ├── 01_H_ChannelCheck_22.2-2_polar.bw64
    ├── 01_H_ChannelCheck_22.2-2_polar.csv
    ├── 01_H_ChannelCheck_22.2-2_polar.wav
    └── 01_H_ChannelCheck_22.2-2_polar.xml
```

Duration: (Cartesian) 159 [s] / (polar) 147 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-25	AO_1002	DirectSpeakers (22.2)	AP_00011001 (polar)※ or AP_00010809 (Cartesian)
26-49	AO_1003	DirectSpeakers (22.2)	AP_00011001 (polar)※ or AP_00010809 (Cartesian)

※ All tracks except for the LFE track (Tracks 2-4, 6-9 and 11-25) are silent.

※ Since an audio pack with polar coordinates in the channel order shown below is not included in the common definitions[32], it has been uniquely defined as

[32] Elements and attributes such as audio packs and audio channels that can be used without explicit definition in the Audio Definition Model (ADM) specified in

AP_00011001. Therefore, the ADM file “01_H_ChannelCheck_22.2-2_polar.xml” does not conform to the ARIB ADM Completed Program Profile.

Channel order（22.2-2）:

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Left	(45, 0)	(-1, 1, 0)
2	Front Right	(-45, 0)	(1, 1, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE1	(45, -30)	(-1, 1, -1)
5	Back Left	(135, 0)	(-1, -1, 0)
6	Back Right	(-135, 0)	(1, -1, 0)
7	Front Center Left	(30, 0)	(-0.5, 1, 0)
8	Front Center Right	(-30, 0)	(0.5, 1, 0)
9	Back Center	(180, 0)	(0, -1, 0)
10	LFE2	(-45, -30)	(1, 1, -1)
11	Side Left	(90, 0)	(-1, 0, 0)
12	Side Right	(-90, 0)	(1, 0, 0)
13	Top Front Left	(45, 30)	(-1, 1, 1)
14	Top Front Right	(-45, 30)	(1, 1, 1)
15	Top Front Center	(0, 30)	(0, 1, 1)
16	Top Center	(0, 90)	(0, 0, 1)
17	Top Back Left	(135, 30)	(-1, -1, 1)
18	Top Back Right	(-135, 30)	(1, -1, 1)
19	Top Side Left	(90, 30)	(-1, 0, 1)
20	Top Side Right	(-90, 30)	(1, 0, 1)
21	Top Back Center	(180, 30)	(0, -1, 1)
22	Bottom Front Center	(0, -30)	(0, 1, -1)
23	Bottom Front Left	(45, -30)	(-1, 1, -1)
24	Bottom Front Right	(-45, -30)	(1, 1, -1)

Recommendation ITU-R BS.2094-2: “Common definitions for the audio definition model” (2025).

I) 22.2-3

Files:

```
01_ChannelCheck
├── 01_I_ChannelCheck_22.2-3_cartesian
│   ├── 01_I_ChannelCheck_22.2-3_cartesian.bw64
│   ├── 01_I_ChannelCheck_22.2-3_cartesian.csv
│   ├── 01_I_ChannelCheck_22.2-3_cartesian.wav
│   └── 01_I_ChannelCheck_22.2-3_cartesian.xml
└── 01_I_ChannelCheck_22.2-3_polar
    ├── 01_I_ChannelCheck_22.2-3_polar.bw64
    ├── 01_I_ChannelCheck_22.2-3_polar.csv
    ├── 01_I_ChannelCheck_22.2-3_polar.wav
    └── 01_I_ChannelCheck_22.2-3_polar.xml
```

Duration: (Cartesian) 161 [s] / (polar) 151 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001
2-25	AO_1002※	DirectSpeakers (22.2)	AP_0001001a (polar) or AP_0001081a (Cartesian)
26-49	AO_1003	DirectSpeakers (22.2)	AP_0001001a (polar) or AP_0001081a (Cartesian)

※ All tracks except for the LFE track (Tracks 2-4, 6-9 and 11-25) are silent.

※ Since the audio packs in polar and Cartesian coordinates for the channel order shown below are not included among the available audio packs specified in ARIB TR-B48, the audio-related metadata files “01_I_ChannelCheck_22.2-3_polar.xml”

and “01_I_ChannelCheck_22.2-3_cartesian.xml” do not conform to the ARIB ADM Completed Program Profile.

Channel order（22.2-3）:

No.	Loudspeaker Name	Position (in polar)	Position (in Cartesian)
1	Front Wide Left	(60, 0)	(-1, 0.667, 0)
2	Front Wide Right	(-60, 0)	(1, 0.667, 0)
3	Front Center	(0, 0)	(0, 1, 0)
4	LFE1	(45, -30)	(-1, 1, -1)
5	Back Left	(135, 0)	(-1, -1, 0)
6	Back Right	(-135, 0)	(1, -1, 0)
7	Front Left	(30, 0)	(-1, 1, 0)
8	Front Right	(-30, 0)	(1, 1, 0)
9	Back Center	(180, 0)	(0, -1, 0)
10	LFE2	(-45, -30)	(1, 1, -1)
11	Side Left	(90, 0)	(-1, 0, 0)
12	Side Right	(-90, 0)	(1, 0, 0)
13	Top Front Left	(45, 30)	(-1, 1, 1)
14	Top Front Right	(-45, 30)	(1, 1, 1)
15	Top Front Center	(0, 30)	(0, 1, 1)
16	Top Center	(0, 90)	(0, 0, 1)
17	Top Back Left	(135, 30)	(-1, -1, 1)
18	Top Back Right	(-135, 30)	(1, -1, 1)
19	Top Side Left	(90, 30)	(-1, 0, 1)
20	Top Side Right	(-90, 30)	(1, 0, 1)
21	Top Back Center	(180, 30)	(0, -1, 1)
22	Bottom Front Center	(0, -30)	(0, 1, -1)
23	Bottom Front Left	(45, -30)	(-1, 1, -1)
24	Bottom Front Right	(-45, -30)	(1, 1, -1)

No.2 Localization （02_Localization/）

A) Dynamic Object – Left-Right Motion

Files:

```
02_Localization
├── 02_A_Localization_LtoR_30_polar_detailed
│   ├── 02_A_Localization_LtoR.csv
│   ├── 02_A_Localization_LtoR.wav
│   ├── 02_A_Localization_LtoR_30_polar_detailed.bw64
│   └── 02_A_Localization_LtoR_30_polar_detailed.xml
├── 02_A_Localization_LtoR_30_polar_simple
│   ├── 02_A_Localization_LtoR.csv
│   ├── 02_A_Localization_LtoR.wav
│   ├── 02_A_Localization_LtoR_30_polar_simple.bw64
│   └── 02_A_Localization_LtoR_30_polar_simple.xml
├── 02_A_Localization_LtoR_45_polar_detailed
│   ├── 02_A_Localization_LtoR.csv
│   ├── 02_A_Localization_LtoR.wav
│   ├── 02_A_Localization_LtoR_45_polar_detailed.bw64
│   └── 02_A_Localization_LtoR_45_polar_detailed.xml
├── 02_A_Localization_LtoR_45_polar_simple
│   ├── 02_A_Localization_LtoR.csv
│   ├── 02_A_Localization_LtoR.wav
│   ├── 02_A_Localization_LtoR_45_polar_simple.bw64
│   └── 02_A_Localization_LtoR_45_polar_simple.xml
├── 02_A_Localization_LtoR_cartesian_detailed
│   ├── 02_A_Localization_LtoR.csv
│   ├── 02_A_Localization_LtoR.wav
│   ├── 02_A_Localization_LtoR_cartesian_detailed.bw64
│   └── 02_A_Localization_LtoR_cartesian_detailed.xml
└── 02_A_Localization_LtoR_cartesian_simple
    ├── 02_A_Localization_LtoR.csv
    ├── 02_A_Localization_LtoR.wav
    ├── 02_A_Localization_LtoR_cartesian_simple.bw64
    └── 02_A_Localization_LtoR_cartesian_simple.xml
```

Duration: 96.0 [s]

Program structure:

Audio program	Audio object	Remarks
APR_1001: Left-Right Motion	AO_1001	·

Audio track configuration:

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001

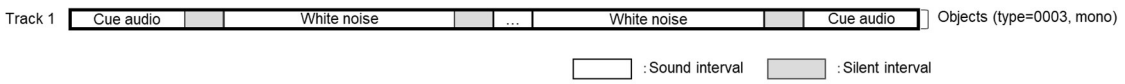


Figure 4 Audio track configuration (Localization)

Description:

The test material, “No.2 Localization A) Dynamic Object – Left-Right Motion,” consists of white noise moving horizontally (left to right and then back to the left) based on the positional information described in the audio-related metadata. This enables evaluation of localization accuracy by perceiving the position of the white noise (red circles in the figure indicate the audio object locations). In the polar coordinate system, initially, the white-noise audio object moves left and right twice at a constant angular velocity between top front left and top front right positions at azimuths of $\pm 45°$ and $\pm 30°$, in separate conditions. The elevation is then varied (five levels in total), and the same horizontal motion is repeated at each elevation level. In the Cartesian coordinate system, the white-noise audio object moves left and right twice at a constant speed between top front left and top front right positions at (X, Y) = (−1.0, 1.0) and (X, Y) = (1.0, 1.0), respectively. The elevation is then varied (five levels in total), and the same horizontal motion is repeated at each elevation level. Note that the terms “constant angular velocity” and “constant speed” used here refer to the motion within the virtual space assumed in ADM. Depending on the actual loudspeaker arrangement in the playback environment, the reproduced motion may not necessarily be perceived as constant angular velocity

or constant speed. For each condition, two versions of the test material are provided, differing in the specification of positional information in ADM:

- “_simple”: the entire left-to-right motion trajectory is defined as a single audio block.
- “_detailed”: the motion is defined using two audio blocks: one from the front left to the front center position, and another from the front center position to the front right.

This test material is suitable for evaluating on-screen sound localization, particularly the perception of horizontal motion, and comparing perceptual differences in sound localization resulting from different specifications of positional information of audio objects. The audio track configuration is shown in Figure 4 and consists of a single track, with Track 1 containing the audio signal for the mono audio object, including cue audio and white noise.

