International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Report ITU-R BS.2159

(11/2009)

Multichannel sound technology in home and broadcasting applications

BS Series

Broadcasting service (sound)

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at http://www.itu.int/publ/R-REP/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	**Broadcasting service (sound)**
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management

Note: *This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2010

REPORT ITU-R BS.2159

Multichannel sound technology in home and broadcasting applications

(2009)

1 Introduction

ITU-R has developed Recommendation ITU-R BS.775 for multichannel stereophonic sound system with and without accompanying picture. Multichannel stereo as well as 2-channel stereo audio services are widely used as part of digital broadcasting services. Recommendation ITU-R BS.775 specifies a hierarchy of compatible multichannel sound systems to enhance the directional stability of the frontal sound image and the sensation of spatial reality (ambience), and each loudspeaker is set at the same height as a listener's ears.

Some television applications with higher resolution imagery and large screen digital imaginary (LSDI) application , both providing wider viewing angle, may need multichannel stereophonic sound systems that can reproduce the sound sources, which are localized at a higher position over the listener and a lower position below the screen, and vertical movements of the sound sources. Several multichannel stereophonic sound systems are currently applied or studied for higher resolution imagery, and some of them have loudspeakers arranged above and below the viewer.

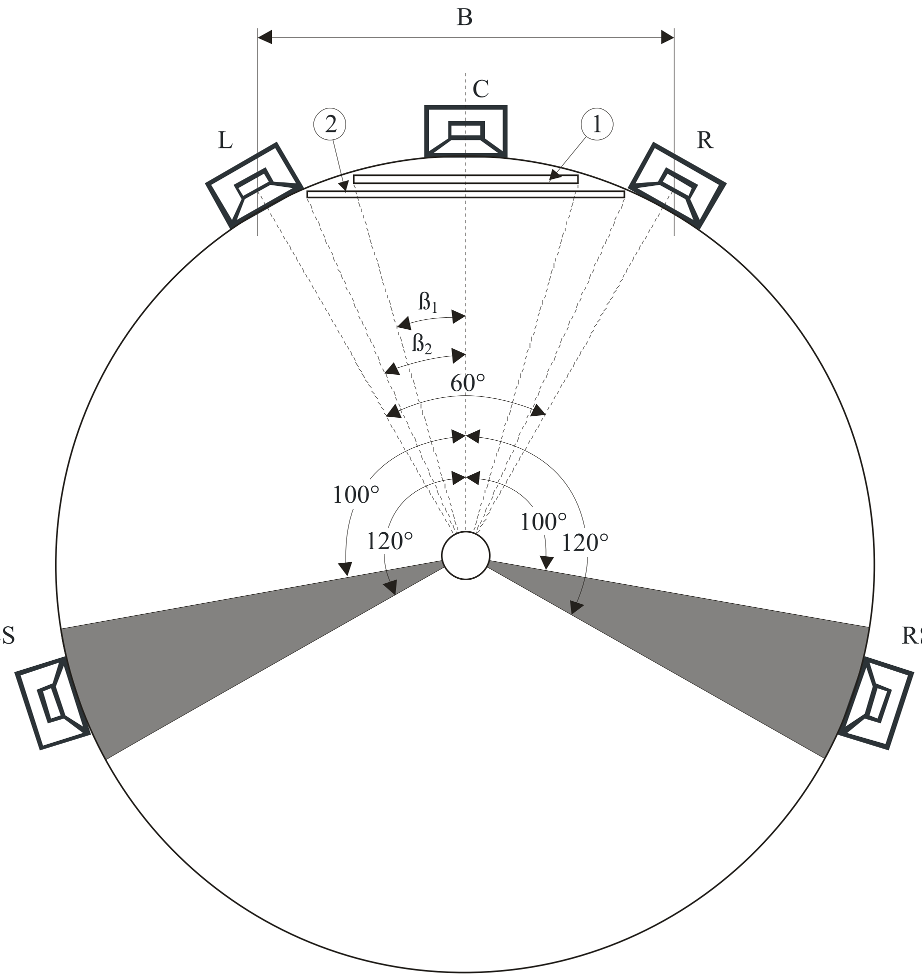
There would be value in continued studies in this area for future broadcasting applications in order to evolve beyond the current 5.1 channel sound system. This Report contains information on the subject of – Multichannel sound technology in home and broadcasting applications, beyond the current 5.1 channel sound system specified in Recommendation ITU-R BS.775.

2 5.1 multichannel sound system

The 5.1 channel sound system has been specified in Recommendation ITU-R BS.775. The system is widely used as a part of digital broadcasting services. It enhances the directional stability of the frontal sound image and the sensation of spatial reality (ambience). The reference loudspeaker arrangement is shown in Fig. 1, in which each loudspeaker is set at the same height as a listener's ears.

FIGURE 1

Reference loudspeaker arrangement with loudspeakers L/C/R and LS/RS



Screen 1 HDTV - Reference distance = 3 H ($2ß_1 = 33°$)

Screen 2 = 2 H ($2ß_2 = 48°$)

H: height of screen

B: loudspeaker base width

Loudspeaker	Horizontal angle from centre (degrees)	Height (meters)	Inclination (degrees)
C	0	1.2	0
L, R	30	1.2	0
LS, RS	100 ... 120	≥ 1.2	0 ... 15 down

Report 2159-01

3 Basic requirements of multichannel sound systems beyond the 5.1 sound system

The following requirements are related to the multichannel sound system beyond the current 5.1 channel sound system specified in Recommendation ITU-R BS.775.

1 The directional stability of the frontal sound image should be maintained over the entire higher resolution imagery area. Coincidence of position between sound image and video image also should be maintained over a wide imagery area.

2 The sound image should be reproduced in all directions around the listener, including elevation.

3 The sensation of three-dimensional spatial impression that augments a sense of reality should be significantly enhanced. This may be achieved by the use of side and/or back, top and/or bottom loudspeakers.

4 Exceptional sound quality should be maintained over wider listening area than that provided by current 5.1 channel sound system.

5 Compatibility with the current 5.1 channel sound system specified in Recommendation ITU-R BS.775 should be ensured to an acceptable degree.

6 Live recording, mixing and transmission should be possible.

4 Multichannel sound systems beyond the 5.1 sound system under development for broadcasting applications

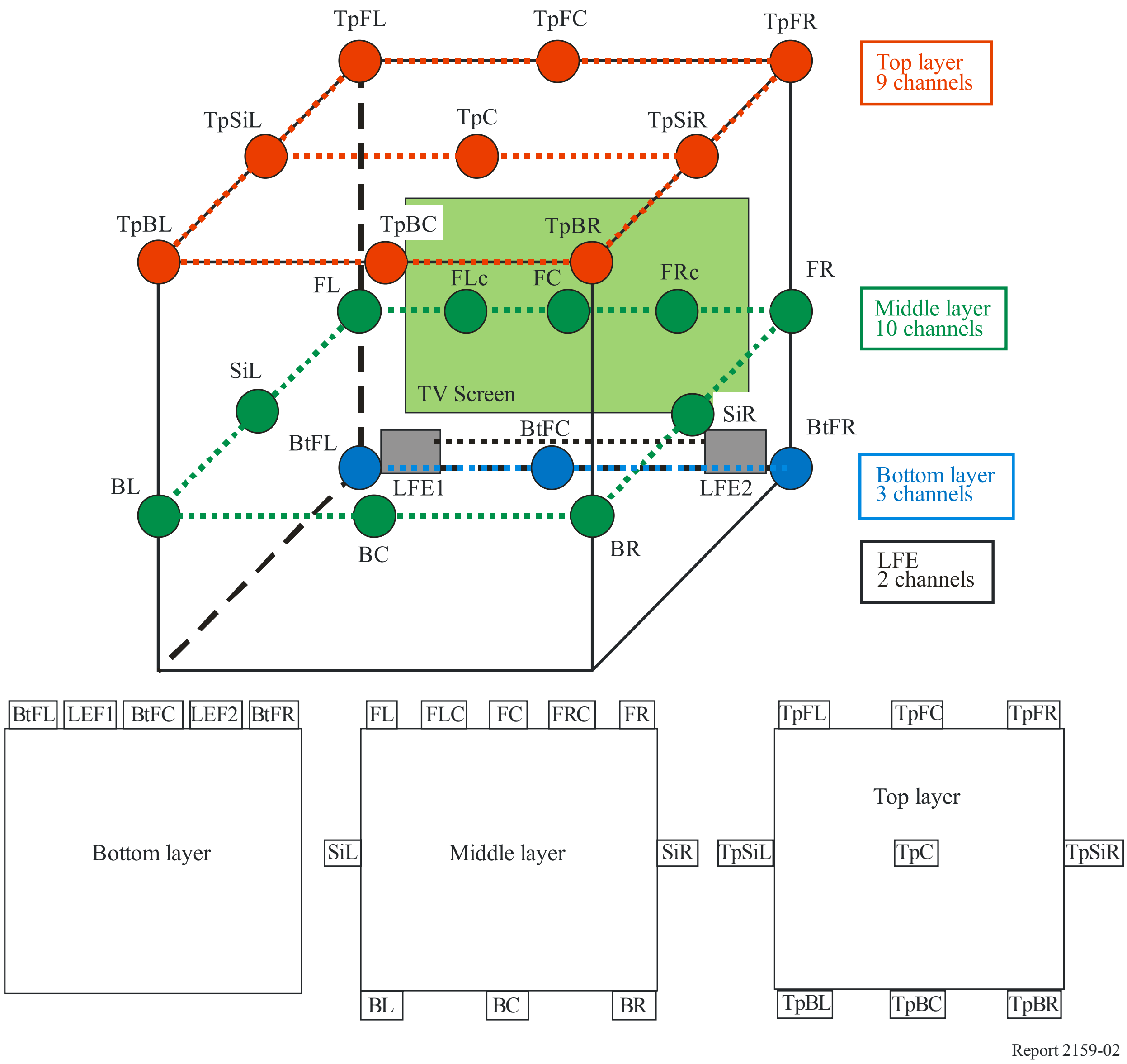
Several multichannel sound systems have been studied to improve the spatial impression of sound. The following systems seem to have the capability for practical use.

