

Reproduction of 22.2 multichannel audio with virtual rendering

ITU-R Workshop "Topics on the Future of Audio in Broadcasting"

Wed 15th July 2015

Popov Room 16:30 - 20:00

Satoshi OODE

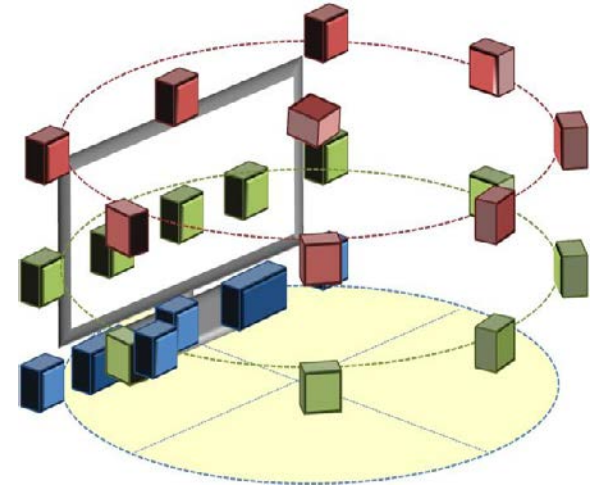
**Science and Technology Research Laboratories
Japan Broadcasting Corporation (NHK)**

Overview

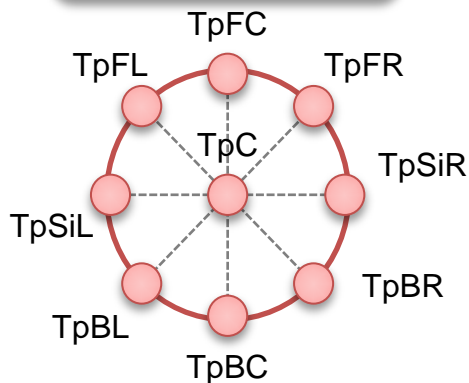
- 8K SHV broadcasting service
 - is planned to begin in 2016.
 - is composed of stereo, 5.1ch and 22.2ch with metadata related to dialogue level control.
- Requirement of 22.2 multichannel audio
 - Three dimensional spatial impression is achieved by loudspeakers placed in 30-45 degree intervals.
- Reproductions of 22.2 multichannel audio for home use
 - Theatrical environment using 24 loudspeakers or more.
 - Rendering to other channel configurations such as 9.1ch.
 - Loudspeakers integrated with the display using virtual rendering (Binaural reproduction over loudspeakers).
 - Headphone using virtual rendering (Binaural reproduction).

22.2 multichannel audio specified in Rec. BS.2051

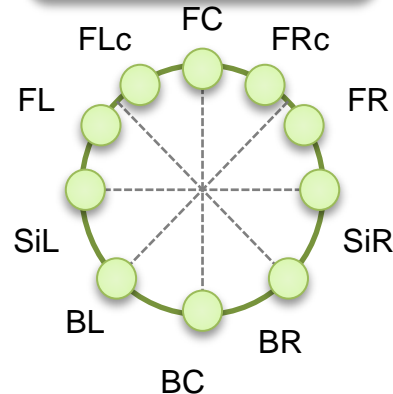
- The 22.2 multichannel sound system
 - is specified as system H in Rec. BS.2051.
 - consists of three layers.
- Top layer: 9 channels including overhead loudspeaker.
- Middle layer: 10 channels.
- Bottom layer: 3 channels including 2 LFE channels