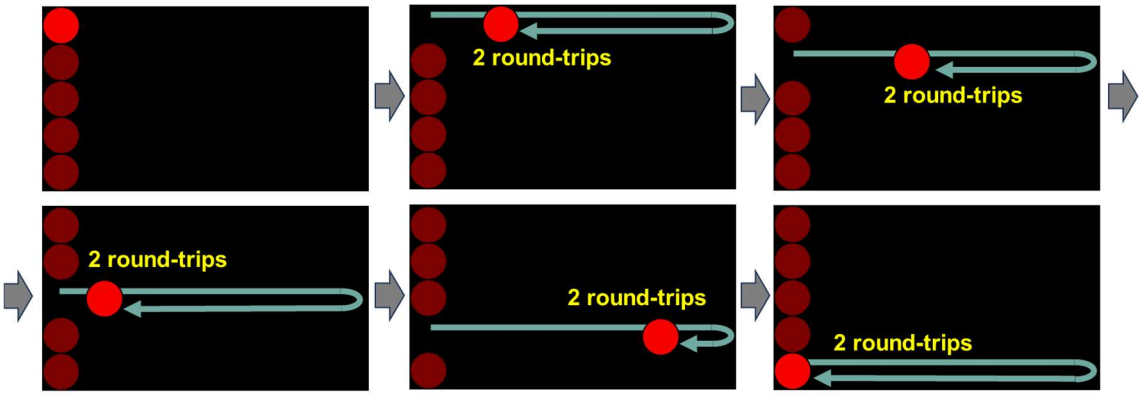


Figure 5 Illustration of audio object motion (left-right motion on screen)

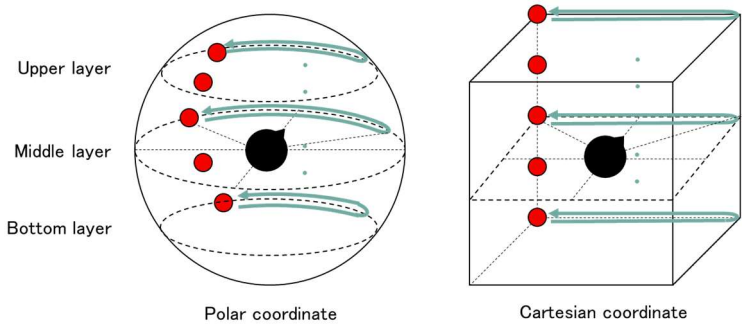


Figure 6 Transition of audio object positional information (left-right motion)

B) Dynamic Object – Up-Down Motion

Files:

```
02_Localization
├── 02_B_Localization_UtoD_30_polar_detailed
│   ├── 02_B_Localization_UtoD.csv
│   ├── 02_B_Localization_UtoD.wav
│   ├── 02_B_Localization_UtoD_30_polar_detailed.bw64
│   └── 02_B_Localization_UtoD_30_polar_detailed.xml
├── 02_B_Localization_UtoD_30_polar_simple
│   ├── 02_B_Localization_UtoD.csv
│   ├── 02_B_Localization_UtoD.wav
│   ├── 02_B_Localization_UtoD_30_polar_simple.bw64
│   └── 02_B_Localization_UtoD_30_polar_simple.xml
├── 02_B_Localization_UtoD_45_polar_detailed
│   ├── 02_B_Localization_UtoD.csv
│   ├── 02_B_Localization_UtoD.wav
│   ├── 02_B_Localization_UtoD_45_polar_detailed.bw64
│   └── 02_B_Localization_UtoD_45_polar_detailed.xml
├── 02_B_Localization_UtoD_45_polar_simple
│   ├── 02_B_Localization_UtoD.csv
│   ├── 02_B_Localization_UtoD.wav
│   ├── 02_B_Localization_UtoD_45_polar_simple.bw64
│   └── 02_B_Localization_UtoD_45_polar_simple.xml
├── 02_B_Localization_UtoD_cartesian_detailed
│   ├── 02_B_Localization_UtoD.csv
│   ├── 02_B_Localization_UtoD.wav
│   ├── 02_B_Localization_UtoD_cartesian_detailed.bw64
│   └── 02_B_Localization_UtoD_cartesian_detailed.xml
└── 02_B_Localization_UtoD_cartesian_simple
    ├── 02_B_Localization_UtoD.csv
    ├── 02_B_Localization_UtoD.wav
    ├── 02_B_Localization_UtoD_cartesian_simple.bw64
    └── 02_B_Localization_UtoD_cartesian_simple.xml
```

Duration: 128 [s]

Program structure:

Audio program	Audio object	Remarks
APR_1001: Up–Down Motion	AO_1001	

Audio track configuration:

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001

Description:

The test material, “No.2 Localization B) Dynamic Object – Up-Down Motion,” consists of white noise moving vertically (from the upper to the bottom layers) based on the positional information described in the audio-related metadata. This enables evaluation of localization accuracy by perceiving the position of the white noise (red circles in the figure indicate the audio object locations). In the polar coordinate system, the white-noise audio object moves down and then up twice at a constant angular velocity between top front left and bottom front left positions at azimuths of $\pm 45°$ and $\pm 30°$, in separate conditions. The horizontal position is then shifted to the right (nine steps in total), and the same vertical motion is repeated at each azimuth. In the Cartesian coordinate system, the white-noise audio object moves down and up twice at a constant speed between top front left and bottom front left positions at (X, Z) = (−1.0, 1.0) and (X, Z) = (−1.0, −1.0), respectively. The horizontal position is then shifted to the right (nine steps in total), and the same vertical motion is repeated. Note that the terms “constant angular velocity” and “constant speed” used here refer to the motion within the virtual space assumed in ADM. Depending on the actual loudspeaker arrangement in the playback environment, the reproduced motion may not necessarily be perceived as constant angular velocity or constant speed. For each condition, two versions of the test material are provided, differing in the specification of positional information in ADM:

- “_simple”: the entire motion from the upper to the bottom layers is defined as a single audio block.
- “_detailed”: the motion is defined using two audio blocks: one from the upper layer to the middle layer, and another from the middle layer to the bottom layer.

This test material is suitable for evaluating on-screen sound localization, particularly the perception of vertical motion, and comparing perceptual differences in sound localization resulting from different specifications of positional information of audio objects. The audio track configuration is the same as “No.2 Localization A) Dynamic Object – Left-right Motion” (Figure 4). It consists of a single track which contains the audio signal for the mono audio object, including cue audio and white noise.

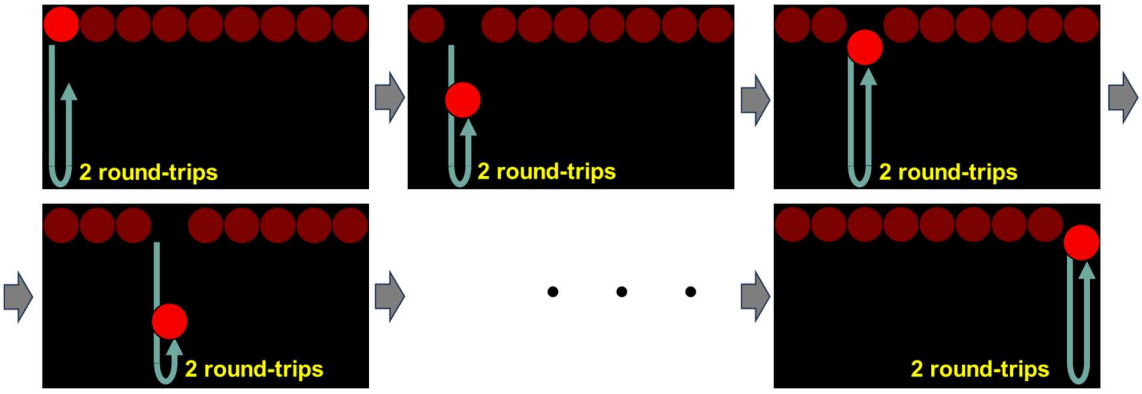


Figure 7 Illustration of audio object motion (up-down motion on screen)

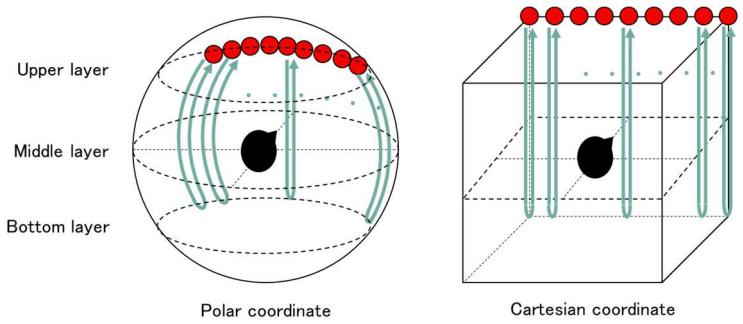


Figure 8 Transition of audio object positional information (up-down motion)

C) Dynamic Object – Rotation

Files:

```
02_Localization
├── 02_C_Localization_Rot_cartesian_detailed
│   ├── 02_C_Localization_Rot.csv
│   ├── 02_C_Localization_Rot.wav
│   ├── 02_C_Localization_Rot_cartesian_detailed.bw64
│   └── 02_C_Localization_Rot_cartesian_detailed.xml
├── 02_C_Localization_Rot_cartesian_simple
│   ├── 02_C_Localization_Rot.csv
│   ├── 02_C_Localization_Rot.wav
│   ├── 02_C_Localization_Rot_cartesian_simple.bw64
│   └── 02_C_Localization_Rot_cartesian_simple.xml
├── 02_C_Localization_Rot_polar_detailed
│   ├── 02_C_Localization_Rot.csv
│   ├── 02_C_Localization_Rot.wav
│   ├── 02_C_Localization_Rot_polar_detailed.bw64
│   └── 02_C_Localization_Rot_polar_detailed.xml
└── 02_C_Localization_Rot_polar_simple
    ├── 02_C_Localization_Rot.csv
    ├── 02_C_Localization_Rot.wav
    ├── 02_C_Localization_Rot_polar_simple.bw64
    └── 02_C_Localization_Rot_polar_simple.xml
```

Duration: 39 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: Rotation	AO_1001	

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001

Description:

The test material, “No.2 Localization C) Dynamic Object – Rotation,” consists of white noise rotating around the listener based on the positional information described in the audio-related metadata. This enables evaluation of localization accuracy by perceiving the position of the white noise (red circles in the figure indicate the audio object locations). In this material, the white-noise audio object rotates counterclockwise from the front in the middle layer and then rotates clockwise back to the front in the middle layer. In the polar coordinate system, the audio object moves at a constant angular velocity. In the Cartesian coordinate system, the audio object moves at a constant speed along the virtual room’s walls. Note that the terms “constant angular velocity” and “constant speed” used here refer to the motion within the virtual space assumed in ADM. Depending on the actual loudspeaker arrangement in the playback environment, the reproduced motion may not necessarily be perceived as constant angular velocity or constant speed. For each condition, two versions of the test material are provided, differing in the specification of positional information in ADM:

- “_simple”: the rotational motion is defined using five audio blocks[33].
- “_detailed”: the rotational motion is defined using eight audio blocks[34].

This test material is suitable for evaluating the perception of horizontal rotational motion, and comparing perceptual differences in sound localization resulting from different specifications of positional information of audio objects. The audio track configuration is the same as “No.2 Localization A) Dynamic Object – Left-right Motion” (Figure 4). It consists of a single track which contains the audio signal for the mono audio object, including cue audio and white noise.

[33] In the counterclockwise direction, the motion is defined by five blocks: (1) front-center to front-left, (2) front-left to rear-left, (3) back-left to back-right, (4) back-right to front-right, and (5) front-right to front-center.

[34] In the counterclockwise direction, the motion is defined by eight blocks: (1) front-center to front-left, (2) front-left to side-left, (3) side-left to back-left, (4) back-left to back, (5) back-center to back-right, (6) back-right to side-right, (7) side-right to front-right, and (8) front-right to front-center.