4.1 22.2 multichannel sound system

The 22.2 multichannel sound system was developed by NHK (Japan Broadcasting Corporation). It has nine channels at the top layer, ten channels at the middle layer, three channels at the bottom layer and two low frequency effects (LFE) channels. This system is suited to wide screens such as 100 inch FPD display, because it can localize two-dimensionally a sound image over the entire screen by using three bottom channels and five middle channels and three top channels around the screen.

FIGURE 2

The 22.2 multichannel sound system



This system has common channels at each three layers with other multichannel systems so that its audio can be easily down-mixed to other multichannel sound systems and has compatibility with every multichannel sound system.

Audio characteristics and audio channel mapping for UHDTV programme production for the 22.2 multichannel sound format has been standardized by SMPTE (SMPTE 2036-2), as mentioned in § 8.1.1.

Several sound reproduction systems based on the 22.2 multichannel sound have been developed and exhibited at international expositions and exhibitions. A large theatrical sound system with a 600 inch screen was exhibited at World Exposition at Aichi, Japan in 2005 (see Fig. 3).

FIGURE 3

Theatrical 22.2 multichannel sound system at World Exposition held at Aichi, Japan in 2005

Report 2159-03

Home sound systems for 22.2 multichannel sound have also been developed. Figure 4 shows the home 22.2 sound system using multiple compact loudspeakers.

FIGURE 4

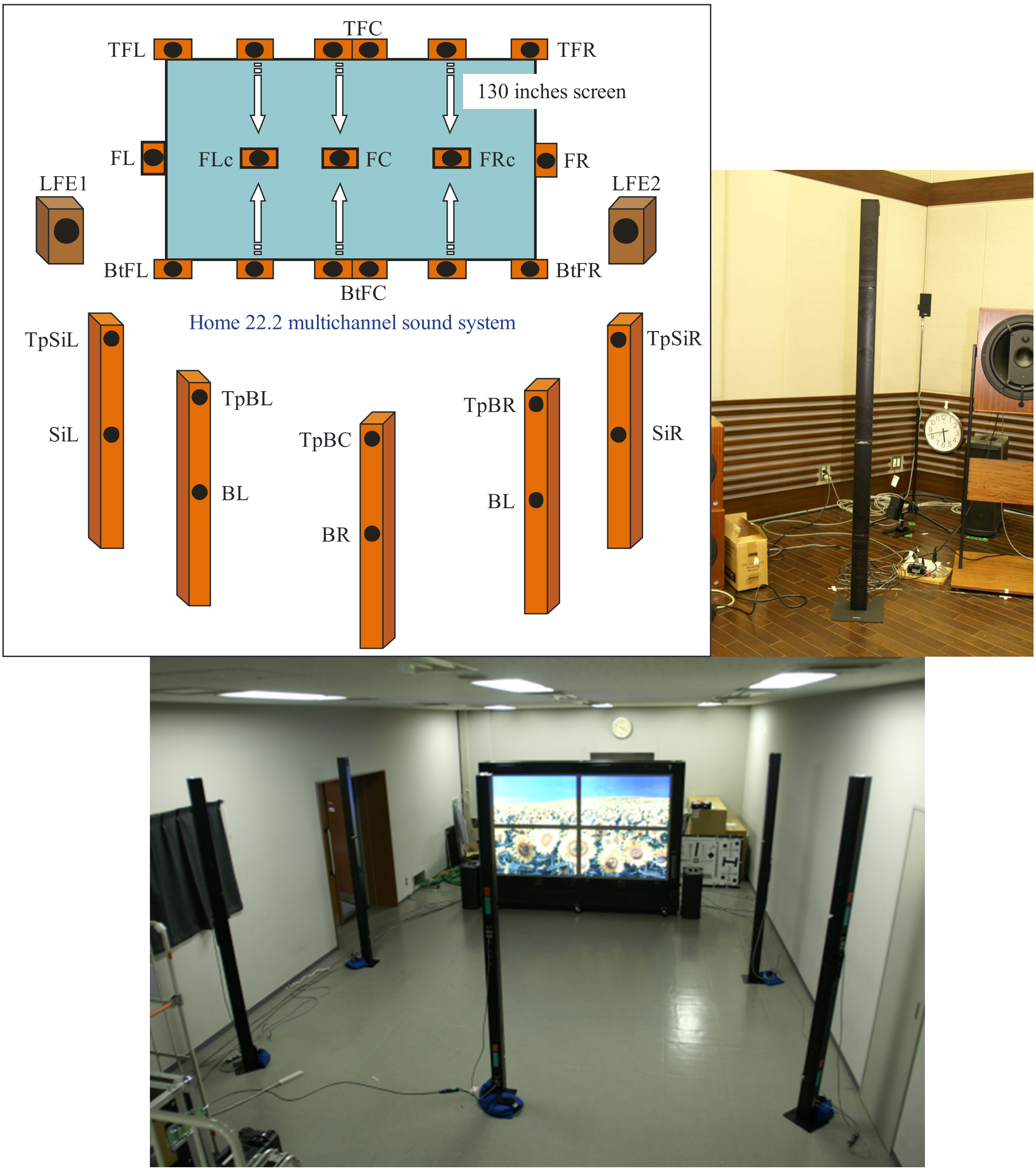
Home 22.2 multichannel sound system using multiple compact loudspeakers

Report 2159-04

A tallboy type loudspeaker has been developed to reproduce three vertical channels (i.e. top, middle and bottom channels) by a single loudspeaker. These loudspeakers are used for the home 22.2 multichannel sound system with UHDTV FPD on which compact loudspeaker units are rigged to reproduce frontal sound channels as Fig. 5.

FIGURE 5

Home 22.2 multichannel sound system using tallboy type loudspeakers



Report 2159-05

A headphone processor to provide 22.2 multichannel sound has been developed; it is shown in Fig. 6. This processor enables listeners to enjoy an accurate immersive 3D sound with ordinary headphones. Because the headphone processor can reproduce the 22.2 multichannel sound without the use of loudspeakers, TV programmes with 3D sound can be efficiently produced on location in places such as a broadcast OB van.

FIGURE 6

22.2 multichannel sound headphone processor

Report 2159-06

4.2 10.2 surround sound system

The 10.2 surround sound format was developed by Tomlinson Holman of TMH Labs. This is actually a 12.2 system that uses a total of 14 channels.

The 14 discrete channels are:

– Five front speakers: left wide, left, centre, right and right wide (in Fig. 7, these are described as left, left c., centre, right c., right).

– Five surround channels: left surround diffuse, left surround direct, back surround, right surround diffuse and right surround direct (in Fig. 7, these are described as left surr., left point surr., centre surround, right point surr., right surr.).

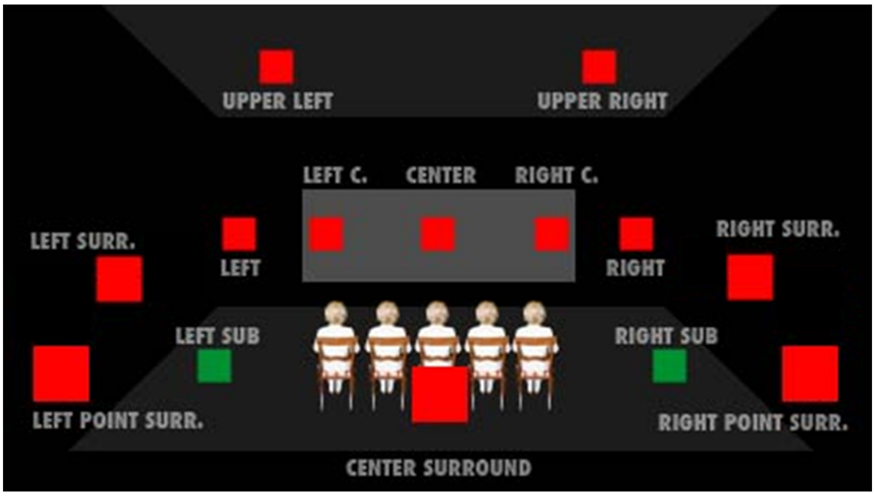
– Two LFE channels: LFE left, LFE right (in Fig. 7, these are described as left sub and right sub).