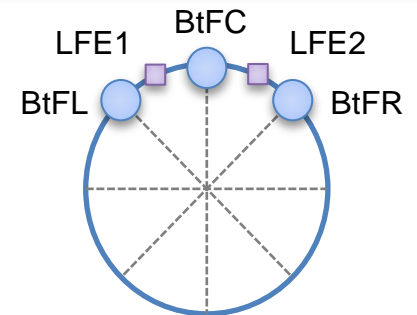
Top layer
9 channels



Middle layer
10 channels

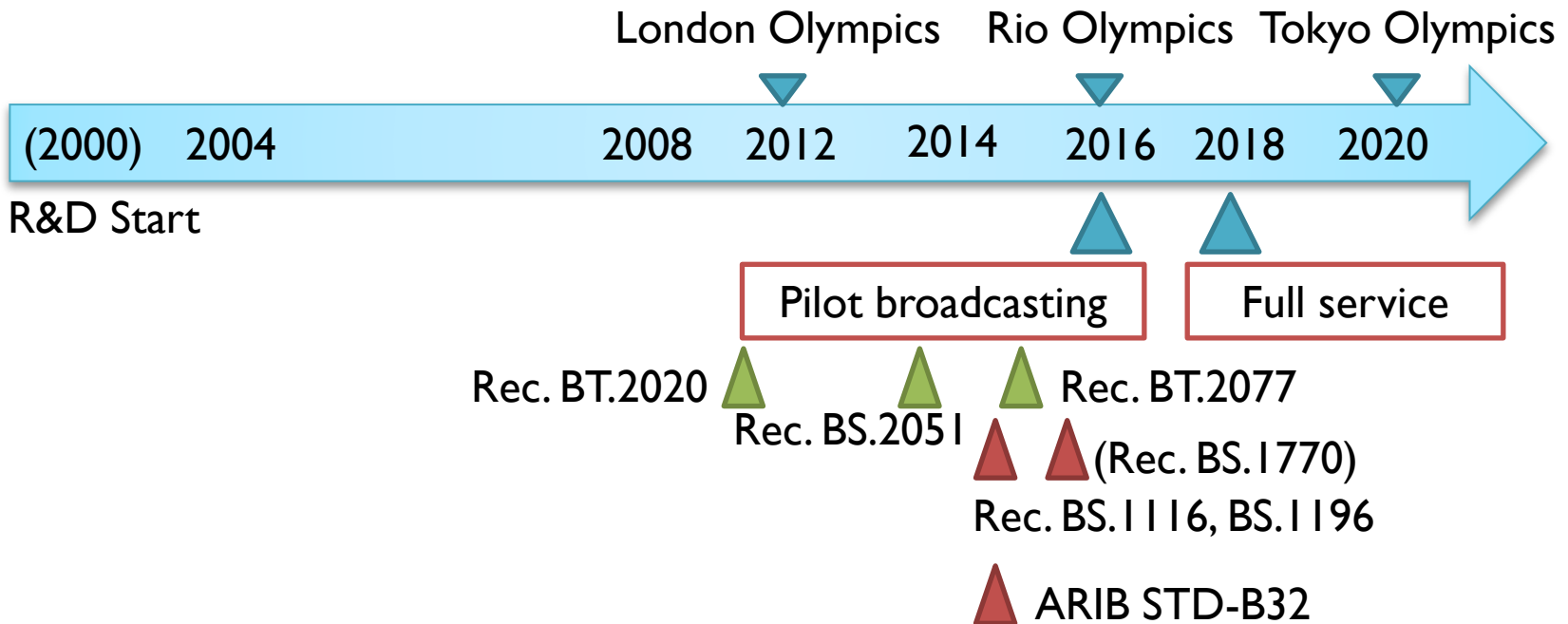


Bottom layer
3 channels + 2 LFE



Roadmap of 8K SHV broadcasting service

- 8K SHV pilot broadcasting service is planned to begin in 2016.
- Recommendations BT.2020 and BS.2051 were developed and some related recommendations were revised.
- ARIB Standard B32 which specifies audio coding was also updated in Japan.