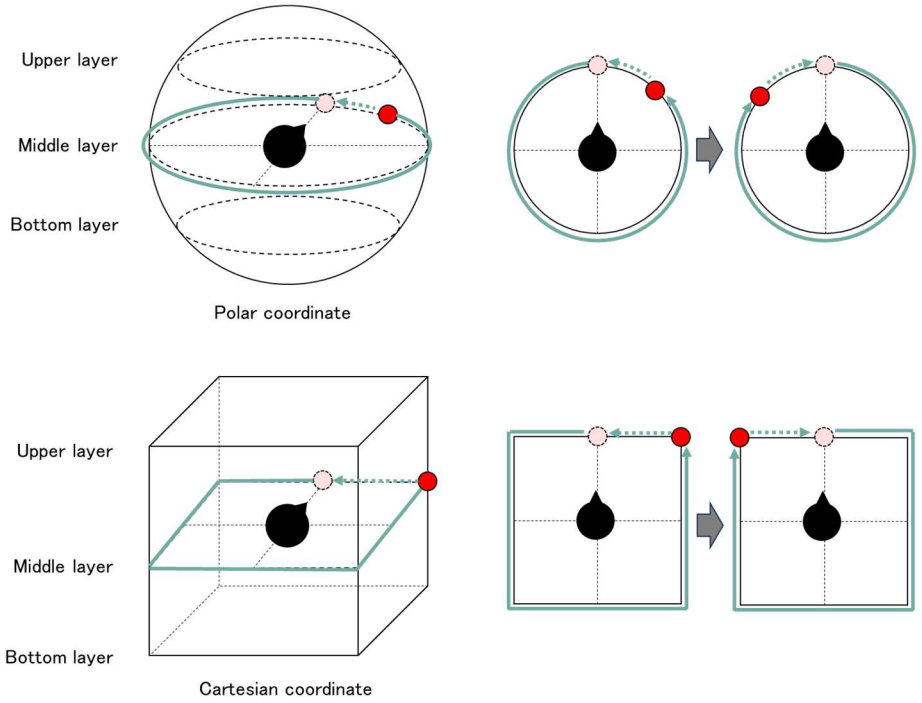


Figure 9 Transition of audio object positional information (rotation)

D) Dynamic Object – Over

Files:

```
02_Localization
├── 02_D_Localization_Over_cartesian
│   ├── 02_D_Localization_Over.csv
│   ├── 02_D_Localization_Over.wav
│   ├── 02_D_Localization_Over_cartesian.bw64
│   └── 02_D_Localization_Over_cartesian.xml
└── 02_D_Localization_Over_polar
    ├── 02_D_Localization_Over.csv
    ├── 02_D_Localization_Over.wav
    ├── 02_D_Localization_Over_polar.bw64
    └── 02_D_Localization_Over_polar.xml
```

Duration: 127 [s]

Program structure:

Audio program	Audio object	Remarks
APR_1001: Over	AO_1001	

Audio track configuration:

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001

Description:

The test material, “No.2 Localization D) Dynamic Object – Over” consists of white noise moving over the listener’s head based on the positional information described in the audio-related metadata. This enables evaluation of localization accuracy by perceiving the position of the white noise (red circles in the figure indicate the audio object locations). The white-noise audio object moves from the front in the middle layer, passes overhead, and then moves to the rear in the middle layer. It then returns from the rear, again passes overhead, back to the front. This sequence is repeated twice. Similarly, this material includes: a left-to-right overhead pass repeated twice, a pass from the front-left position to the back-right position repeated

twice, and a pass from the front-right position to the back-left position repeated twice. In the polar coordinate system, the audio object moves at a constant angular velocity. In the Cartesian coordinate system, the audio object moves at a constant speed along the virtual room's ceiling and walls. Note that the terms "constant angular velocity" and "constant speed" used here refer to the motion within the virtual space assumed in ADM. Depending on the actual loudspeaker arrangement in the playback environment, the reproduced motion may not necessarily be perceived as constant angular velocity or constant speed. This test material is suitable for evaluating the perception of vertical motion in sagittal and transverse planes. The audio track configuration is the same as "No.2 Localization A) Dynamic Object – Left-right Motion" (Figure 4). It consists of a single track which contains the audio signal for the mono audio object, including cue audio and white noise.

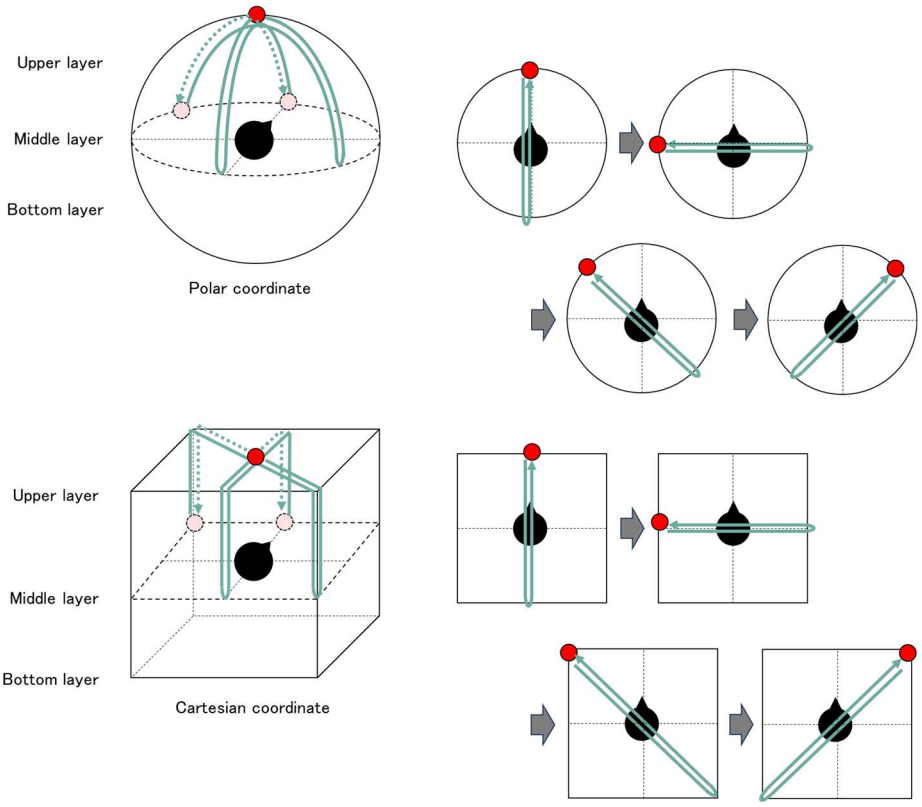


Figure 10 Transition of audio object positional information (Over)

E) Dynamic Object – Helix

Files:

```
02_Localization
├── 02_E_Localization_Helix_cartesian
│   ├── 02_E_Localization_Helix.csv
│   ├── 02_E_Localization_Helix.wav
│   ├── 02_E_Localization_Helix_cartesian.bw64
│   └── 02_E_Localization_Helix_cartesian.xml
└── 02_E_Localization_Helix_polar
    ├── 02_E_Localization_Helix.csv
    ├── 02_E_Localization_Helix.wav
    ├── 02_E_Localization_Helix_polar.bw64
    └── 02_E_Localization_Helix_polar.xml
```

Duration: 39 [s]

Program structure:

Audio program	Audio object	Remarks
APR_1001: Helix	AO_1001	

Audio track configuration:

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001

Description:

The test material, “No.2 Localization E) Dynamic Object – Helix,” consists of white noise moving in a spiral trajectory downward from the upper layer to the middle layer, based on the positional information described in the audio-related metadata. This enables evaluation of localization accuracy perceiving the position of the white noise (red circles in the figure indicate the audio object locations). The white-noise audio object starts at the upper-front-center position, moves downward with rotating counterclockwise for one full rotation until it reaches the front-center position, then returns to the upper-front-center with rotating clockwise. In the polar coordinate system, the audio object moves at a constant angular velocity. In the Cartesian

coordinate system, the audio object moves at a constant speed along the virtual room's walls. Note that the terms "constant angular velocity" and "constant speed" used here refer to the motion within the virtual space assumed in ADM. Depending on the actual loudspeaker arrangement in the playback environment, the reproduced motion may not necessarily be perceived as constant angular velocity or constant speed. This test material is suitable for evaluating the perception of horizontal rotation with upward or downward movement. The audio track configuration is the same as "No.2 Localization A) Dynamic Object – Left-right Motion" (Figure 4). It consists of a single track which contains the audio signal for the mono audio object, including cue audio and white noise.

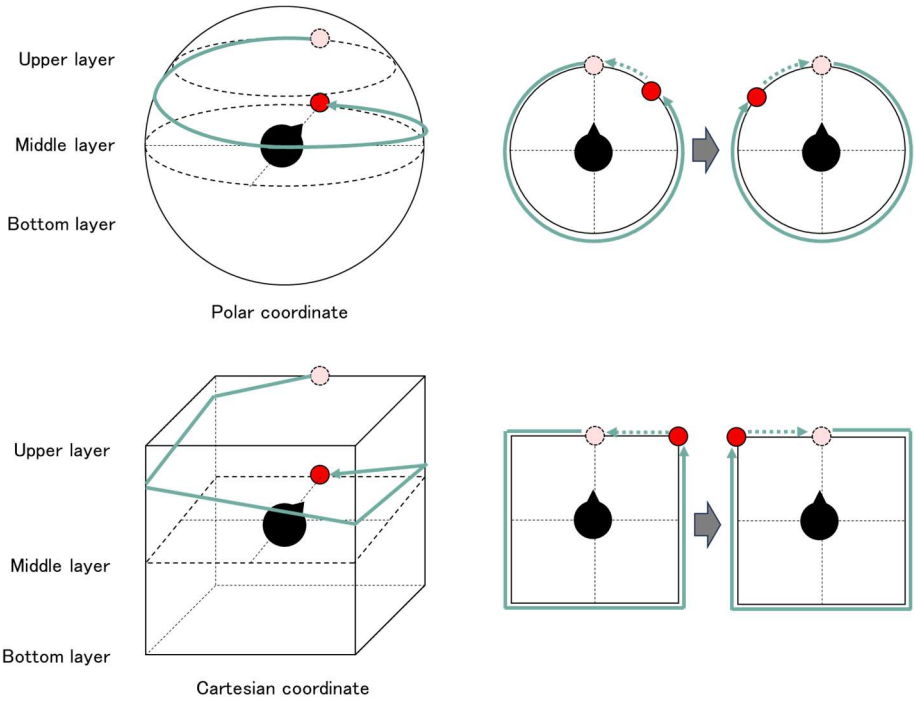


Figure 11 Transition of audio object positional information (Helix)

F) Static Object

Files:

```
02_Localization
├── 02_F_Localization_Static_cartesian
│   ├── 02_F_Localization_Static.csv
│   ├── 02_F_Localization_Static.wav
│   ├── 02_F_Localization_Static_cartesian.bw64
│   └── 02_F_Localization_Static_cartesian.xml
└── 02_F_Localization_Static_polar
    ├── 02_F_Localization_Static.csv
    ├── 02_F_Localization_Static.wav
    ├── 02_F_Localization_Static_polar.bw64
    └── 02_F_Localization_Static_polar.xml
```

Duration: 53 [s]

Program structure:

Audio program	Audio object	Remarks
APR_1001: 静的	AO_1001	

Audio track configuration:

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001

Description:

The test material, “No.2 Localization F) Static Objects,” consists of white noise with multiple fixed positions, based on the positional information described in the audio-related metadata. This enables evaluation of localization accuracy by perceiving the position of the white noise (red circles in the figure indicate the audio object locations). Whether polar or Cartesian coordinates are used, the positional information of the white-noise audio object is defined to represent the following seven positions: top front center, left half of the front (1/4 height), left side (3/4 height), back left (1/2 height), left half of the rear (1/2 height), front left on the ceiling (halfway

between the upper layer and the top), and rear (1/4 below overhead). The coordinate values for each audio object are as shown in the following table:

No.	Audio object position	Polar coordinate	Cartesian coordinate
1	Top front center	(0.000, 30.0, 1.00)	(0.000, 1.00, 1.00)
2	Left half of the front (1/4 height)	(15.0, 7.50, 1.00)	(-0.500, 1.00, 0.250)
3	Left side (3/4 height)	(90.0, 22.5, 1.00)	(-1.00, 0.400, 0.750)
4	Back left (1/2 height)	(120, 15.0, 1.00)	(-1.00, -0.670, 0.500)
5	Left half of the rear (1/2 height)	(165, 15.0, 1.00)	(-0.330, -1.00, 0.500)
6	Front left on the ceiling (halfway between the upper layer and the top)	(45.0, 60.0, 1.00)	(-0.500, 0.500, 1.00)
7	Rear (1/4 below overhead)	(180, 75.0, 1.00)	(0.000, -0.250, 1.00)

The white-noise audio object stays at each position for a fixed duration, then jumps to the next position in the sequence (from 1 to 7 in order). This test material is suitable for evaluating the perception of static audio objects located between the middle layer and the top and comparing perceptual differences in sound localization resulting from different coordinates for the same intended positions. The audio track configuration is the same as “No.2 Localization A) Dynamic Object – Left-right Motion” (Figure 4). It consists of a single track which contains the audio signal for the mono audio object, including cue audio and white noise.