– Two height channels: left height, right height (in Fig. 7, these are described as upper left and upper right).

The .2 of the 10.2 refers to the addition of a second subwoofer for LFE. Although low frequencies are not localizable, it was found that splitting the bass on either side of the audience increases the sense of envelopment.

FIGURE 7

10.2 surround sound



Report 2159-07

4.3 Wave-field-synthesis

Wave-field-synthesis (WFS) was invented by the Delft University of Technology, Netherlands in 1989. In the European project CARROUSO components for the complete chain including recording, coding, transmission, decoding, and sound reproduction were developed. Since then WFS has been refined to deliver truly immersive sound. Application areas include cinema (with a priority on combination with 3D video), theme parks, virtual reality (VR) installations (in combination with 3D audio) and, in the long run, home theatres. In February 2003 the first cinema using this system started daily operation (Ilmenau, Germany). In 2004 the first WFS system was installed in a sound stage in Studio City, CA. Since 2008, the Chinese 6 Theatre and the Museum of Tolerance in Los Angeles have been equipped with WFS sound systems. These systems are also used in themed environments. Commercial examples of IOSONO GmbH (a spinoff of Fraunhofer IDMT) include the installation in the 4D cinema at the Bavaria Filmstadt (Munich), the Odysseum Science Adventure Park (Cologne), and the “Haunted Mansion” at Disney World (Orlando). Virtual reality installations at the University of Surrey and the Technical University of Ilmenau use WFS with two loudspeaker arrays in the front to enable the proper reproduction of elevation. These two systems also use stereoscopic video projection. An extension of WFS with additional loudspeakers above the listeners was presented at “the 2008 Expo” in Saragossa.

FIGURE 8

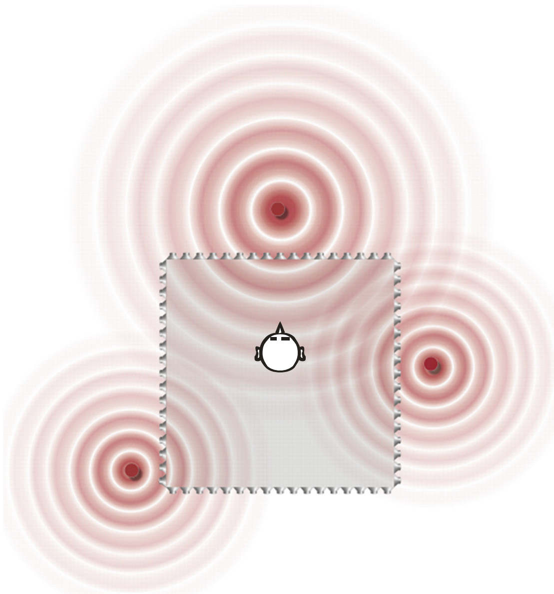
WFS sound system in the German cinema "Linden lichtspiele ilmenau"

Report 2159-08

WFS is an object-oriented approach to accurately recreate a replication of a sound field using the theory of waves and of the generation of wave fronts. This concept is best explained by the well known Huygens principle: points on a wave front serve as individual point sources of spherical secondary waves. This principle is applied in acoustics by using a large number of small and closely spaced loudspeakers (loudspeaker arrays). The driving signal is calculated for each of the loudspeakers in real time at the reproduction site. The number of loudspeakers is independent from the number of transmission channels and only related to the size of the reproduction room. Loudspeaker arrays controlled by WFS reproduce wave fields that originate from any combination of (virtual) sound sources like an acoustic hologram. When manipulated properly, the system recreates wave fronts approaching perfect temporal, spectral and spatial properties throughout the listening room.

FIGURE 9

Working principle of WFS

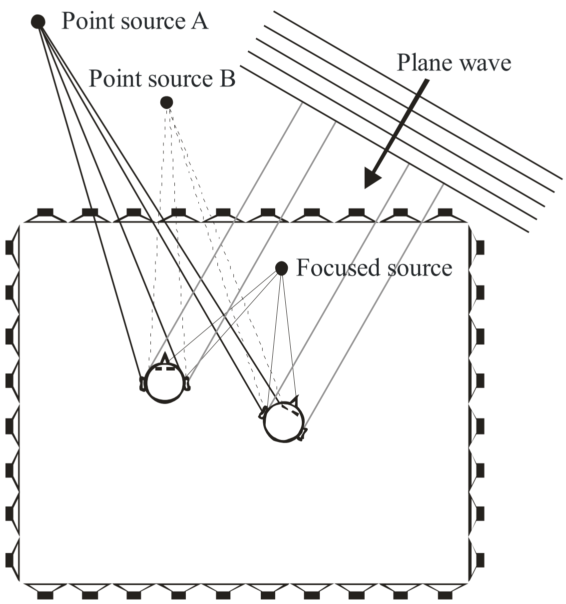


Report 2159-09

Three representations of sound sources are possible in WFS (see Fig. 10). In the first two, virtual sound sources can be placed behind the loudspeaker arrays (so-called point sources) as well as in front of loudspeaker arrays (so-called focused sources). In the case of sound sources in front of the array, the array radiates convergent waves toward a focus point, from which divergent waves propagate into the listening area. The third type is the so-called plane waves. Plane waves come from the same angular direction for all positions in the listening space.

FIGURE 10

Reproduction of point sources, focused sources and plane waves



Report 2159-10

The other commonly used channel-oriented sound reproduction approaches require a well-defined loudspeaker setup, i.e., the number and positions of the loudspeakers are predefined. In the mastering process the target setup must be determined and the loudspeaker signals prepared in a way that allows them to they perfectly fit the assumed setup. This implies that it is difficult to feed the generated signals into another sound system.

This problem can be solved by the object-oriented sound reproduction paradigm, which was develop for WFS but which is not restricted to it. In this method, the audio content is represented as audio objects containing the pure audio content together with metadata describing the position of the object in real time along with the properties of the audio object like directivity. On the rendering site the driving signal for each individual loudspeaker is calculated taking into account its exact position in the reproduction room. Besides the positioning of direct sound, a position-dependent calculation of early reflections and diffuse reverberations is possible, which enables the generation of realistic but also artificial spatial environments. Through the availability of the direct sound of each source and a parametric description of the properties of the room, an optimal reproduction can be adapted to the given spatial environment. This can be a WFS setup of any size (and number of loudspeakers) but also an arbitrary loudspeaker configuration. Increasing the number of loudspeakers increases the size of the sweet spot and makes the sound sources more stable. This results in an increased freedom when deciding which loudspeaker setup to install, because the actual loudspeaker signals are calculated at the reproduction site through a process called rendering.

WFS overcomes the restrictions of a sweet-spot and enable the location of sound objects at any position outside and inside the reproduction room without problems of phase or sound coloration. All formats mentioned in §§ 4.1, 4.2 and 5 can be reproduced using WFS by the concept of virtual loudspeakers enabling an enlarged sweet-spot for any content already produced.

5 Multichannel sound systems in use for home audio release media

The following multichannel sound systems are used in home audio entertainment.

5.1 DVD audio

DVD audio is a digital format for delivering exceptionally high-fidelity audio content on a DVD. It offers many channel configurations of audio channels, ranging from mono to 5.1-channel surround sound, at various sampling frequencies and bit resolution per sample (from compact disc 44.1 kHz/16 bits up to 192 kHz/24 bits). Compared with the CD format, the much higher capacity DVD format enables the inclusion of considerably more music (with respect to total running time and quantity of songs) and/or far higher audio quality (reflected by higher sampling frequencies and greater bit resolution per sample, and/or additional channels for spatial sound reproduction).

Audio is stored on the disc in linear pulse code modulation (PCM) format, which is either uncompressed or losslessly compressed with Meridian Lossless Packing (MLP). The maximum permissible total bit rate is 9.6 Mbit/s. In uncompressed modes, it is possible to get up to 96 kHz/16 bits or 48 kHz/24 bits in 5.1-channel surround sound. To store 5.1-channel surround sound tracks in 88.2 kHz/20 bits, 88.2 kHz/24 bits, 96 kHz/20 bits or 96 kHz/24 bits MLP encoding is mandatory.