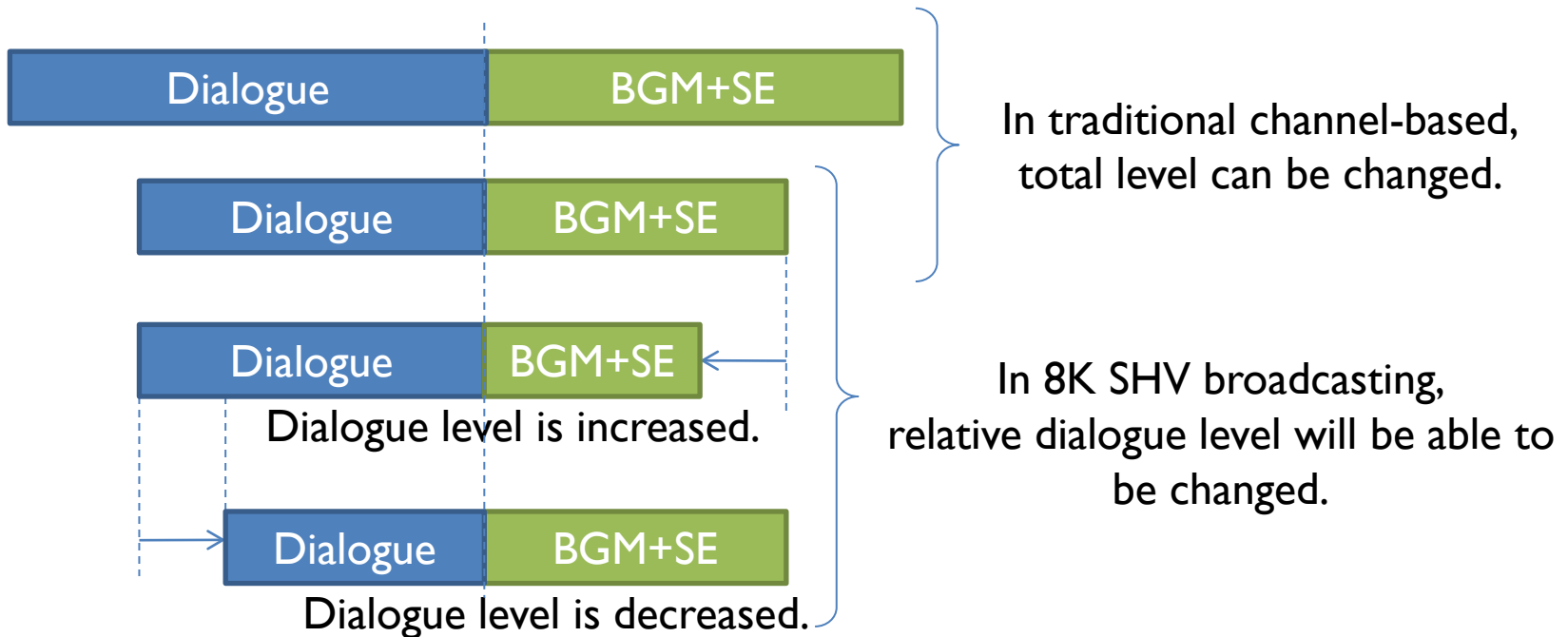


Audio service in 8K SHV broadcasting

- The revised ARIB Standard STD-B32 provides specifications of audio coding for the advanced satellite broadcasting system.
- New features are as follows.
 - Transmission of the down-mixing coefficients for each programme.
 - Dialogue level control and dialogue replacement.
 - Audio coding including lossless transmission.
- NHK plans to broadcast 8K SHV...
 - using MPEG-4 AAC in satellite broadcasting.
 - using stereo, 5.1ch and 22.2ch audio formats.
 - with metadata related to down-mixing for each programme.
 - with metadata related to dialogue level control function.
- The information is reported in Report ITU-R BS.2159-7.

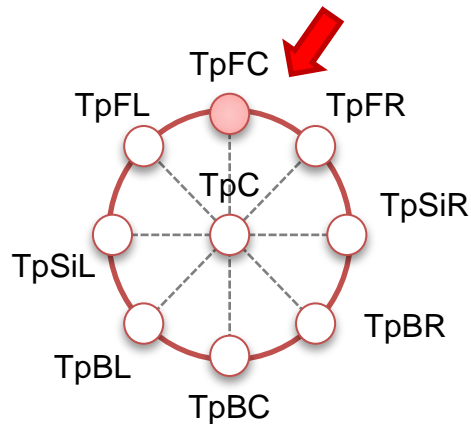
Dialogue level control function

- ◆ Many complaints with regard to the intelligibility of dialogue although “Dialogue” is the most important contents.
- ◆ The listeners can separately control the level of dialogue and that of total level.

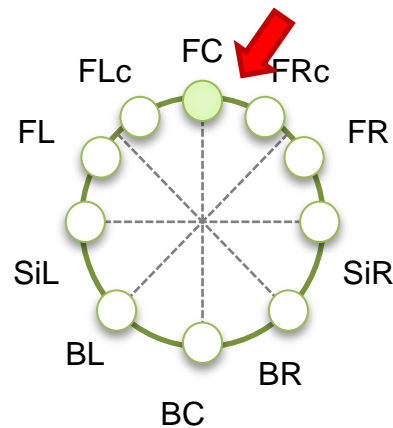


Dialogue level control in 22.2 multichannel audio

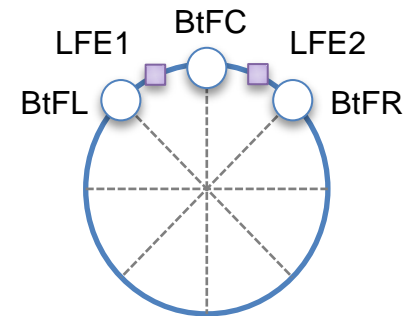
- ◆ Some channels are used only for dialogue depending on individual programmes.
- ◆ Dialogue channels have a flag of “dialogue”



Top layer
9 channels



Middle layer
10 channels



Bottom layer
3 channels + 2 LFE

Metadata for dialogue level control specified in ARIB STD-B32

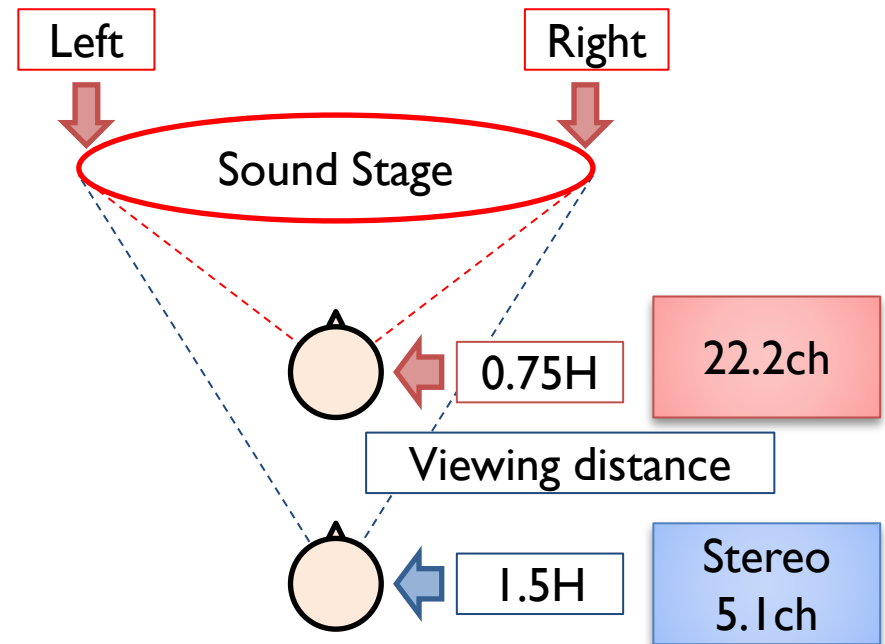
Descriptor	Explanation
ext_dialogue_status	Existence of dialogue channels.
num_dialogue_chans	Number of main dialogue channels.
num_additional_lang_chans	Number of alternative dialogue channels.
dialogue_src_index[i]	Index of dialogue channels.
dialogue_main_lang_comment_bytes	Byte count of characters indicating content of main dialogue.
dialogue_main_lang_comment_data	Byte data of characters indicating content of main dialogue.
dialogue_main_lang_code	Language code of main dialogue.
dialogue_additional_lang_code[i]	Language code of <i>i</i> th alternative dialogue.
dialogue_additional_lang_comment_bytes[i]	Byte count of characters indicating content of <i>i</i> th alternative dialogue.
dialogue_additional_lang_comment_data[i]	Byte data of characters indicating content of <i>i</i> th alternative dialogue.
dialogue_gain_index[i]	Gain of alternative dialogue channels. (0000: 0 dB, 0001: -1 dB, 0010: -2 dB, ..., 1110: -14 dB, 1111: $-\infty$ dB)
sn_dialogue_plus_index	Maximum gain of dialogue channels in receiver. (000: 0 dB, 001: +3 dB, 010: +6 dB, ..., 110: +18 dB, 111: $+\infty$ dB)
sn_dialogue_minus_index	Minimum gain of dialogue channels in receiver. (000: 0 dB, 001: -3 dB, 010: -6 dB, ..., 110: -18 dB, 111: $-\infty$ dB)
additional_dialogue_data_sync	Data stream element in which alternative dialogue data is packed.
additional_dialogue_index	Index of alternative dialogue channels corresponding to the “i” of dialogue_additional_lang_code[i].