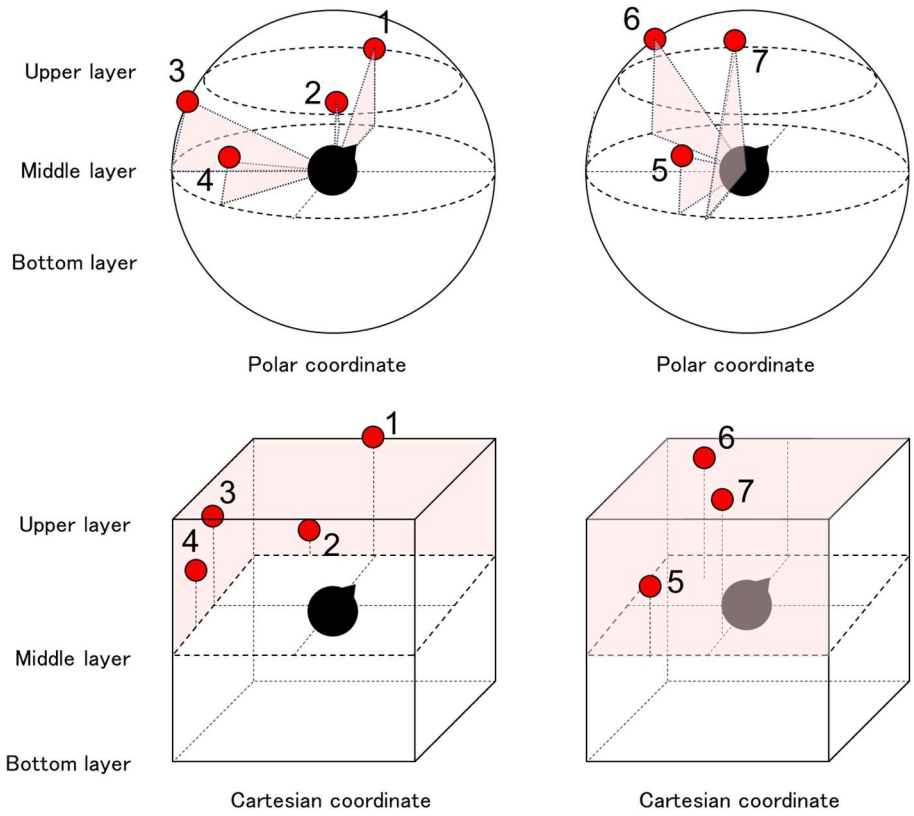


Figure 12 audio objects positional information (Static)

No.3 Function check （03_Function/）

A) Amusement park

Files：

```
03_Function
├── 03_A_Function_AmusementPark_cartesian
│   ├── 03_A_Function_AmusementPark.wav
│   ├── 03_A_Function_AmusementPark_cartesian.bw64
│   ├── 03_A_Function_AmusementPark_cartesian.csv
│   └── 03_A_Function_AmusementPark_cartesian.xml
└── 03_A_Function_AmusementPark_polar
    ├── 03_A_Function_AmusementPark.wav
    ├── 03_A_Function_AmusementPark_polar.bw64
    ├── 03_A_Function_AmusementPark_polar.csv
    └── 03_A_Function_AmusementPark_polar.xml
```

Duration： 41.7 [s]

Program structure：

Audio program	**Audio object**	**Remarks**
APR_1001: Standard Version	AO_1001：Cue audio (JPN) AO_1002：Cue audio (ENG) AO_1003：Cue audio (FRE)	・ User control not permitted. ・ AO_1001 can be substituted with AO_1002 or AO_1003
	AO_1011 : Ambient Sound (Multi-microphone) AO_1012 : Ambient Sound (One-point microphone)	・ AO_1011 can be substituted with AO_1012
	AO_1021：Narration (JPN) AO_1022：Narration (ENG) AO_1023：Narration (FRE)	・ AO_1021 can be substituted with AO_1022 or AO_1023
APR_1002: With Audio Commentary	AO_1001：Cue audio (JPN)	・ User control not permitted.
	AO_1011 : Ambient Sound (Multi-microphone) AO_1012 : Ambient Sound (One-point microphone)	・ AO_1011 can be substituted with AO_1012
	AO_1021：Narration (JPN)	

	AO_1031: Audio commentary (JPN)	
APR_1003: Without Narration	AO_1001：Cue audio (JPN) AO_1002：Cue audio (ENG) AO_1003：Cue audio (FRE)	・ User control not permitted ・ AO_1001 can be substituted with AO_1002 or AO_1003
	AO_1011 : Ambient Sound (Multi-microphone) AO_1012 : Ambient Sound (One-point microphone)	・ AO_1011 can be substituted with AO_1012

Audio track configuration：

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001
2	AO_1002	Objects (mono)	AP_00031002
3	AO_1003	Objects (mono)	AP_00031003
4-27	AO_1011	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
28-51	AO_1012	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
52	AO_1021	Objects (mono)	AP_00031021
53	AO_1022	Objects (mono)	AP_00031022
54	AO_1023	Objects (mono)	AP_00031023
55	AO_1031	Objects (mono)	AP_00031031

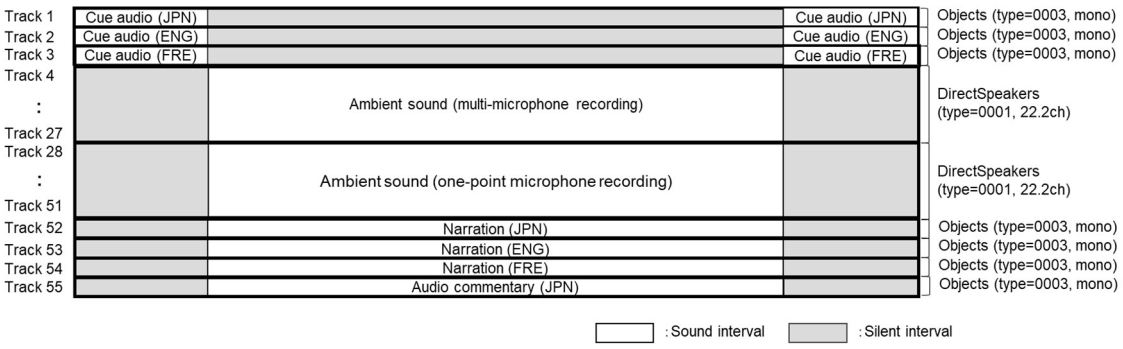


Figure 13 Audio track configuration (Function Check – Amusement Park)

Description:

The test material, “No.3 Function check A) Amusement Park,” simulates an on-site report at an amusement park. This test material contains three switchable audio programs: APR_1001 (“Standard”), APR_1002 (“With Commentary Audio”), and APR_1003 (“Without Narration”). In APR_1001, one object can be selected from each of the following groups: three types of cue audio, two types of ambient sound, and three types of narration, resulting in 18 possible combinations of audio objects. Similarly, APR_1002 allows selection from 2 combinations (different ambient sounds), and APR_1003 allows selection from 6 combinations (cue audio × ambient sound variations). The audio objects AO_1001, AO_1002, and AO_1003 are cue audio spoken respectively by a Japanese female voice, an English male voice, and a French female voice, providing an opening phrase at the start and a closing phrase at the end of the test material. The audio objects AO_1011 and AO_1012 are recordings from the Standard Test Materials Series A. These are channel-based audio objects (type: 0001) representing amusement-park ambience: AO_1011 was recorded using a one-point array of 22 directional microphones, and AO_1012 was recorded using 22 omnidirectional microphones placed at the positions corresponding to the 22 main loudspeakers of the 22.2 multichannel system. These recordings include environmental sounds such as a roller coaster above and a rolling tin can below. This material is therefore suitable for evaluating vertical localization (upper vs. lower movement), as well as perceptual differences caused by the use of different microphone recording systems. The audio objects AO_1021, AO_1022, and AO_1023 are mono audio objects (type: 0003) containing Japanese-female, English-male, and French-female reporters introducing the amusement park. AO_1031 is a mono audio object representing commentary audio for users with visual impairments, spoken by a Japanese male voice describing visual information such as the reporter’s movements and park scenery.

The audio track configuration is as shown in Figure 13 and consists of multiple tracks: Tracks 1–3 contain three mono audio signals corresponding to AO_1001, AO_1002, and AO_1003, respectively; Tracks 4–51 contain two 22.2-multichannel audio signals corresponding to AO_1011 and AO_1012; Tracks 52–54 contain three mono audio signals corresponding to AO_1021, AO_1022, and AO_1023, respectively; and Track 55 contains the mono audio signal corresponding to AO_1031.

This test material is designed to evaluate the following renderer functions:

(1) Simultaneous reproduction of mono audio objects (type: 0003)

In APR_1002 (“With Commentary Audio”), simultaneous reproduction of two mono audio objects can be confirmed: AO_1021 (Japanese narration) and AO_1031 (Japanese commentary audio).

(2) Dialog switching by substitution of mono audio objects (type: 0003)

For example, in APR_1001 (“Standard”), substitution among AO_1021, AO_1022, and AO_1023 can be confirmed. This is achieved by defining AO_1021 as the reference audio object, with AO_1022 and AO_1023 as its alternatives via *audioComplementaryObject*. Initially, the reference object (AO_1021) is reproduced.

(3) Background-sound switching by substitution of channel-based audio objects (type: 0001)

For example, in APR_1001, substitution between the 22.2 channel-based audio objects AO_1011 and AO_1012 can be confirmed. This is achieved by defining AO_1011 as the reference audio object and AO_1012 as its alternative via *audioComplementaryObject*.

(4) Restriction of user control (non-interactive objects)

The audio objects AO_1001, AO_1002, and AO_1003 are non-interactive mono audio objects (type: 0003) designed to emulate emergency-broadcast audio. This is achieved by setting the ADM attribute interact = “0” in *audioObject*.

(5) User interface (UI)

Tags defining selectable presets are specified. For example, if the user selects audioPreset00, the reproduced output corresponds to program APR_1001 with audio objects AO_1001, AO_1011, and AO_1021. This test material is designed to evaluate receiver-side functionality for presenting selectable playback presets to the user, particularly for verifying UI behavior related to additional objects, accessibility functions, and dialog control.