5.2 SACD

Super Audio CD (SACD) is a read-only optical audio disc format that provides higher fidelity digital audio reproduction. SACD audio is stored in a format called Direct Stream Digital (DSD), which differs from the conventional PCM used by compact disc or conventional computer audio systems. DSD is 1-bit and has a sampling frequency of 2.8224 MHz. This gives the format a greater dynamic range and wider frequency response than that of the CD. The system is capable of delivering a dynamic range of 120 dB from 20 Hz to 20 kHz and an extended frequency response up to 100 kHz.

SACD supports up to six channels at full bandwidth. In its current form the SACD standard does not precisely specify how the channels shall be used.

222 sound currently uses SACD to provide 2 + 2 + 2 sound contents consist of 6 channels including 4 channels (front left, front right, rear left and rear right) and 2 height channels (top front left and top front right).

5.3 BD

BD is an optical disc format. The format was developed to enable recording, rewriting and playback of high-definition (HD) video, as well as storing large amounts of data. BD pre-recorded application format (BD-ROM) is designed not only for pre-packaged HD movie content but also as a key component of a consumer HD platform. The BD platform is designed to provide access to HD content throughout the home via HD digital broadcast recording and HD playback functions.

One of the key features offered by BD-ROM is

– Industry standard high definition video and surround sound audio:
  – MPEG-2, MPEG-4 AVC, and SMPTE VC-1 video formats;
  – LPCM as well as Dolby Digital, Dolby Digital Plus, Dolby Lossless, DTS digital surround, and DTS-HD audio formats.

BD-ROM supports six types of audio stream formats ranging from a low bit rate to high audio quality, as shown in Table 1.

TABLE 1

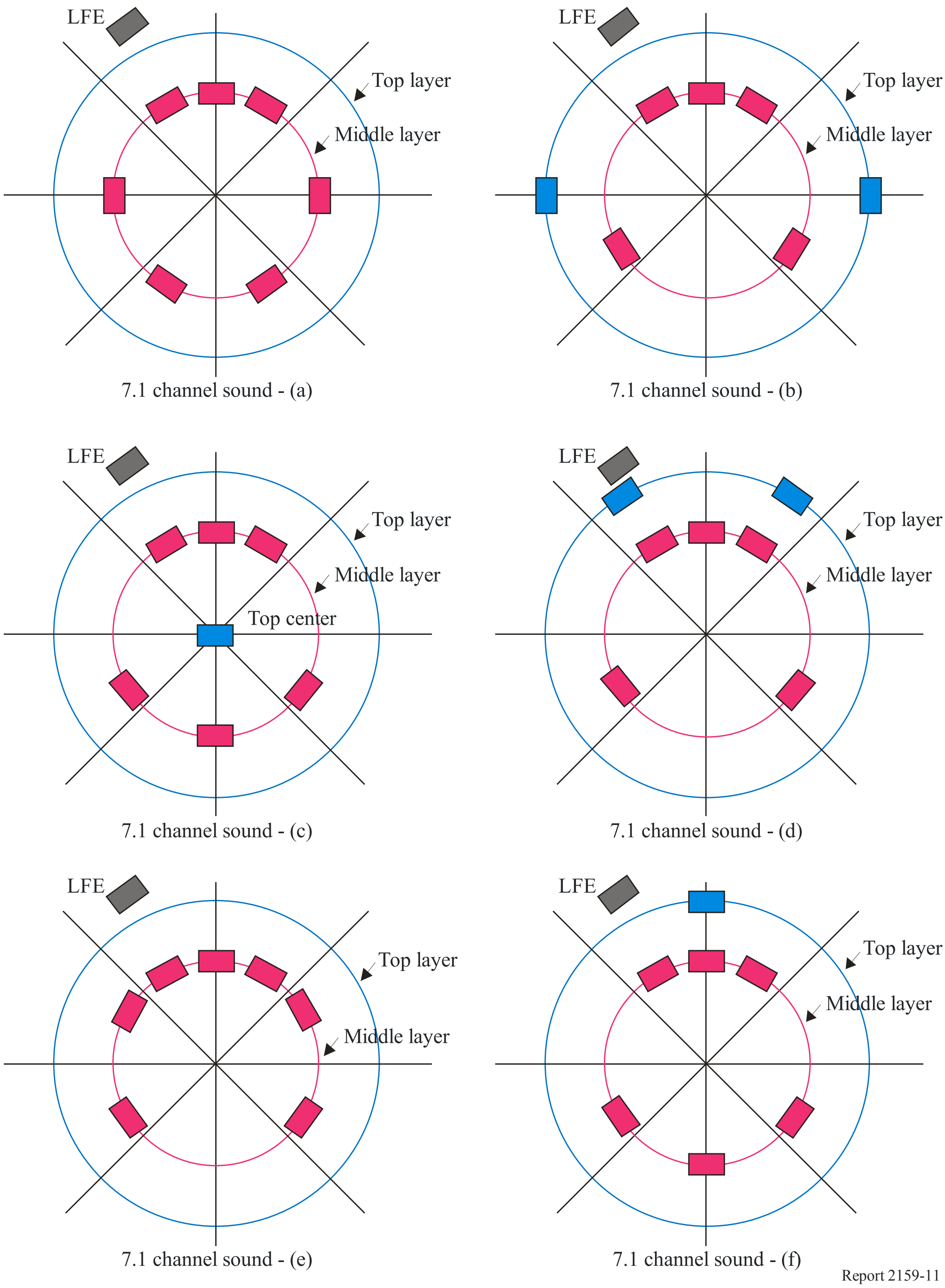
Specification of BD-ROM audio streams

CODEC	**LPCM**	**Dolby Digital**	**Dolby Digital Plus**	**Dolby lossless**	**DTS digital surround**	**DTS-HD**
Max. bit rate	27.648 Mbit/s	640 kbit/s	4.736 Mbit/s	18.64 Mbit/s	1.524 Mbit/s	24.5 Mbit/s
Max.ch	8(48 kHz, 96 kHz), 6(192 kHz)	5.1	7.1	8(48 kHz, 96 kHz), 6(192 kHz)	5.1	8(48 kHz, 96 kHz), 6(192 kHz)
bits/sample	16, 20, 24	16-24	16-24	16-24	16, 20, 24	16-24
Sampling frequency	48 kHz, 96 kHz, 192 kHz	48 kHz	48 kHz	48 kHz, 96 kHz, 192 kHz	48 kHz	48 kHz, 96 kHz, 192 kHz

Whilst 7.1 channel sound is available in Dolby Digital Plus and DTS-HD, several channel mappings are proposed in terms of 7.1 channel sound as shown in Fig. 11. The proposed mappings consist of two layers of loudspeaker positions, middle and top layer. The middle layer is basically at the same height with the listener's ear and the top layer is at a higher position such as at ceiling level.

FIGURE 11

Examples of loudspeaker mapping of 7.1 channel sound



6 Multichannel sound programme production in studio for home audio

6.1 Production of 5.1, 6.1 and 7.1 channels

Many countries are currently producing 5.1 channel sound programmes for broadcasting and audio and video releases. Production of 6.1 channel and 7.1 channel sound programmes is also increasing for audio and video releases. Several microphone techniques had been already proposed by many sound engineers and audio researchers for 5.1 channel sound recording. As described above, 7.1 channel sound is functional with the loudspeakers at a higher position. Several issues regarding how to use height channel properly or effectively were discussed in various workshops.

6.2 Production of 22.2 multichannel sound

6.2.1 Principles of three-dimensional sound mixing

NHK has already produced several UHDTV programmes with 3D sound using the 22.2 multichannel sound mixing system. Sound engineers and designers have been developing know-how and experience in the 3-D sound field. The current, conventional applications of layers on 22.2 multichannel sound used for mixing are enumerated below.

Top layer

– Reverberation and ambience.

– Sound localized above, such as loudspeakers hung in gymnasiums and airplanes and at fireworks shows.

– Unusual sound, such as meaningless sound.

Middle layer

– Basic sound field formation.

– Envelopment reproduction.

Bottom layer

– Sounds of water such as the sea, rivers, and drops of water.

– Sound on the ground in scenes with bird's-eye views.

Sound engineers have also been discussing several issues in 3D sound mixing. The principal issues are as follows.

– Effective use of the top and bottom layer.

– 3D movement of sound images.

– Creating a sense of elevation.

– Interaction between immersive audio and visual cues.

6.2.2 22.2 multichannel sound post-production system

A 22.2 multichannel sound post-production system has been developing for producing 3D sound. This system currently includes a Digital Audio Workstation and a sound mixing console with 3D pan on each channel strip and a down-mixing function. It can mix over 800 sound tracks to produce 22.2 multichannel sound.