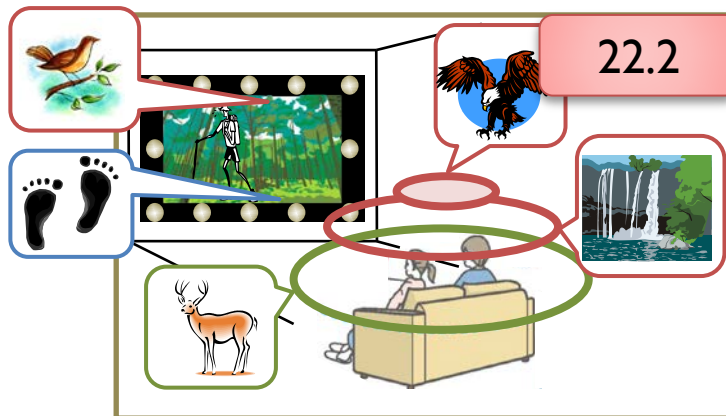
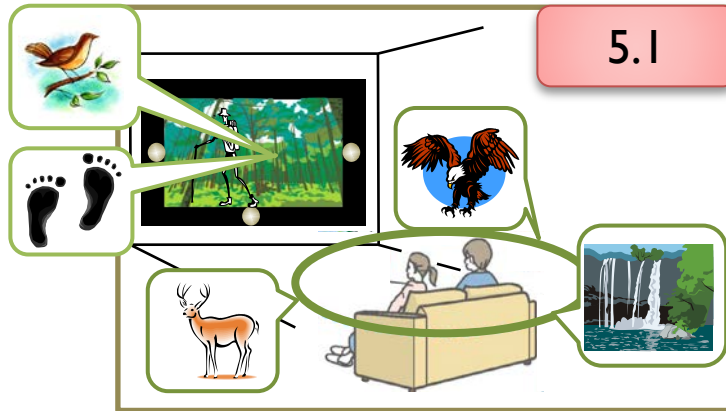
Limitation of dialogue level control is important for broadcaster because dialogue source is not always clean.

Viewing angle of 8K Super Hi-Vision

- System parameters are specified in Rec. ITU-R BT.2020
- 8K SHV has viewing angle of 100° in azimuth and 60° in elevation when the listener is positioned at $0.75H$ Height of the display.
- 8K SHV requires wider and higher sound fields than 5.1ch or stereo to match the visual image with the sound image .



Characteristics of 22.2 multichannel audio

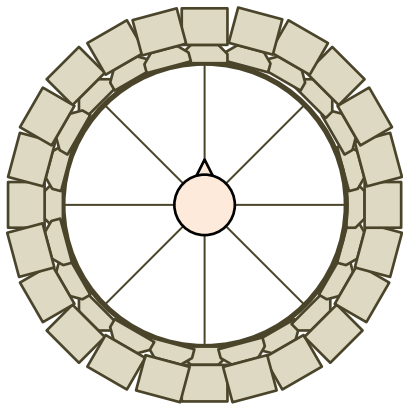


- ◆ Stable localization of frontal sound over the entire area of the large-screen image
- ◆ Sound image reproduced in all directions around the listener, including elevation
- ◆ 3D spatial impression augmenting the listener's sense of reality
- ◆ Wide listening area with excellent sound quality
- ◆ Compatible with existing multichannel sound systems
- ◆ Suitable for live recording, mixing, and transmission