Preset Audio	**Audio Program**	**Audio Object**	**Labels**
audioPreset00: Main Audio (Japanese, Ambient 1)	APR_1001	AO_1001 AO_1011 AO_1021	・ Language: Japanese
audioPreset01: Sub Audio 1 (English, Ambient 1)	APR_1001	AO_1002 AO_1001 AO_1022	・ Language: English

audioPreset02: Sub Audio 2 (French, Ambient 1)	APR_1001	AO_1003 AO_1011 AO_1023	・ Language: French
audioPreset03: Sub Audio 3 (Japanese, Ambient 2)	APR_1001	AO_1001 AO_1012 AO_1021	・ Language: Japanese
audioPreset04: Sub Audio 4 (English, Ambient 2)	APR_1001	AO_1002 AO_1012 AO_1022	・ Language: English
audioPreset05: Sub Audio 5 (French, Ambient 2)	APR_1001	AO_1003 AO_1012 AO_1023	・ Language: French
audioPreset06: Sub Audio 6 (Japanese, Audio commentary, Ambient 1)	APR_1002	AO_1001 AO_1011 AO_1021 AO_1031	・ Language: Japanese ・ Additional Object: Commentary Audio ・ Accessibility: For visually‐impaired users
audioPreset07: Sub Audio 7 (Japanese, Audio commentary, Ambient 2)	APR_1002	AO_1001 AO_1011 AO_1022 AO_1031	・ Language: Japanese ・ Additional Object: Commentary Audio ・ Accessibility: For visually-impaired users
audioPreset08: Sub Audio 8 (Japanese, No narration 、 Ambient 1)	APR_1003	AO_1001 AO_1011	・ Language: Japanese ・ Dialog Control: Muted

audioPreset09: Sub Audio 9 (English, No narration, Ambient 1)	APR_1003	AO_1002 AO_1011	・ Language: English ・ Dialog Control: Muted
audioPreset10: Sub Audio 10 (French, No narration, Ambient 1)	APR_1003	AO_1003 AO_1011	・ Language: French ・ Dialog Control: Muted
audioPreset11: Sub Audio 11 (Japanese, No narration, Ambient 2)	APR_1003	AO_1001 AO_1012	・ Language: Japanese ・ Dialog Control: Muted
audioPreset12: Sub Audio 12 (English, No narration, Ambient 2)	APR_1003	AO_1002 AO_1012	・ Language: English ・ Dialog Control: Muted
audioPreset13: Sub Audio 13 (French, No narration, Ambient 1)	APR_1003	AO_1003 AO_1012	・ Language: French ・ Dialog Control: Muted

B) Volleyball

Files:

```
03_Function
├── 03_B_Function_Volleyball_cartesian
│   ├── 03_B_Function_Volleyball.wav
│   ├── 03_B_Function_Volleyball_cartesian.bw64
│   ├── 03_B_Function_Volleyball_cartesian.csv
│   └── 03_B_Function_Volleyball_cartesian.xml
└── 03_B_Function_Volleyball_polar
    ├── 03_B_Function_Volleyball.wav
    ├── 03_B_Function_Volleyball_polar.bw64
    ├── 03_B_Function_Volleyball_polar.csv
    └── 03_B_Function_Volleyball_polar.xml
```

Duration: 31.6 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: Live Commentary	AO_1001:Cue audio (JPN) AO_1002:Cue audio (ENG)	・ User control not permitted ・ AO_1001 can be substituted with AO_1002
	AO_1011:Ambient Sound (End Line)	
	AO_1021 : Live Commentary (JPN) AO_1022 : Live Commentary (ENG)	・ AO_1021 can be substituted with AO_1022
APR_1002: Court-Side Report	AO_1001:Cue audio (JPN) AO_1002:Cue audio (ENG)	・ User control not permitted ・ AO_1001 can be substituted with AO_1002
	AO_1012 : Ambient Sound (Side Line)	

	AO_1023：Reporter (JPN) AO_1024：Reporter (ENG)	・ AO_1023 can be substituted with AO_1024

Audio track configuration：

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001
2	AO_1002	Objects (mono)	AP_00031002
3-26	AO_1011	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
27-50	AO_1012	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
51	AO_1021	Objects (mono)	AP_00031021
52	AO_1022	Objects (mono)	AP_00031022
53	AO_1023	Objects (mono)	AP_00031023
54	AO_1024	Objects (mono)	AP_00031024

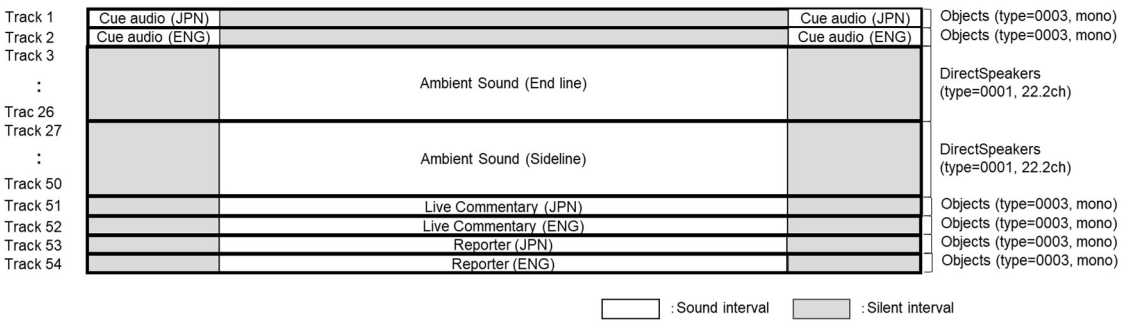


Figure 14 Audio track configuration (Function Check – Volleyball)

Description：

The test material, “No.3 Function Check B) Volleyball,” simulates a sports broadcast for a volleyball game. This test material contains two switchable audio programs: APR_1001 (“Live Commentary”) and APR_1002 (“Court-Side Report”). In APR_1001, one object can be selected from each of the following groups: two types of cue audio, one type of ambient sound, two types of live commentary, resulting in 4 possible combinations of audio objects. Similarly, APR_1002 allows

selection from 4 combinations (cue audio × ambient sound × reporter audio variations). Depending on the selected audio program, the test material allows switching between different ambient sound (recorded on end line vs. sideline), and different spoken content types (live commentary vs. reporter audio). AO_1001 and AO_1002 are non-interactive mono audio objects (type: 0003) designed to emulate emergency-broadcast audio. They are cue audio spoken respectively by a Japanese female voice and an English male voice, providing an opening phrase at the start and a closing phrase at the end of the test material. User control is disabled by setting the ADM audioObject attribute *interact = "0"*. AO_1011 and AO_1012 are 22.2 multichannel channel-based audio objects (type: 0001) recorded for the Standard Test Materials Series A. AO_1011 was recorded using 22 omnidirectional microphones arranged in a cuboid around the volleyball court (end-line forward). AO_1012 was recorded using a one-point microphone placed directly under the net (sideline forward). These sounds include spike impacts, shouts, and gymnasium reverberation; sounds arriving from above and below help evaluate vertical localization and codec-related degradation. For details of microphones and microphone layouts, please refer to the manual of the Standard Test Materials Series A. AO_1021, AO_1022, AO_1023, and AO_1024 are mono audio objects (type: 0003) containing Japanese-male and English-male live commentary, and Japanese-female and English-female reporter audio for one of the teams.

The audio track configuration is as shown in Figure 14 and consists of multiple tracks: Tracks 1–2 contain two mono audio signals corresponding to AO_1001 and AO_1002; Tracks 3–50 contain two 22.2-multichannel audio signals corresponding to AO_1011 and AO_1012; Tracks 51–54 contain four mono audio signals corresponding to AO_1021, AO_1022, AO_1023, and AO_1024.

This test material is designed to evaluate the following renderer functions. Note that user-control restrictions (e.g., non-interactive objects) follow the same rules as in the Amusement Park test materials.

(1) Dialog switching via audio program switching
Switching between APR_1001 and APR_1002 confirms switching between AO_1021 / AO_1022 (for live commentary), and AO_1023 / AO_1024 (for reporter audio).

(2) Background-sound switching via audio program switching
Switching between APR_1001 and APR_1002 confirms switching between AO_1011 (end-line ambient sound), and AO_1012 (sideline ambient sound).

(3) Dialog switching by substituting mono audio objects (type: 0003)

For example, in APR_1001, substitution between AO_1021 and AO_1022 can be confirmed. This is achieved by defining AO_1021 as the reference audio object and AO_1022 as its alternative via audioComplementaryObject.

(4) User interface

Tags are defined to provide selectable presets. This enables evaluation of receiver-side functions that present preset playback patterns to the user, particularly those related to keyword-based filtering. By specifying keywords, user preferences can be reflected in the preset selection.

Preset Audio	**Audio Program**	**Audio Object**	**Labels**
audioPreset00: Main Audio (Japanese)	APR_1001	AO_1001 AO_1011 AO_1021	・ Language: Japanese
audioPreset01: Sub Audio 1 (English)	APR_1001	AO_1002 AO_1001 AO_1022	・ Language: English
audioPreset02: Sub Audio 2 (Japanese, Pitch Report)	APR_1002	AO_1001 AO_1012 AO_1023	・ Language: Japanese ・ Keyword: Team A
audioPreset03: Sub Audio 3 (English, Pitch Report)	APR_1002	AO_1002 AO_1012 AO_1024	・ Language: English ・ Keyword: Team A

C) Train

Files:

```
03_Function
├── 03_C_Function_Train_cartesian
│   ├── 03_C_Function_Train.wav
│   ├── 03_C_Function_Train_cartesian.bw64
│   ├── 03_C_Function_Train_cartesian.csv
│   └── 03_C_Function_Train_cartesian.xml
└── 03_C_Function_Train_polar
    ├── 03_C_Function_Train.wav
    ├── 03_C_Function_Train_polar.bw64
    ├── 03_C_Function_Train_polar.csv
    └── 03_C_Function_Train_polar.xml
```

Duration: 54.5 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: Japanese	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・
	AO_1021: Dialog (JPN)	・
	AO_1031: Sound effects (Road)	・
APR_1002: English	AO_1001: Cue audio (ENG)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・
	AO_1021: Dialog (ENG)	・
	AO_1031: Sound effects (Road)	・
APR_1003: Director's talk	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・

	AO_1021: Director’s talk (JPN)	・
	AO_1031: Sound effects (Road)	・
APR_1004: Japanese (Dialogue-Enhanced)	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・ AVS_1011_0001: Static gain control applied to AO_1011 (−6 dB)
	AO_1021: Dialog (JPN)	
	AO_1031: Sound effects (Road)	・ AVS_1031_0002: Static gain control applied to AO_1031 (−12 dB)
APR_1005: English (Dialogue-Enhanced)	AO_1001: Cue audio (ENG)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・ AVS_1011_0001 : Static gain control applied to AO_1011 (−6 dB)
	AO_1021: Dialog (ENG)	
	AO_1031: Sound effects (Road)	・ AVS_1031_0002: Static gain control applied to AO_1031 (−12 dB)
APR_1006: Director’s talk (Dialogue-Enhanced)	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・ AVS_1011_0001: Static gain control applied to AO_1011 (−6 dB)
	AO_1021: Director’s talk (JPN)	
	AO_1031: Sound effects (Road)	・ AVS_1031_0002: Static gain control applied to AO_1031 (−12 dB)

Audio track configuration:

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1	AO_1001	Objects (mono)	AP_00031001

2	AO_1002	Objects (mono)	AP_00031002
3-14	AO_1011	DirectSpeakers (7.1.4)	AP_00010017 (polar) or AP_00010817 (Cartesian)
15	AO_1021	Objects (mono)	AP_00031021
16	AO_1022	Objects (mono)	AP_00031022
17	AO_1023	Objects (mono)	AP_00031023
18-27	AO_1031	DirectSpeakers (5.1.4)	AP_00010005 (polar) or AP_00010805 (Cartesian)

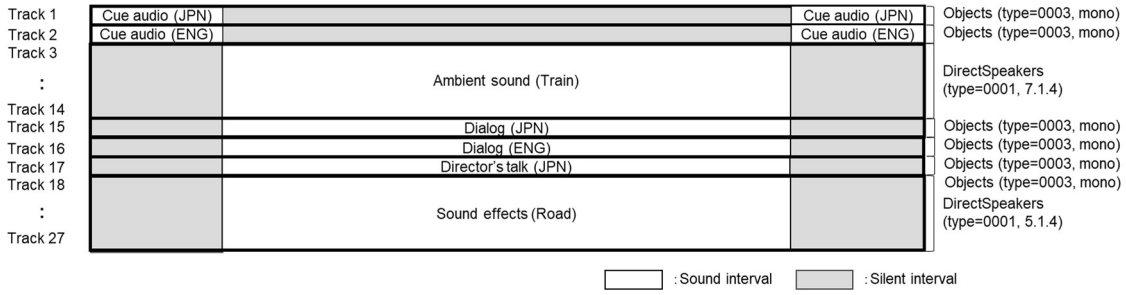


Figure 15 Audio track configuration (Function Check – Train)

Description:

The test material, “No.3 Function Check C) Train,” simulates a drama-style scene. This test material contains six switchable audio programs: APR_1001 (“Japanese”), APR_1002 (“English”), APR_1003 (“Director’s Talk”), APR_1004 (“Japanese (Dialogue-Enhanced)”), APR_1005 (“English (Dialogue-Enhanced)”), and APR_1006 (“Director’s Talk (Dialogue-Enhanced)”). Switching among APR_1001 - APR_1003 selects between Japanese, English, and Director’s Talk (assumed to be commentary by the director on the filming and direction). The audio programs APR_1004 - APR_1006 contain the same base content as APR_1001 - APR_1003, respectively, but with modified reproduction conditions to enhance dialog intelligibility. AO_1001 and AO_1002 are non-interactive mono audio objects (type: 0003) designed to emulate emergency-broadcast audio. They are cue audio spoken respectively by a Japanese female voice (AO_1001) and an English male voice (AO_1002), providing an opening phrase at the start and a closing phrase at the end of the test material. AO_1011 is a channel-based audio object (type: 0001) recorded for the Standard Test Materials Series A. It represents a soundscape recorded along a riverside, where a railway track passes overhead and trains move from left to right. Twenty-two omnidirectional microphones were placed at the positions of the 22 main

loudspeakers of the 22.2 multichannel system. This test material was created using the signals from 12 microphones closest to the 7.1.4 loudspeaker layout. Because the train sound approaches with increasing level, it is suitable for evaluating natural distance perception. Since it uses multi-microphone recording rather than panning, it is also suited for evaluating localization characteristics specific to multi-point capture. AO_1021, AO_1022, and AO_1023 are mono audio objects (type: 0003) containing Japanese-female, English-male, and Japanese-male dialog, respectively. AO_1031 is a channel-based audio object (type: 0001) representing a 5.1.4 ambient soundscape recorded near a road with passing cars and motorbikes. It was created using the signals from 10 microphones positioned near the 5.1.4 loudspeaker locations within a 22-microphone multichannel recording. For details of microphones and microphone layouts, please refer to the manual of the Standard Test Materials Series A.