FIGURE 12

22.2 multichannel sound mixing console

Report 2159-12

6.2.3 Examples of three-dimensional sound live mixing

6.2.3.1 A large-scale musical TV programme at NHK Hall

The 22.2 multichannel sound of a large-scale musical TV programme at NHK Hall was live mixed using about 150 sound input feeds. Multiple microphones were arranged in the manner of standard pop recording, basically as a “close setup near the sound sources”, so the 22.2 multichannel sound mixing was also done with the conventional pop music mixing technique, i.e., the multi-microphone recording technique. The major difference in microphone arrangement and mixing between 5.1 channel sound and 22.2 multichannel sound is in how the ambience of a concert hall is recorded and mixed. It is important to reproduce the acoustical impression of the huge dimensions of the NHK Hall, which has a 4 000-seat capacity and the impression of being surrounded by an enormous audience. Spatial sound reproduction advantages of the 22.2 multichannel system include the improvement of the listeners’ sense of envelopment and the enlargement of the listening area with exceptional sound quality. For the achievement of these new features of spatial reproduction with a 22.2 multichannel sound system, the following concept was planned as shown in Fig. 13.

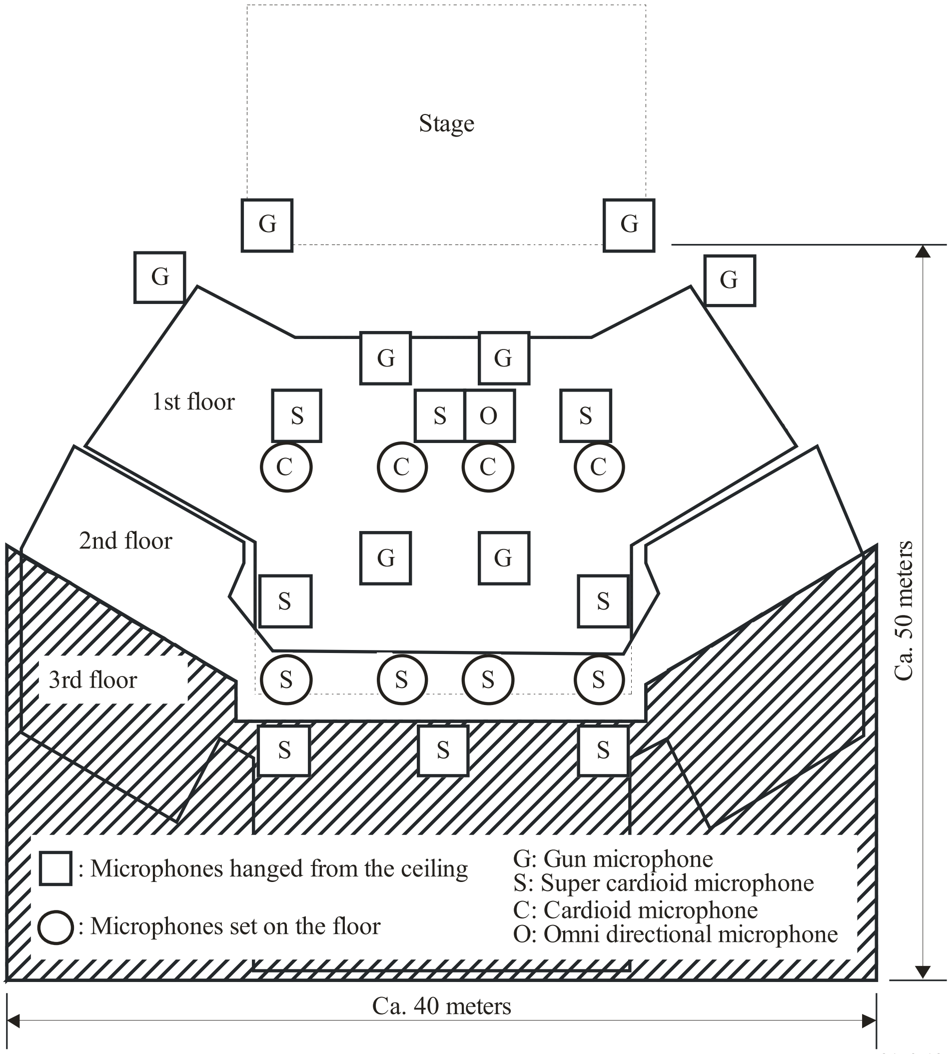
– Reflection and reverberation in the auditorium of NHK Hall, which are captured by microphones hung from the ceiling, are reproduced by the top layer loudspeakers of the 22.2 multichannel sound system to widen the listening area and create a sense of the listener being enveloped.

– The sounds of the audience, such as applause and shouts of encouragement, which are captured by several microphones set close to the audience, are reproduced by the middle layer loudspeakers to give the viewers a good sense of presence, as if they were sitting in the audience in NHK Hall.

– As the sound of musical instruments and vocals are reproduced by the sound reinforcement (SR) loudspeaker system in NHK Hall, reproduced sound reflected by the wall, ceiling, and floor of the hall is captured by the ambience microphones and reproduced by the top and middle layer loudspeakers to give the viewers the same sense of presence.

FIGURE 13

Arrangement of microphone for live mixing of a large-scale musical UHDTV programme by the 22.2 multichannel sound at NHK Hall



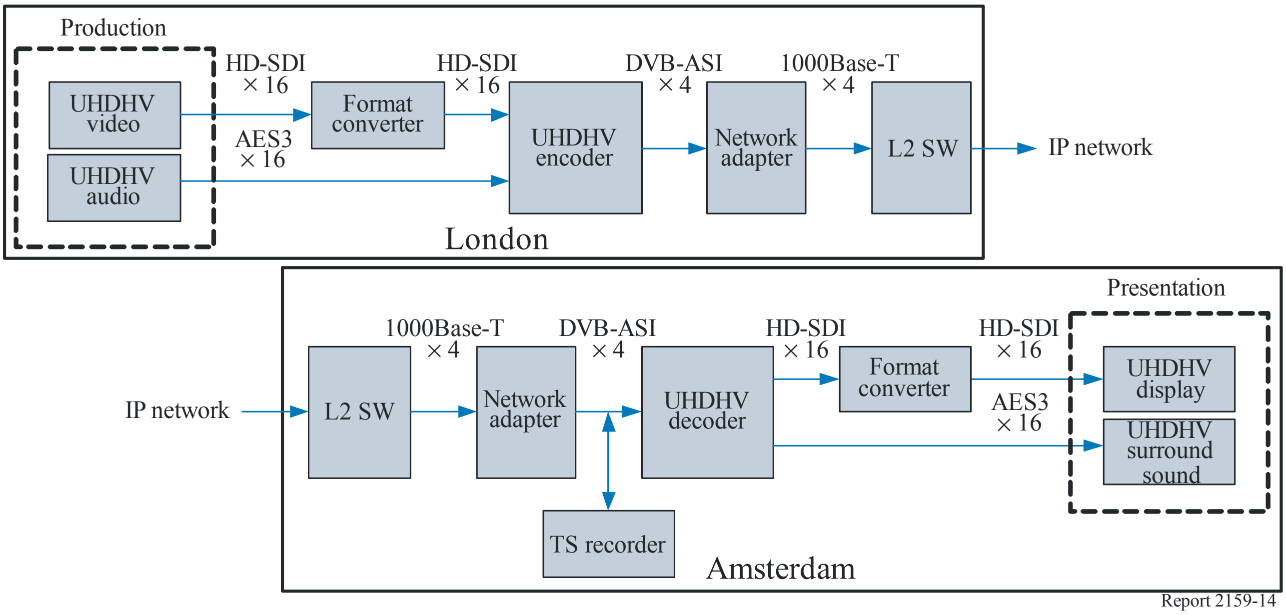
Report 2159-13

6.2.3.2 Emulated live news reports demonstrated at IBC2008

Ultra High Definition Television (UHDTV) and 22.2 multichannel sound technologies were demonstrated at IBC2008 by the international collaborative group called the Broadcast Technology Futures group (BTF), which included international live contribution link over an ultra-broadband IP network. The outline is depicted in Fig. 14.

FIGURE 14

UHDTV and 22.2 multichannel sound IP transmission system at IBC2008



UHDTV live pictures and sound captured in central London were carried to Amsterdam over an ultra-broadband IP network. In order to demonstrate the live nature of the link, the scenario set up was to emulate live news reports from London to Amsterdam with two-way interaction between a reporter in London and a presenter in the theatre in Amsterdam.

Sound acquisition system adopted in London was a microphone array with 15.2 system rather than a full-blown 22.2 system due to limitation on number of channels in mixing desk. This meant that there would be a middle layer containing eight of the ten specified 22.2 microphone complement, the top layer would be reduced to four microphones from nine, and the lower layer would have the full complement of five microphones, of which two were LFE channels. The microphone array is shown in Fig. 15. The total of 18 audio channels, including the 15.2 channels and one commentary channel, were sent to Amsterdam to be reproduced for the 22.2 multichannel system in the viewing theatre there. The 3D surround sound quality reproduced by 22.2 multichannel sound in Amsterdam was completely convincing; ambient sounds of London were reproduced effectively, even the sounds of airplanes and helicopters flying overhead sounded as if they were flying over the theatre.