Loudspeaker Intervals

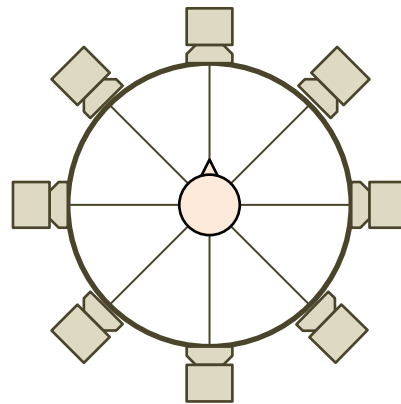
-Requirements of 22.2 multichannel audio-

Reference



24 loudspeakers
15 degree intervals

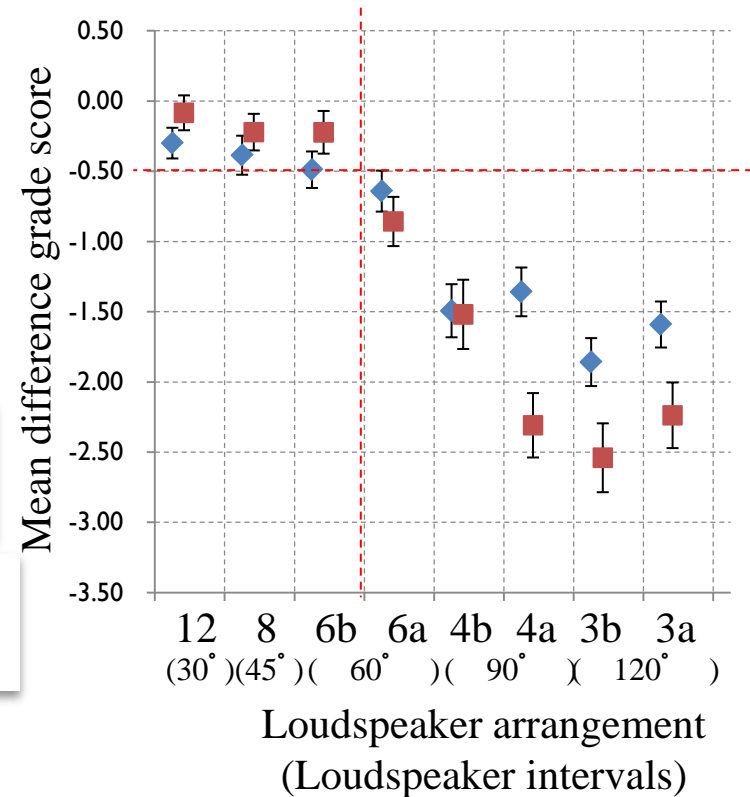
Test item



8 loudspeakers
45 degree intervals

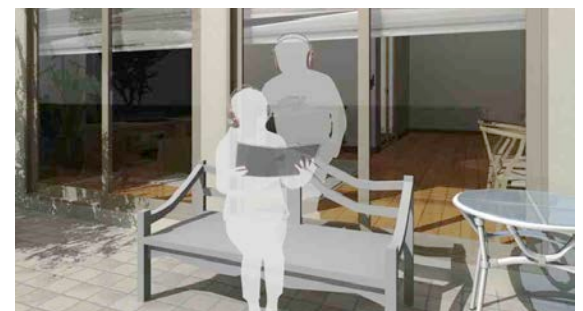
Spatial impression by 8 loudspeakers is almost the same as that by 24 loudspeakers.

◆ Top layer (45°) (40 subjects)
■ Middle layer (0°) (20 subjects)



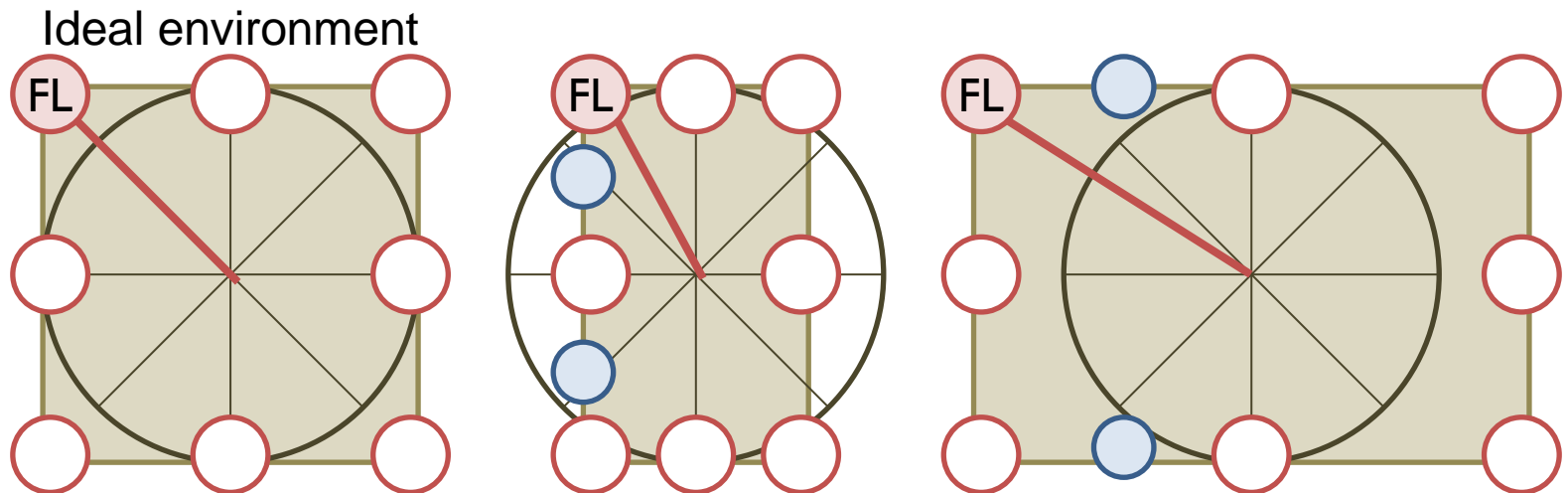
Reproductions of 22.2 multichannel audio for home use