The audio track configuration is as shown in Figure 14 and consists of multiple tracks: Tracks 1–2 contain two mono audio signals corresponding to AO_1001 and AO_1002; Tracks 3–14 contain a 7.1.4 audio signal corresponding to AO_1011; Tracks 15–17 contain three mono audio signals corresponding to AO_1021, AO_1022, and AO_1023; Tracks 18–27 contain a 5.1.4 audio signal corresponding to AO_1031.

This test material is designed to evaluate the following renderer functions. Note that user-control restrictions (e.g., non-interactive objects) follow the same rules as in the Amusement Park and Volleyball test materials.

(1) Simultaneous reproduction of multiple channel-based audio objects (type: 0001)
Reproducing any of APR_1001 - APR_1006 confirms simultaneous reproduction of AO_1011 (7.1.4) and AO_1031 (5.1.4).

(2) Dialog enhancement through static gain control
Reproducing APR_1004 - APR_1006 confirms that AO_1011 is reproduced at −6 dB and AO_1031 at −12 dB. This is implemented by defining exclusive adjustment values (alternativeValueSet): AO_1011 is attenuated by 6 dB and AO_1031 by 12 dB, respectively.

(3) User Interface
Tags defining selectable presets are provided. This test material is suitable for evaluating receiver-side functions that present selectable playback presets to the user, especially UI behavior related to preset presentation and dialog control. In audioPreset02 and audioPreset05, restrictions may be applied depending on the user’s status (e.g., service subscription).

Preset Audio	Audio Program	Audio Object	Labels
audioPreset00: Main Audio (Japanese)	APR_1001	AO_1001 AO_1011 AO_1021 AO_1031	・ Language: Japanese
audioPreset01: Sub Audio 1 (English)	APR_1002	AO_1002 AO_1001 AO_1022 AO_1031	・ Language: English
audioPreset02: Sub Audio 2 (Japanese, Director's talk, Restricted)	APR_1003	AO_1001 AO_1011 AO_1023 AO_1031	・ Language: Japanese ・ Presentation: Restricted
audioPreset03: Sub Audio 3 (Japanese, Dialogue-Enhanced)	APR_1001	AO_1001 AO_1011 AO_1021 AO_1031	・ Language: Japanese ・ Dialog Control: Enhanced
audioPreset04: Sub Audio 4 (English, Dialogue-Enhanced)	APR_1002	AO_1002 AO_1001 AO_1022 AO_1031	・ Language: English ・ Dialog Control: Enhanced
audioPreset05: Sub Audio 5 (Japanese, Director's talk, Restricted, Dialogue-Enhanced)	APR_1003	AO_1001 AO_1011 AO_1023 AO_1031	・ Language: Japanese ・ Presentation: Restricted ・ Dialog Control: Enhanced

D) Octet

Files:

```
03_Function
├── 03_D_Function_Octet_cartesian
│   ├── 03_D_Function_Octet.wav
│   ├── 03_D_Function_Octet_cartesian.bw64
│   ├── 03_D_Function_Octet_cartesian.csv
│   └── 03_D_Function_Octet_cartesian.xml
└── 03_D_Function_Octet_polar
    ├── 03_D_Function_Octet.wav
    ├── 03_D_Function_Octet_polar.bw64
    ├── 03_D_Function_Octet_polar.csv
    └── 03_D_Function_Octet_polar.xml
```

Duration: 80.3 [s]

Program structure:

Audio program	**Audio object**	**Remarks**
APR_1001: Japanese	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Music (Octet)	・
	AO_1021: Narration (JPN)	・ User-adjustable control range (−12 dB to +6 dB)
APR_1002: English	AO_1002: Cue audio (ENG)	・ User control not permitted
	AO_1011: Music (Octet)	・ GC_1011_0001: Dynamic gain control for English narration
	AO_1022: Narration (ENG)	・ User-adjustable control range (−12 dB to +6 dB)
APR_1003: French	AO_1003: Cue audio (FRE)	・ User control not permitted
	AO_1011: Music (Octet)	・ GC_1011_0002: Dynamic gain control for French narration

	AO_1023：Narration (FRE)	・ User‐adjustable control range (−12 dB to +6 dB)

Audio track configuration：

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001
2	AO_1002	Objects (mono)	AP_00031002
3	AO_1003	Objects (mono)	AP_00031003
4-27	AO_1011	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
28	AO_1021	Objects (mono)	AP_00031021
29	AO_1022	Objects (mono)	AP_00031022
30	AO_1023	Objects (mono)	AP_00031023

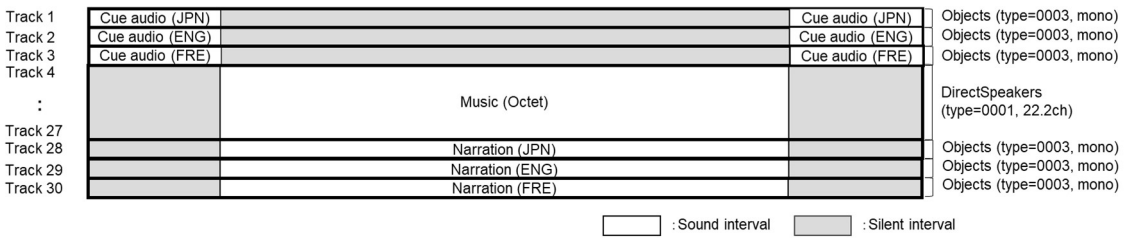


Figure 16 Audio track configuration (Function Check – Octet)

Description：

The test material, “No.3 Function Check D) Octet,” simulates an educational music program. This test material contains three switchable audio programs: APR_1001 (“Japanese”), APR_1002 (“English”), and APR_1003 (“French”). Switching among APR_1001 - APR_1003 enables switching between Japanese, English, and French narration. For APR_1002 and APR_1003, the background sound is dynamically adjusted according to the timing of each narration segment. AO_1001 and AO_1002 are non-interactive mono audio objects (type: 0003) designed to emulate emergency-broadcast audio. They are cue audio spoken respectively by a Japanese female voice (AO_1001), an English male voice (AO_1002), and a French female voice (AO_1003), providing an opening phrase at the start and a closing phrase at the end of the test material. AO_1011 is a channel-based audio object (type: 0001) recorded for the Standard Test Materials Series A. It was recorded using a one-point microphone placed at the center of an octet ensemble, and artificial reverberation

was added to simulate a large concert hall. This object is suitable for evaluating instrument localization, timbre, and envelopment. For details of microphones and microphone layouts, please refer to the manual of the Standard Test Materials Series A. AO_1021, AO_1022, and AO_1023 are mono audio objects (type: 0003) containing narration in Japanese-female, English-male, and French-female narration, respectively.

The audio track configuration is as shown in Figure 16 and consists of multiple tracks: Tracks 1-3 contain three mono audio signals corresponding to AO_1001, AO_1002 and AO_1003; Tracks 4-27 contain a 22.2 audio signal corresponding to AO_1011; Tracks 28-30 contain three mono audio signals corresponding to AO_1021, AO_1022, and AO_1023.

This test material is designed to evaluate the following renderer functions. Note that user-control restrictions for non-interactive objects follow the same rules as in the Amusement Park, Volleyball, Train test materials.

(1) Dialog switching via audio program switching
Switching among APR_1001, APR_1002, and APR_1003 confirms switching among the narration audio objects AO_1021, AO_1022, and AO_1023.

(2) Specification of user-adjustable dialog-level range
For AO_1021, AO_1022, and AO_1023, setting user-adjustable gain limits confirms that the gain can be changed only within the range from 6 dB to −12 dB. This is achieved by defining the maximum and minimum gain values in the ADM element audioObjectInteraction.

(3) Dynamic gain control of audio objects
Switching from APR_1001 to APR_1002 or APR_1003 confirms that the temporal gain behavior applied to the channel-based audio object AO_1011 differs depending on the program. This is achieved by defining two dynamic gain-control patterns: GC_1011_0001 for APR_1002 (English narration) and GC_1011_0002 for APR_1003 (French narration). This enables the background sound level to be adjusted in synchrony with the timing of each narration.

(4) User interface
Tags defining selectable presets are provided. This test material is suitable for evaluating receiver-side functions that present preset playback options to the user, particularly those related to language selection.

Preset Audio	Audio Program	Audio Object	Labels
audioPreset00: Main Audio (Japanese)	APR_1001	AO_1001 AO_1011 AO_1021	・ Language: Japanese
audioPreset01: Sub Audio 1 (English)	APR_1002	AO_1002 AO_1011 AO_1022	・ Language: English
audioPreset02: Sub Audio 2 (French)	APR_1003	AO_1003 AO_1011 AO_1023	・ Language: French

No.4 Simulcast（04_Simul/）

A) Drama - Stereo+5.1+22.2

Files：

```
04_Simul
├── 04_A_Simul_Drama_cartesian
│   ├── 04_A_Simul_Drama.wav
│   ├── 04_A_Simul_Drama_cartesian.bw64
│   ├── 04_A_Simul_Drama_cartesian.csv
│   └── 04_A_Simul_Drama_cartesian.xml
└── 04_A_Simul_Drama_polar
    ├── 04_A_Simul_Drama.wav
    ├── 04_A_Simul_Drama_polar.bw64
    ├── 04_A_Simul_Drama_polar.csv
    └── 04_A_Simul_Drama_polar.xml
```

Duration： 23.2 [s]

Program structure：

Audio program	**Audio object**	**Remarks**
APR_1001：stereo	AO_1011	·
APR_1002：5.1	AO_1012	·
APR_1003：22.2	AO_1013	·

Audio track configuration：

Audio track	**Audio object**		**Audio pack**
	ID	**Type**	
1-2	AO_1011	DirectSpeakers（stereo）	AP_00010002(polar) or AP_00010802(Cartesian)
3-8	AO_1012	DirectSpeakers（5.1）	AP_00010003(polar) or AP_00010803(Cartesian)
9-32	AO_1013	DirectSpeakers (22.2)	AP_00010009(polar)

			or AP_00010809(Cartesian)

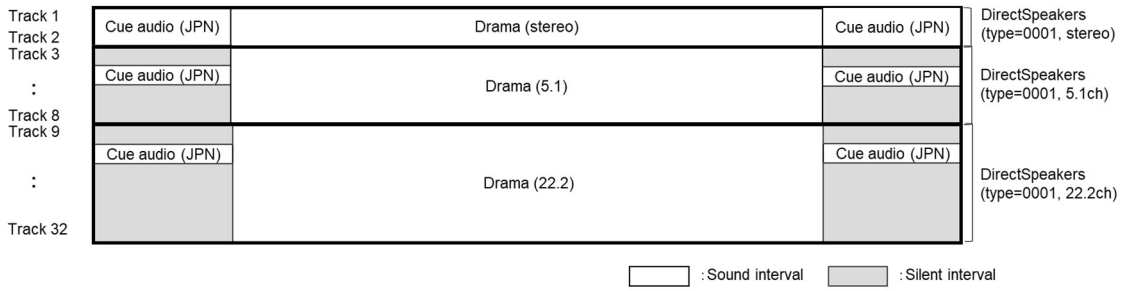


Figure 17 Audio track configuration (Simulcast)

Description:

The test material, “No.4 Simulcast A) Drama,” simulates simulcast broadcasting using stereo, 5.1ch, and 22.2ch. This test material contains three switchable audio programs: APR_1001, APR_1002, and APR_1003. The user can switch among different audio formats to reproduce each channel-based audio object individually. This allows verification of the audio-related metadata used when providing a simulcast broadcast of a conventional channel-based program in an MGA.