FIGURE 15

Microphone array

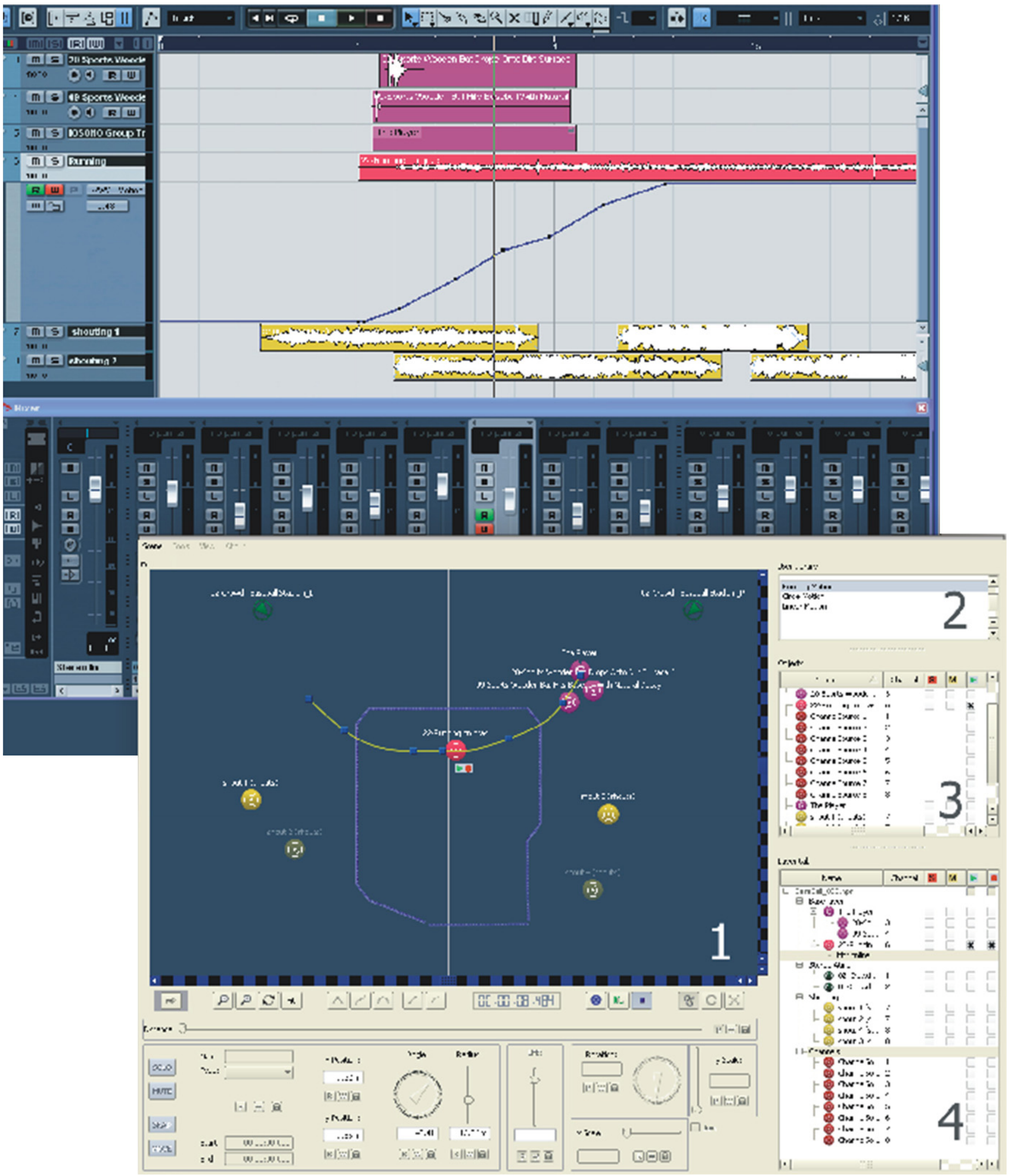
Report 2159-15

6.3 Production of WFS multichannel sound

The creation of a WFS sound scene involves associating spatial information with the sound signals composing the scene. The IOSONO Spatial Audio Workstation (SAW) is a tool for object-oriented production, editing and mastering of auditory scenes for reproduction in a WFS-equipped environment. It is realized as a plug-in for the well-known digital audio workstation Steinberg Nuendo 4. While Nuendo enables the editing and post production of audio streams the SAW plug-in enables the sound engineer to create advanced sound source movements and complex audio scenes based on the audio material loaded into the Nuendo session. Using the SAW, sound sources are positioned on the scene like marking points on a map. In addition, the SAW allows the definition of motion trajectories for up to 32 simultaneous sound objects. The mixer can assign a discrete position to each sound object for its x, y and z coordinates. For moving sound objects the position information is accompanied with a timestamp (SMPTE time code). The user has full control over the audio objects and the motion lines. Moreover the plug-in offers a wide range of functions for audio objects and motion lines, e.g. move, rotate, scale and group. The SAW is equipped with a powerful graphical user interface that allows the mixer to easily assign a discrete position to each sound object in the listeners' space. This gives the sound engineer a real intuitive view compared to traditional channel-oriented loudspeaker panning techniques. With this tool, even live mixing is possible.

FIGURE 16

The IOSONO spatial audio workstation for production of object oriented spatial audio scenes



Report 2159-16

7 Quality performance of the multichannel sound systems

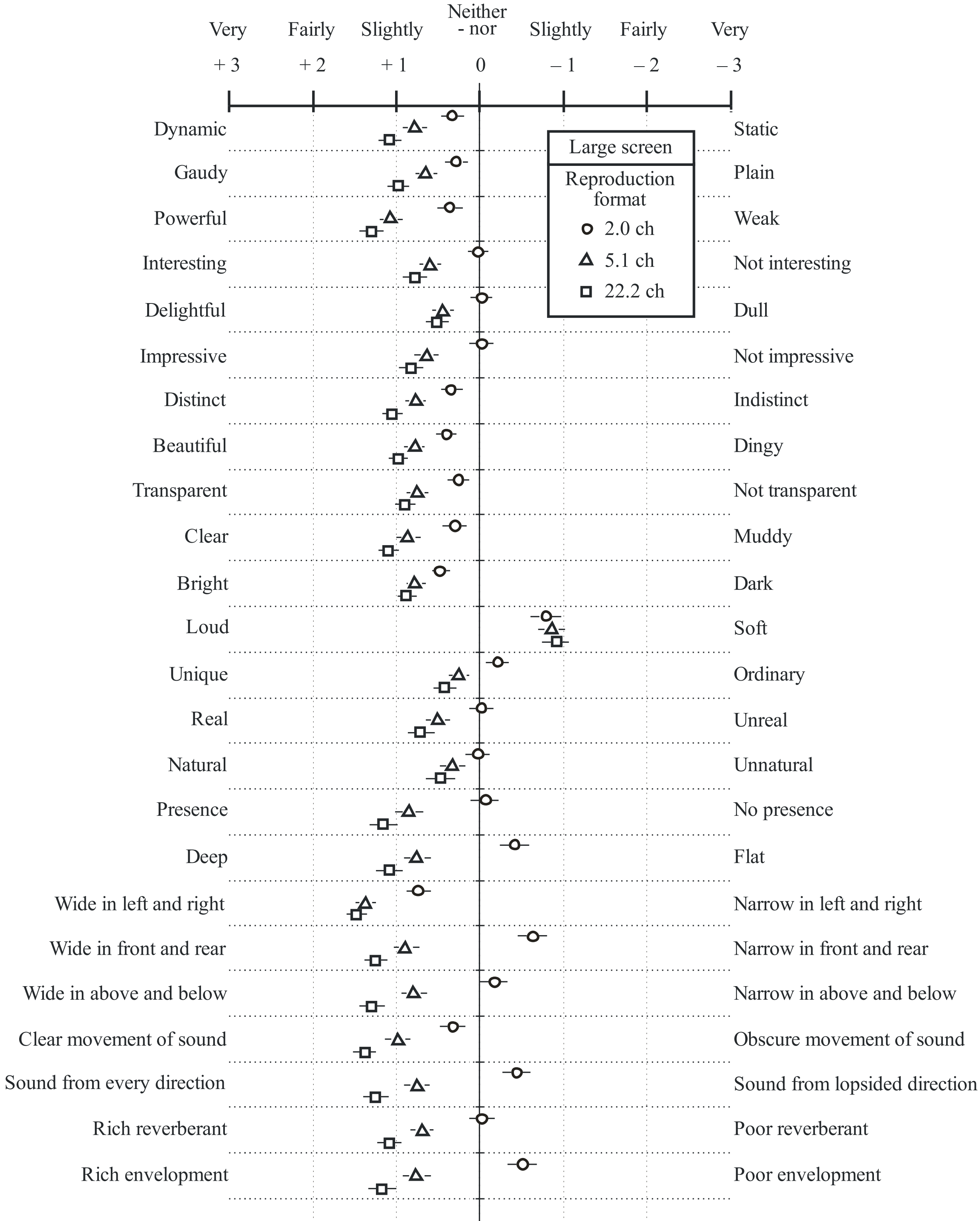
Subjective evaluations were conducted to assess the performance of three different multichannel audio systems: two-channel stereo, 5.1 channel sound, and 22.2 multichannel sound. The stimuli for the subjective evaluations were selected from the World Expo 2005 programmes. The sounds and video were screened at two locations: NHK Labs’ UHDTV theatre (a 450-in. screen that is 14 m long, 15 m wide, and 10 m high) and NHK Labs’ small post-production studio (a 50-in. screen that is 8 m long, 7.5 m wide, and 4 m high). The subjects were asked to report their impressions of the sound provided by the different sound systems when shown with different images on the screen. They were also asked to sit in different positions so that differences in impressions based on position with regard to the screen could be discussed.