- Theatrical environment using 24 loudspeakers or more.
- Rendering to other channel configurations such as 9.1ch.
- Down-mixing to stereo or 5.1ch
- Loudspeakers integrated with display using virtual rendering.
- Headphone with virtual rendering (Binaural).



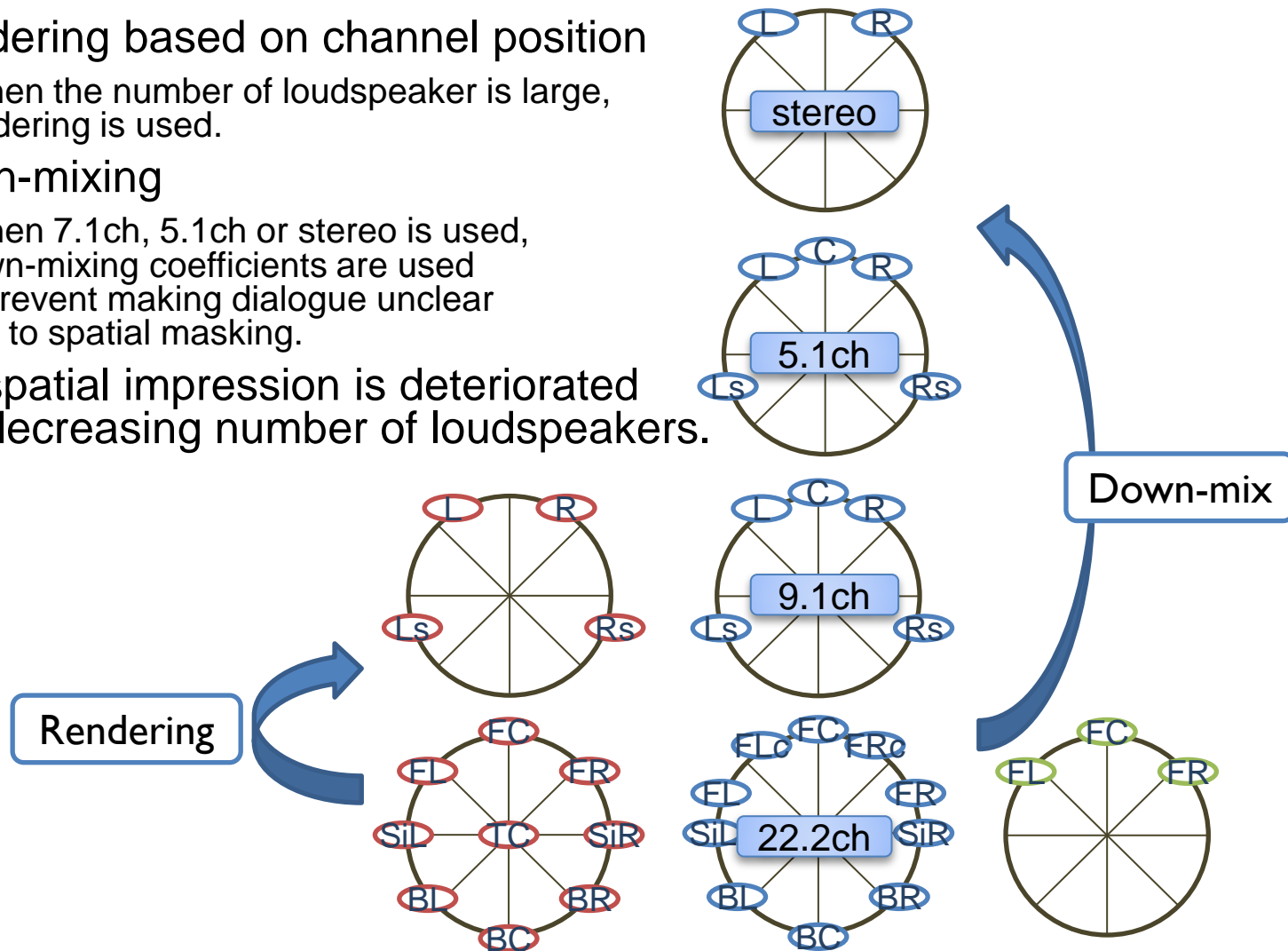
Theatrical environment using 24 loudspeakers or more

- ◆ 22.2 multichannel audio is usually reproduced by 24 loudspeakers
- ◆ Additional loudspeakers are added depending on listening environment.
 - ◆ Subwoofers of base management for full range channels (room size).
 - ◆ Full range loudspeakers on the side to keep uniformity (room shape).
- ◆ Not absolute positions of channels but relative positions are important.