The audio object AO_1013 is a recording from the Standard Test Materials Series A. It was recorded using a one-point microphone placed at the center of a meeting room. This recording includes applause, tableware sounds, and voices of speakers located in multiple directions. This test material is suitable for evaluating spatial resolution, instrument/voice directionality, and degradation from audio processing. The flutter echoes following the applause are caused by reflections from walls and glass panes. For details of microphones and microphone layouts, please refer to the manual of the Standard Test Materials Series A. The audio objects AO_1011 (stereo) and AO_1012 (5.1ch) were produced by downmixing AO_1013 (22.2ch).

The audio track configuration is as shown in and consists of multiple tracks: Tracks 1–2 contain one stereo audio signal corresponding to AO_1001; Tracks 3–8 contain a 5.1 audio signal corresponding to AO_1002; Tracks 9–32 contain a 22.2 audio signal corresponding to AO_1003.

No.5 Downmix（05_Downmix/）

A) Channel check (22.2) – coefficient set 1

B) Channel check (22.2) – coefficient set 2

Files：

```
05_Downmix
├── 05_A_Downmix_ChannelCheck_22.2_coefficient1_cartesian
│   ├── 05_A_Downmix_ChannelCheck_22.2_cartesian.csv
│   ├── 05_A_Downmix_ChannelCheck_22.2_cartesian.wav
│   ├── 05_A_Downmix_ChannelCheck_22.2_coefficient1_cartesian.bw64
│   └── 05_A_Downmix_ChannelCheck_22.2_coefficient1_cartesian.xml
├── 05_A_Downmix_ChannelCheck_22.2_coefficient1_polar
│   ├── 05_A_Downmix_ChannelCheck_22.2_coefficient1_polar.bw64
│   ├── 05_A_Downmix_ChannelCheck_22.2_coefficient1_polar.xml
│   ├── 05_A_Downmix_ChannelCheck_22.2_polar.csv
│   └── 05_A_Downmix_ChannelCheck_22.2_polar.wav
├── 05_B_Downmix_ChannelCheck_22.2_coefficient2_cartesian
│   ├── 05_B_Downmix_ChannelCheck_22.2_cartesian.csv
│   ├── 05_B_Downmix_ChannelCheck_22.2_cartesian.wav
│   ├── 05_B_Downmix_ChannelCheck_22.2_coefficient2_cartesian.bw64
│   └── 05_B_Downmix_ChannelCheck_22.2_coefficient2_cartesian.xml
└── 05_B_Downmix_ChannelCheck_22.2_coefficient2_polar
    ├── 05_B_Downmix_ChannelCheck_22.2_coefficient2_polar.bw64
    ├── 05_B_Downmix_ChannelCheck_22.2_coefficient2_polar.xml
    ├── 05_B_Downmix_ChannelCheck_22.2_polar.csv
    └── 05_B_Downmix_ChannelCheck_22.2_polar.wav
```

Duration：（Cartesian） 159 [s] / (polar） 152 [s]

Program structure：

Audio program	**Audio object**	**Remarks**
APR_1001: position check	AO_1001 AO_1002	
APR_1002: channel check	AO_1003	

Audio track configuration：

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001
2-25	AO_1002	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)
26-49	AO_1003	DirectSpeakers (22.2)	AP_00010009 (polar) or AP_00010809 (Cartesian)

Description:

The test material, “No.5 Function Check A) Downmix Coefficients (22.2) – Channel Check,” uses the same audio content as “No.1 Channel Check G) 22.2-1.” The audio track configuration consists of multiple tracks: Tracks 1 contain one mono audio signal corresponding to AO_1001; Tracks 2–49 contain two 22.2 audio signals corresponding to AO_1002 and AO_1003.

Within the audio-related metadata, downmix coefficients are defined for converting the 22.2 multichannel channel-based audio object into the following loudspeaker configurations: 7.1.4, 7.1, 5.1.4, 5.1.2, 5.1, stereo. This enables the user to reproduce audio downmixed using coefficients specified by the content producer, rather than renderer-implemented default coefficients. In the coefficient set 1, the 22.2 audio signal is downmixed according to the following custom rules: (1) the M±060 and FL/FR channels are treated as loudspeakers located at the listener’s sides; (2) within the same layer, the signal is distributed using coefficients that take into account the spatial proximity of nearby loudspeakers; and (3) if loudspeakers in different layers have similar azimuth angles or XY coordinates, the signal is also distributed across layers using coefficients that reflect this proximity. In coefficient set 2, the 22.2 audio signal is downmixed according to the following custom rules: (1) the M±060 and FL/FR channels are treated as loudspeakers located in front of the listener; (2) within the same layer, the signal is evenly distributed to nearby loudspeakers; and (3) the signal is distributed so as to minimize spreading across different layers. Detailed coefficients are provided in Tables 1 and 2.

As in the “No.1 Channel Check” test material, switching between APR_1001 and APR_1002 enables confirmation of the differences in localization of a mono audio object (type: 0003) using panning and a channel-based audio object (type: 0001) using downmixing. In addition, by comparing each coefficient set during the reproduction of APR_1002, differences in reproduction resulting from the defined downmix coefficients can be confirmed. In coefficient set 1, it can be clearly

observed that audio signals located at $\pm 60°$ in the 22.2 multichannel audio are mapped toward left/right front side, while those located at $\pm 90°$ are mapped to the listener's sides. In coefficient set 2, it can also be clearly observed that audio signals located at $\pm 60°$ in 22.2 multichannel audio are mapped toward $\pm 30°$, while those located at $\pm 90°$ are mapped to positions intermediate between $\pm 30°$ and $\pm 110°$, or $\pm 30°$ and $\pm 135°$, depending on the target loudspeaker configuration.

Downmix coefficient set 1 (stereo, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	0.707	0.000	0.707	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.707	0.000	1.000	0.000	0.707	0.500	0.707	0.000	0.707	0.000	0.500	0.707	1.000	0.000
M-030	0.000	0.707	0.707	0.000	0.000	0.707	0.000	1.000	0.500	0.000	0.000	0.707	0.000	1.000	0.707	0.500	0.000	0.707	0.000	0.707	0.500	0.707	0.000	1.000

Downmix coefficient set 1 (5.1, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	0.817	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.577	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.707	0.000	0.000	0.000	1.000	0.000
M-030	0.000	0.817	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.577	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+110	0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.817	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
M-110	0.000	0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.817	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (5.1.2, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	0.817	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.577	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	0.817	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.577	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+110	0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.817	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.577	0.000	0.707	0.000	0.000	0.000
M-110	0.000	0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.817	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.577	0.707	0.000	0.000	0.000
U+030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.817	0.000	0.000	0.000	0.000	0.000
U-030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.817	0.000	0.000	0.000	0.000

Downmix coefficient set 1 (5.1.4, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	0.817	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.577	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	0.817	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.577	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+110	0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.817	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-110	0.000	0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.817	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
U+030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.577	0.000	0.000	0.000	0.000	0.000
U-030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.577	0.000	0.000	0.000	0.000
U+110	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.817	0.000	0.707	0.000	0.000	0.000
U-110	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.817	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (7.1, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+090	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
M-090	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
M+135	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000
M-135	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (7.1.4, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+090	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-090	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+135	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-135	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
U+045	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000
U-045	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000
U+135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
U-135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.0000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (stereo, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	0.707	0.000	0.707	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.707	0.000	1.000	0.000	0.707	0.500	0.707	0.000	0.707	0.000	0.500	0.707	1.000	0.000
FR	0.000	0.707	0.707	0.000	0.000	0.707	0.000	1.000	0.500	0.000	0.000	0.707	0.000	1.000	0.707	0.500	0.000	0.707	0.000	0.707	0.500	0.707	0.000	1.000

Downmix coefficient set 1 (5.1, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.707	0.000	0.000	0.000	1.000	0.000
FR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.707	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
BR	0.000	0.707	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (5.1.2, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.707	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
BR	0.000	0.707	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000
TpFL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000
TpFR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000

Downmix coefficient set 1 (5.1.4, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.707	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BR	0.000	0.707	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TpFL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000
TpFR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000
TpBL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
TpBR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (7.1, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SiL	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
SiR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000

Downmix coefficient set 1 (7.1.4, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SiL	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SiR	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Downmix coefficient set 2 (stereo, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	1.000	0.000	0.707	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.707	0.000	1.000	0.000	0.707	0.500	0.707	0.000	0.707	0.000	0.500	0.707	1.000	0.000
M-030	0.000	1.000	0.707	0.000	0.000	0.707	0.000	1.000	0.500	0.000	0.000	0.707	0.000	1.000	0.707	0.500	0.000	0.707	0.000	0.707	0.500	0.707	0.000	1.000

Downmix coefficient set 2 (5.1, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.707	0.000	0.000	0.000	1.000	0.000