The semantic difference evaluation method was used in this experiment. Subjects were asked to rate their impressions on a 7-point scale for the pairs of evaluation terms. Each pair contained two terms with opposite meanings, such as "dynamic" and "static." Subjects were asked to select a score from 3 to −3, in which 3 meant very dynamic, 2 meant fairly dynamic, 1 meant slightly dynamic, 0 meant neither dynamic nor static, −1 meant slightly static, −2 meant fairly static, and −3 meant very static. They rated each stimulus (segment of content) using the 24 evaluation pairs. There were 53 subjects (28 university students of music or audio engineering and 25 audio professionals) for this experiment.

Figures 17 and 18 show the total mean values of all the results from the 24 pairs of evaluation terms for each sound system in the large theatre and the small studio, respectively. Each mean value is marked with a 95% confidence interval for different terms and for different sound systems. The horizontal axis represents the scale of evaluation, and the vertical axis contains each pair of evaluation terms. Both figures show that the 22.2 multichannel sound system was rated significantly better (larger value) than the two-channel stereo system, for every evaluation term except "loud". Figure 17 shows that, in the large theatre, the 22.2 multichannel sound system was rated significantly better than the 5.1 channel sound system for the terms "gaudy", "distinct", "wide in front and rear", "wide in above and below", "clear movement of sound", "sound from every direction", "rich reverberant" and "rich envelopment". Figure 18 shows that, in the small studio, the 22.2 multichannel sound system was rated significantly better than the 5.1 channel sound system for every evaluation term except "loud", "dynamic", "gaudy" and "natural". The results also show that the 5.1 channel sound system was rated significantly better than the two-channel stereo system for every term except "loud". The results suggest that there was no difference in the loudness between each system. They also suggest that the 22.2 multichannel sound system provided a better 3D spatial sound quality than both the two-channel stereo and the 5.1 surround sound systems in both a large theatre and a small studio. Furthermore, the difference of rate between the 22.2 multichannel sound system and the 5.1 channel sound system or two-channel stereo system is basically bigger in a small studio than in a large theatre. Therefore, in a small studio, the 22.2 multichannel sound system may provide a better 3D spatial sound quality than both other sound systems than in a large theatre.

FIGURE 17

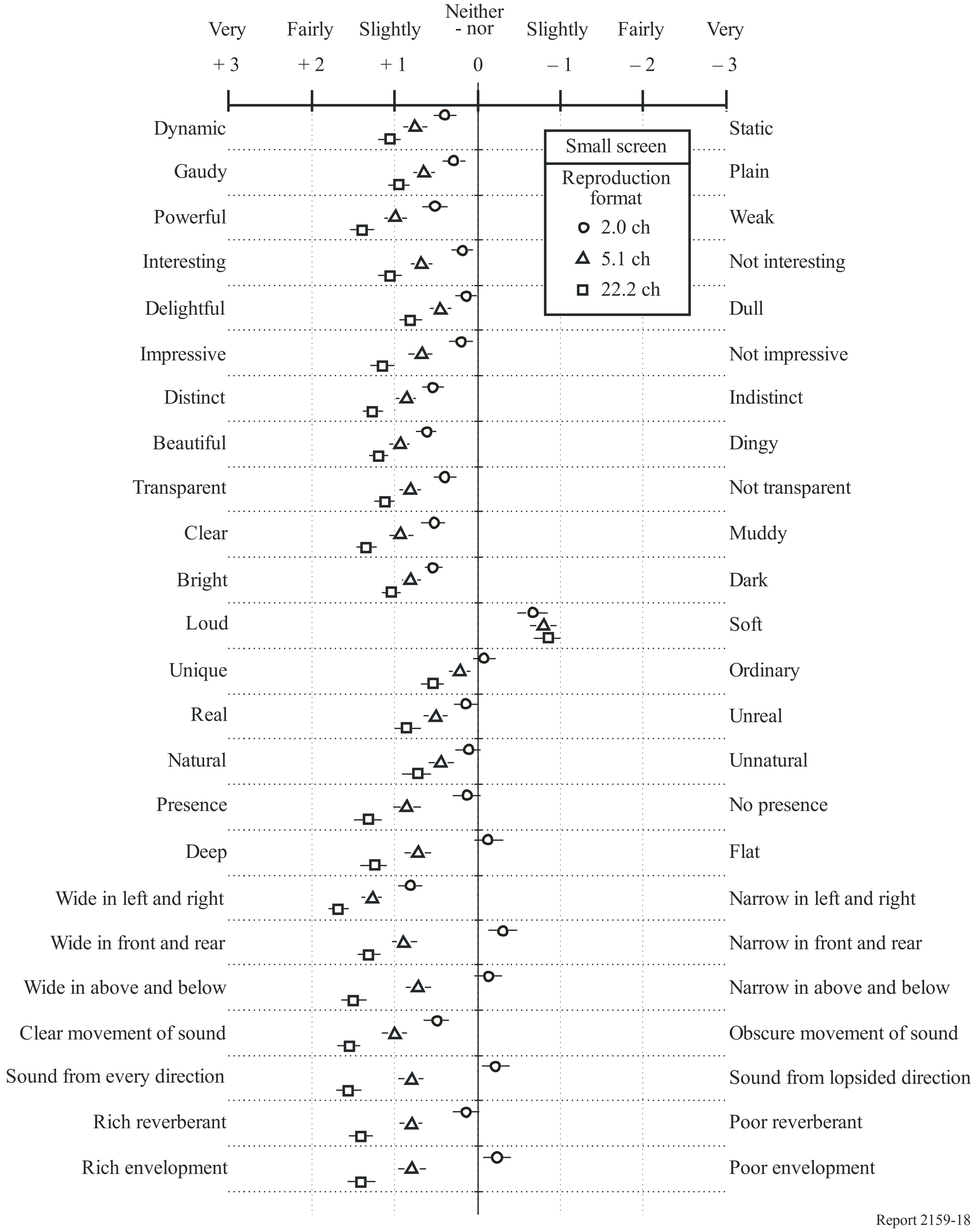
Results of subjective evaluation comparing three different sound systems with a large screen in a large theatre



Report 2159-17

FIGURE 18

Results of subjective evaluation comparing three different sound systems with a small screen in a small studio



8 Relevant documents concerning the multichannel sound systems developed by organizations outside ITU

8.1 SMPTE

8.1.1 SMPTE 2036-2-2008, "Ultra High Definition Television – Audio characteristics and audio channel mapping for programme production"

SMPTE 2036 Ultra High Definition Television (UHDTV) suite of documents is in multiple parts:

– Part 1: Image parameter values for programme production.

– Part 2: Audio characteristics and audio channel mapping for programme production.

SMPTE Standard 2036-2-2008 is Part 2 of SMPTE 2036 and describes the audio characteristics and audio channel mapping for programme production. This document specifies the characteristics of digital audio for UHDTV programme production and distribution, and also defines the mapping and labelling of 22.2 multichannel audio for UHDTV programme production.

The audio specifications are as follows:

1 Digital signal characteristics

UHDTV audio shall support a channel count of 24 full-bandwidth channels.

NOTE 1 – The two LFE channels are transported as full-bandwidth channels.