Rendering to other channel configurations such as 9.1ch

- ◆ Rendering based on channel position
 - ◆ When the number of loudspeaker is large, rendering is used.
- ◆ Down-mixing
 - ◆ When 7.1ch, 5.1ch or stereo is used, down-mixing coefficients are used to prevent making dialogue unclear due to spatial masking.
- ◆ The spatial impression is deteriorated with decreasing number of loudspeakers.



Loudspeakers integrated with display using virtual rendering

- How to introduce 8K SHV Audio into the home environment
 - High-quality sound requires 24 discrete loudspeakers.
 - Installing 24 loudspeakers is over equipped for home environment.
- Compact and convenient system
 - Loudspeaker system should be integrated with SHV display.
- 12 loudspeakers system integrated with 85 inches SHV FPD was developed.

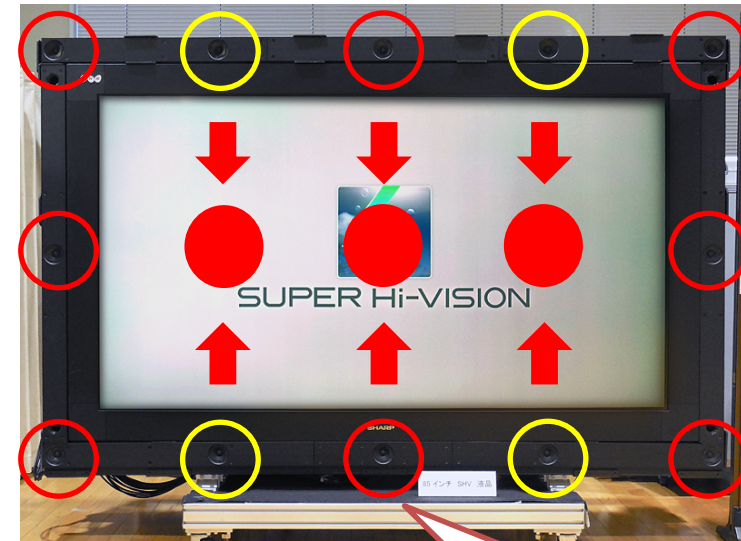
12 loudspeakers integrated with 8K SHV Flat Panel Display



Loudspeakers integrated with display using virtual rendering

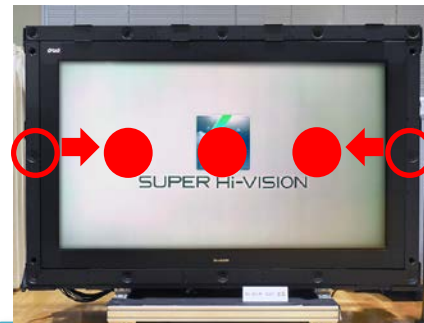
- ◆ For front 11 channels
 - ◆ 8 channels around the display are directly reproduced.
(marked as red circles)
 - ◆ 3 front channels on the display are reproduced using amplitude panning of vertical pair-wise.

12 loudspeakers integrated with 8K SHV Flat Panel Display



Vertical pair-wise panning method

loudspeaker



Horizontal pair-wise panning method

Loudspeakers integrated with display using virtual rendering

- ◆ 11 side, back and overhead channels around the listener are reproduced by binaural reproduction over 11 front loudspeakers.
(* to simulate acoustical propagation characteristics from the loudspeaker to each ear.)
- ◆ Binaural reproduction over loudspeakers has been studied since the 1960s.
- ◆ Some studies suggested horizontally arrayed loudspeakers can operate binaural reproduction very well.
- ◆ The system realizes immersive audio with only front loudspeakers.

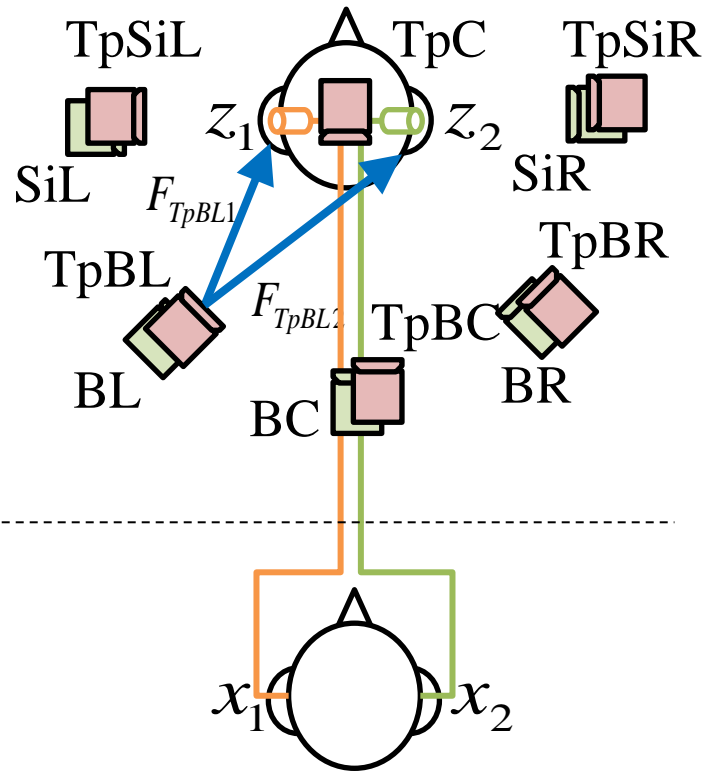
12 loudspeakers integrated with 8K SHV Flat Panel Display