M-030	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+110	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
M-110	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (5.1.2, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+110	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-110	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
U+030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.707	1.000	0.000	1.000	0.000	0.707	0.000	0.000	0.000
U-030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.707	0.000	1.000	0.000	1.000	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (5.1.4, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+110	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-110	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
U+030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000
U-030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000
U+110	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
U-110	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (7.1, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+090	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
M-090	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
M+135	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000
M-135	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (7.1.4, polar)

	M+060	M-060	M+000	LFE1	M+135	M-135	M+030	M-030	M+180	LFE2	M+090	M-090	U+045	U-045	U+000	T+000	U+135	U-135	U+090	U-090	U+180	B+000	B+045	B-045
M+030	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
M-030	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
M+000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+090	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-090	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M+135	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M-135	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
U+045	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000
U-045	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000
U+135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
U-135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (stereo, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	1.000	0.000	0.707	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.707	0.000	1.000	0.000	0.707	0.500	0.707	0.000	0.707	0.000	0.500	0.707	1.000	0.000
FR	0.000	1.000	0.707	0.000	0.000	0.707	0.000	1.000	0.500	0.000	0.000	0.707	0.000	1.000	0.707	0.500	0.000	0.707	0.000	0.707	0.500	0.707	0.000	1.000

Downmix coefficient set 2 (5.1, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.707	0.000	0.000	0.000	1.000	0.000
FR	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (5.1.2, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TpFL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.707	1.000	0.000	1.000	0.000	0.707	0.000	0.000	0.000
TpFR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.707	0.000	1.000	0.000	1.000	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (5.1.4, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TpFL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.707	0.500	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000
TpFR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.707	0.500	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000
TpBL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.707	0.000	0.707	0.000	0.000	0.000
TpBR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.707	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (7.1, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SiL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
SiR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000

Downmix coefficient set 2 (7.1.4, Cartesian)

	FL	FR	FC	LFE1	BL	BR	FLc	FRc	BC	LFE2	SiL	SiR	TpFL	TpFR	TpFC	TpC	TpBL	TpBR	TpSiL	TpSiR	TpBC	BtFL	BtFR	BtFC
FL	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
FR	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
FC	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
LFE	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SiL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SiR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BL	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BR	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.707	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

No.6 Loudness （06_Loudness/）

A) Train – Measurement (Baseline +3dB)

B) Train – Measurement (Baseline +0 dB）

C) Train – Measurement (Baseline -3 dB）

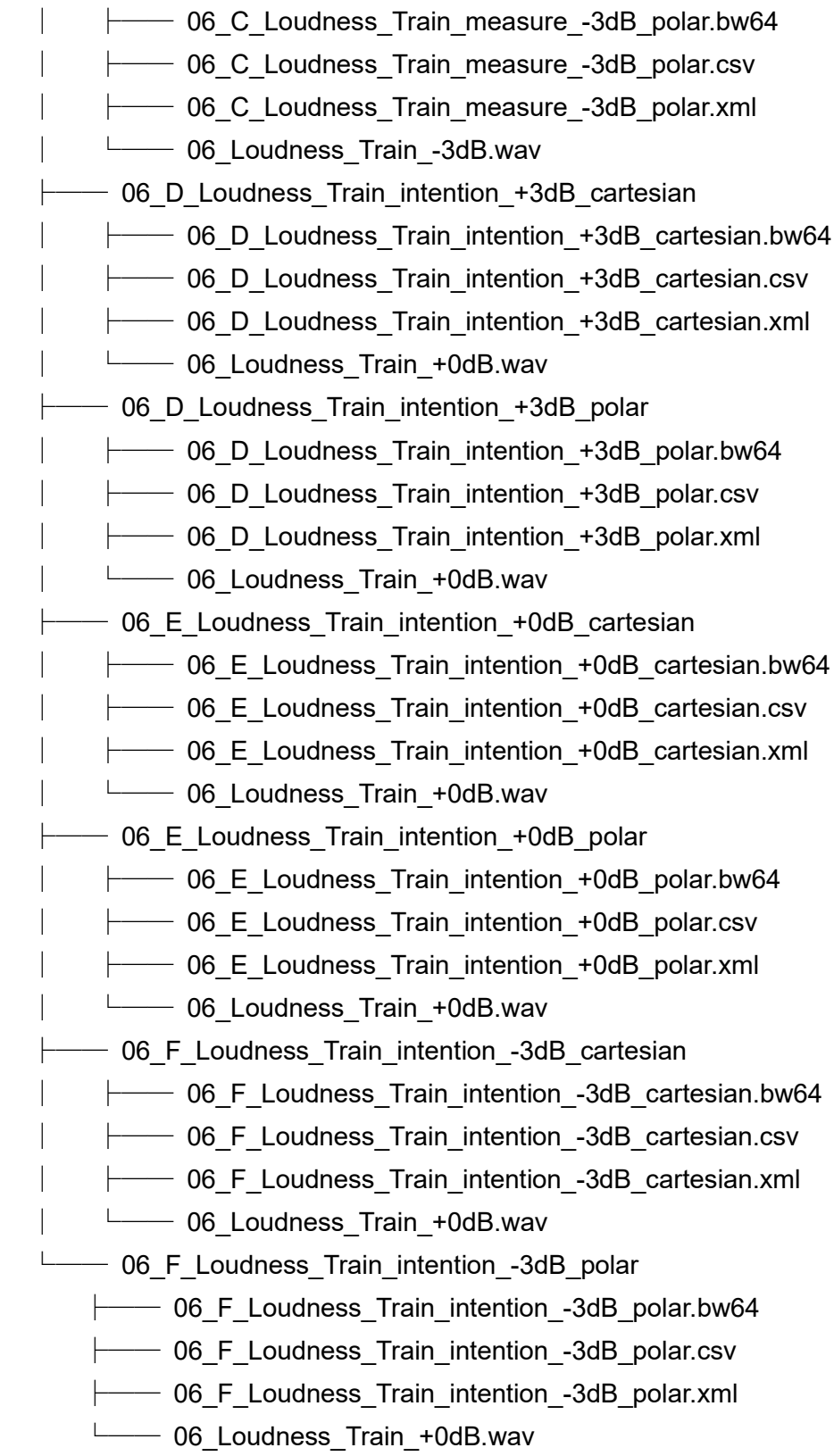
D) Train – Creator’s intent (Baseline +3 dB）

E) Train – Creator’s intent (Baseline +0 dB）

F) Train – Creator’s intent (Baseline -3 dB）

Files：

```
06_Loudness
├── 06_A_Loudness_Train_measure_+3dB_cartesian
│   ├── 06_A_Loudness_Train_measure_+3dB_cartesian.bw64
│   ├── 06_A_Loudness_Train_measure_+3dB_cartesian.csv
│   ├── 06_A_Loudness_Train_measure_+3dB_cartesian.xml
│   └── 06_Loudness_Train_+3dB.wav
├── 06_A_Loudness_Train_measure_+3dB_polar
│   ├── 06_A_Loudness_Train_measure_+3dB_polar.bw64
│   ├── 06_A_Loudness_Train_measure_+3dB_polar.csv
│   ├── 06_A_Loudness_Train_measure_+3dB_polar.xml
│   └── 06_Loudness_Train_+3dB.wav
├── 06_B_Loudness_Train_measure_+0dB_cartesian
│   ├── 06_B_Loudness_Train_measure_+0dB_cartesian.bw64
│   ├── 06_B_Loudness_Train_measure_+0dB_cartesian.csv
│   ├── 06_B_Loudness_Train_measure_+0dB_cartesian.xml
│   └── 06_Loudness_Train_+0dB.wav
├── 06_B_Loudness_Train_measure_+0dB_polar
│   ├── 06_B_Loudness_Train_measure_+0dB_polar.bw64
│   ├── 06_B_Loudness_Train_measure_+0dB_polar.csv
│   ├── 06_B_Loudness_Train_measure_+0dB_polar.xml
│   └── 06_Loudness_Train_+0dB.wav
├── 06_C_Loudness_Train_measure_-3dB_cartesian
│   ├── 06_C_Loudness_Train_measure_-3dB_cartesian.bw64
│   ├── 06_C_Loudness_Train_measure_-3dB_cartesian.csv
│   ├── 06_C_Loudness_Train_measure_-3dB_cartesian.xml
│   └── 06_Loudness_Train_-3dB.wav
├── 06_C_Loudness_Train_measure_-3dB_polar
```



Duration: 54.5 [s]

Program structure:

Audio program	Audio object	Remarks
APR_1001: Japanese	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・
	AO_1021: Dialog (JPN)	・
	AO_1031: Sound effects (Road)	・
APR_1002: English	AO_1001: Cue audio (ENG)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・
	AO_1021: Dialog (ENG)	・
	AO_1031: Sound effects (Road)	・
APR_1003: Director’s talk	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・
	AO_1021: Director’s talk (JPN)	・
	AO_1031: Sound effects (Road)	・
APR_1004: Japanese (Dialogue-Enhanced)	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・ AVS_1011_0001: Static gain control applied to AO_1011 (−6 dB)
	AO_1021: Dialog (JPN)	
	AO_1031: Sound effects (Road)	・ AVS_1031_0002: Static gain control applied to AO_1031 (−12 dB)
APR_1005:	AO_1001: Cue audio (ENG)	・ User control not permitted

English (Dialogue-Enhanced)	AO_1011: Ambient sound (Train)	・ AVS_1011_0001: Static gain control applied to AO_1011 (−6 dB)
	AO_1021: Dialog (ENG)	
	AO_1031: Sound effects (Road)	・ AVS_1031_0002: Static gain control applied to AO_1031 (−12 dB)
APR_1006: Director's talk (Dialogue-Enhanced)	AO_1001: Cue audio (JPN)	・ User control not permitted
	AO_1011: Ambient sound (Train)	・ AVS_1011_0001: Static gain control applied to AO_1011 (−6 dB)
	AO_1021: Director's talk (JPN)	
	AO_1031: Sound effects (Road)	・ AVS_1031_0002: Static gain control applied to AO_1031 (−12 dB)

Audio track configuration:

Audio track	Audio object		Audio pack
	ID	Type	
1	AO_1001	Objects (mono)	AP_00031001
2	AO_1002	Objects (mono)	AP_00031002
3-14	AO_1011	DirectSpeakers (7.1.4)	AP_00010017 (polar) or AP_00010817 (Cartesian)
15	AO_1021	Objects (mono)	AP_00031021
16	AO_1022	Objects (mono)	AP_00031022
17	AO_1023	Objects (mono)	AP_00031023
18-27	AO_1031	DirectSpeakers (5.1.4)	AP_00010005 (polar) or AP_00010805 (Cartesian)

Description:

The test materials, "No.6 Loudness A) Train – Measured Value (Baseline +3 dB)" through "No.6 Loudness F) Train – Creators' Intended Value (Baseline −3 dB)," use the same configuration as "No.3 Function Check C) Train." The audio track configuration is as shown in Figure 15 and consists of multiple tracks: Tracks 1-2 contain two mono audio signals corresponding to AO_1001 and AO_1002; Tracks

3-14 contain a 7.1.4 audio signal corresponding to AO_1011; Tracks 15-17 contain three mono audio signals corresponding to AO_1021, AO_1022, and AO_1023; Tracks 18-27 contain a 5.1.4 audio signal corresponding to AO_1031.

The ADMs allow specification of the loudness value of program audio using the loudnessMetadata element. In the ARIB Technical Document TR-B48, the loudness value may be specified either as a measured value (loudnessMethod="ITU-R BS.1770"), obtained using the method defined in ITU-R BS.1770[35], or as a creator's intended value[36] (loudnessMethod="Creator's Intention"). In both cases, loudness correction or normalization is performed based on the value specified in the loudness metadata. However, in Japan, from the perspective of program quality management, it is assumed that creators' intended target values will be described in the loudness metadata at the program delivery. This is because, when measured values are specified, the margin is set by broadcasting guideline, and thus the reproduction level at the production may not match the default reproduction level at the receiver. In addition, in live broadcasts such as sports programs, it is not possible to determine the overall program loudness in advance. Therefore, specifying measured loudness is not feasible.

This test material is designed to evaluate differences in level adjustment resulting from loudness correction and normalization functions when different measured values and target values are specified in the loudness metadata. For A, B, and C, the audio signal levels are adjusted to +3 dB, +0 dB, and −3 dB, respectively, relative to "No.3 Function Check C) Train" and measured values (loudnessMethod="ITU-R_BS.1770") are specified in the loudness metadata. For D, E, and F, the audio signal levels identical (+0 dB) to those of "No.3 Functional Evaluation C) Train," while the loudness metadata specifies creator's intended values (loudnessMethod= "Creator's Intention") of +3 dB, +0 dB, and −3 dB relative to the broadcast reference level of −24 LKFS, respectively. For receivers equipped with loudness correction functionality, in cases A, B and C, the measured values specified in the loudness metadata match the actual program loudness. It can therefore be confirmed that the audio is adjusted to the reference audio level (−24 LKFS in broadcasting) based on the specified measured values. In cases D, E, and F, the actual program loudness is approximately −24 LKFS, while the loudness metadata specifies target values of −27 LKFS, −24 LKFS, and −21 LKFS, respectively. Based on these specified target

[35] Recommendation ITU-R BS.1770-5: Algorithms to measure audio programme loudness and true-peak audio level, 2023.

[36] Creator's intended value: A loudness value set according to production intent, regardless of the actual measured loudness of the program audio.

values, it can be confirmed that the level of the rendered audio signal is adjusted to approximately −3 dB, approximately +0 dB, and approximately +3 dB, respectively.