2 Channel mapping and channel labelling of 22.2 multichannel audio

TABLE 2

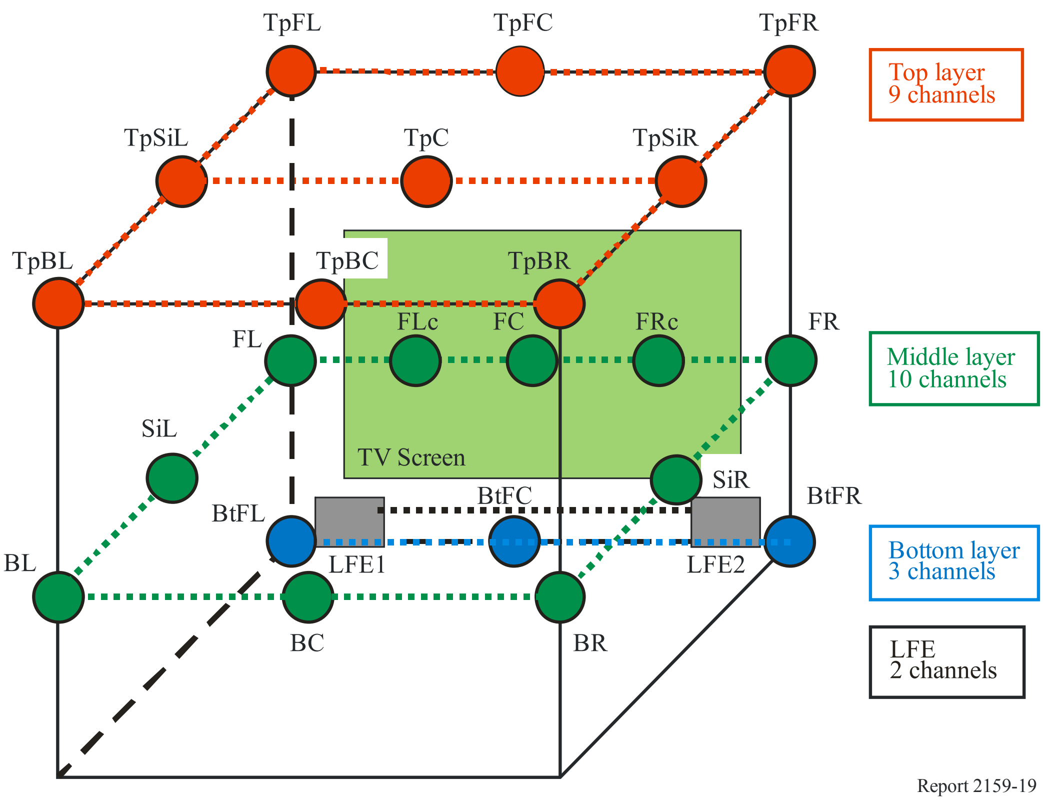
Channel maps and labels of 22.2 multichannel audio

AES Pair No./Ch No.	Channel No.	Label	Name
1/1	1	FL	Front left
1/2	2	FR	Front right
2/1	3	FC	Front centre
2/2	4	LFE1	LFE-1
3/1	5	BL	Back left
3/2	6	BR	Back right
4/1	7	FLc	Front left centre
4/2	8	FRc	Front right centre
5/1	9	BC	Back centre
5/2	10	LFE2	LFE-2
6/1	11	SiL	Side left
6/2	12	SiR	Side right
7/1	13	TpFL	Top front left
7/2	14	TpFR	Top front right
8/1	15	TpFC	Top front centre
8/2	16	TpC	Top centre
9/1	17	TpBL	Top back left
9/2	18	TpBR	Top back right
10/1	19	TpSiL	Top side left
10/2	20	TpSiR	Top side right
11/1	21	TpBC	Top back centre
11/2	22	BtFC	Bottom front centre
12/1	23	BtFL	Bottom front left
12/2	24	BtFR	Bottom front right

3 Loudspeaker layout (informative)

The Fig. 19 illustrates the loudspeaker layout of a 22.2 multichannel sound system.

FIGURE 19



Report 2159-19

8.2 MPEG (ISO/IEC JTC 1/SC 29/WG 11)

8.2.1 MPEG-2 AAC (ISO/IEC 13818-7)

MPEG-2 Advanced Audio Coding (AAC), which has been standardized by ISO and IEC, specifies low bit rate audio coding scheme which able to include up to 48 audio channels in one stream. The technical corrigendum ISO/IEC 13818-7:2006/COR 1 on Advance Audio Coding (AAC) has been approved to add the channel mapping for application specific channel configurations. The audio coding signal of 3D sound system including 22.2 multichannel audio can be transmitted by applying the application specific channel configuration.

Table 3 shows the example of an application specific channel alignment for a 22.2 channel configuration.

TABLE 3

Audio syntactic elements and channel alignment for an application-specific 22.2 channel configuration

Number of channels	Audio syntactic elements, listed in order received	Channel to speaker mapping
22+2	single_channel_element channel_pair_element, channel_pair_element, channel_pair_element channel_pair_element single_channel_element lfe_element, lfe_element, single_channel_element, channel_pair_element, channel_pair_element single_channel_element, channel_pair_element, single_channel_element single_channel_element, channel_pair_element	centre front speaker, left, right front centre speakers, left, right front speakers, left, right side speakers, left, right back speakers, back centre speaker, left front low frequency effects speaker, right front low frequency effects speaker, top centre front speaker, top left, right front speakers, top left, right side speakers, centre of the room ceiling speaker, top left, right back speakers, top centre back speaker, bottom centre front speaker, bottom left, right front speakers

8.3 EBU

8.3.1 EBU TECH 3306-2007, "RF64: An extended File Format for Audio"

The RF64 file format fulfils the longer-term need for multichannel sound in broadcasting and archiving. An RF64 file has additions to the basic Microsoft RIFF/WAVE specification to allow for either, or both:

– more than 4 Gbyte file sizes when needed;

– a maximum of 18 surround channels, stereo down-mix channel and bitstream signals with non-PCM coded data. This specification is based on the *Microsoft Wave Format Extensible* for multichannel parameters.

The file format is designed to be a compatible extension to the Microsoft RIFF/WAVE format and to the BWF format and its supplements and additional chunks. It extends the maximum size capabilities of the RIFF/WAVE and BWF format allowing for multichannel sound in broadcasting and audio archiving.

RF64 can be used in the entire programme chain from capture to editing and play out and for short or long term archiving of multichannel files.

An RF64 file with a bext chunk becomes an MBWF (multichannel BWF) file.

The following are specifications about audio channels:

1 Definition of a new format, RF64.

The wave format extensible channel mask contains 18 “#define” settings specifying different loudspeaker positions (or channel allocations).

Microsoft Wave Format Extensible Channel Mask

```
#define SPEAKER_FRONT_LEFT              0x00000001
#define SPEAKER_FRONT_RIGHT             0x00000002
#define SPEAKER_FRONT_CENTER            0x00000004
#define SPEAKER_LOW_FREQUENCY           0x00000008
#define SPEAKER_BACK_LEFT               0x00000010
#define SPEAKER_BACK_RIGHT              0x00000020
#define SPEAKER_FRONT_LEFT_OF_CENTER    0x00000040
#define SPEAKER_FRONT_RIGHT_OF_CENTER   0x00000080
#define SPEAKER_BACK_CENTER             0x00000100
#define SPEAKER_SIDE_LEFT               0x00000200
#define SPEAKER_SIDE_RIGHT              0x00000400
#define SPEAKER_TOP_CENTER              0x00000800
#define SPEAKER_TOP_FRONT_LEFT          0x00001000
#define SPEAKER_TOP_FRONT_CENTER        0x00002000
#define SPEAKER_TOP_FRONT_RIGHT         0x00004000
#define SPEAKER_TOP_BACK_LEFT           0x00008000
#define SPEAKER_TOP_BACK_CENTER         0x00010000
#define SPEAKER_TOP_BACK_RIGHT          0x00020000
```

8.4 Japan

8.4.1 Advancement of satellite digital broadcasting

The advancement of satellite digital broadcasting has been studied in Japan. Due to its utilization of state-of-the-art technologies, the advanced system is expected to enable the effective use of spectrum and the introduction of new services. The Ministry of Internal Affairs and Communications of Japan finally standardized the advancement of satellite digital broadcasting system in February 2009. The technical specifications were also standardized by the Association of Radio Industries and Businesses (ARIB) of Japan in July 2009.

The advanced system has a number of new features, including a transmission capacity increased by 30% or more due to the LDPC code and less roll-off factor, APSK modulations for a much higher capacity, higher resolution imagery and H.264|MPEG-4 AVC, three-dimensional surround sound, and the transport of IP packets targeted for multimedia storage.

The audio coding specifications are as follows.

TABLE 4

Audio coding for advancement of satellite digital broadcasting

Audio input	Maximum channels	22.2 channels per stream
	Sampling frequency	48 kHz
	Bit resolution	16, 20, or 24 bits
	Audio mode	Mono, stereo, multi-channel stereo (3.0, 4.0, 5.0, 5.1, 6.1, 7.1, 10.2, 22.2), dual- mono
Audio coding	Standard	MPEG-2 AAC, SBR (spectral band replication)
	Profile	LC

____