loudspeaker

Loudspeakers integrated with display using virtual rendering

◆ Binaural Reproduction over loudspeakers

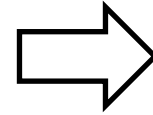
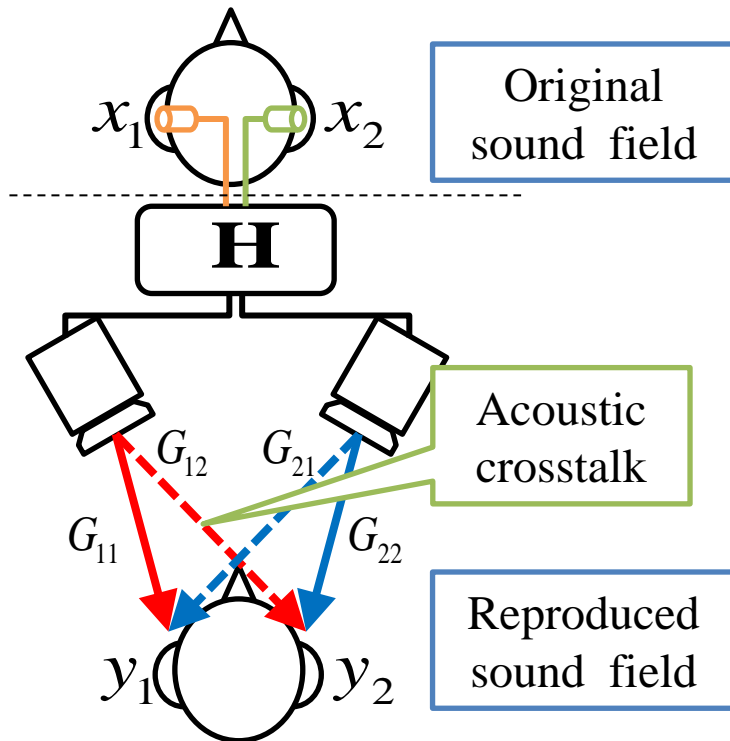


22.2 multichannel sound field* is simulated using HRTFs corresponding to each loudspeaker

*11 side, back and overhead channels

Loudspeakers integrated with display using virtual rendering

◆ Binaural Reproduction over loudspeakers



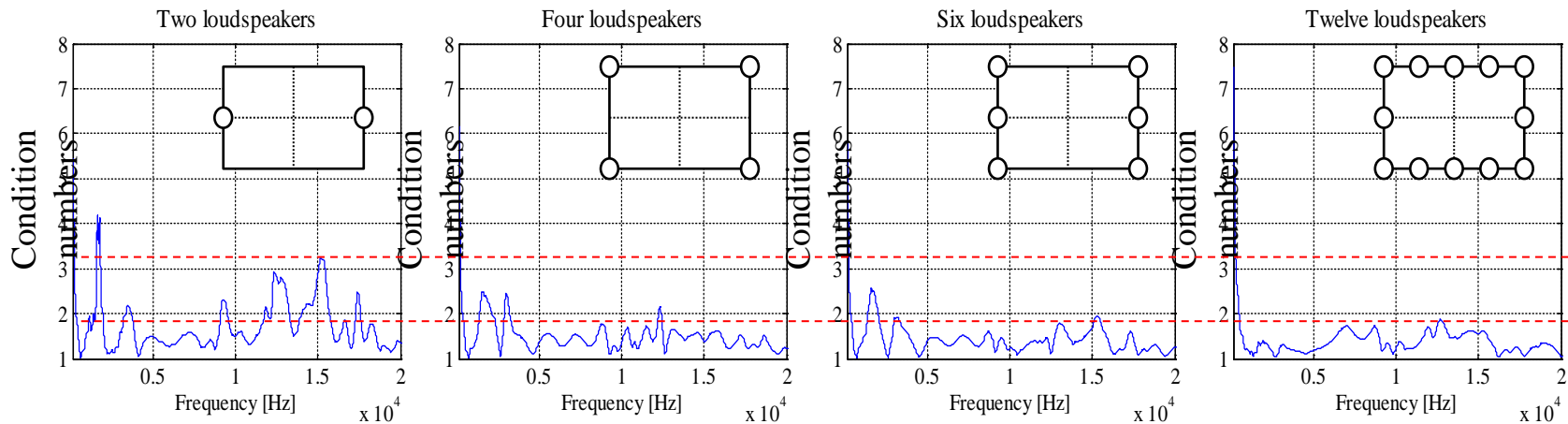
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} H \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Crosstalk cancellation is achieved by unit matrix

Condition number is one of the factors for system stability

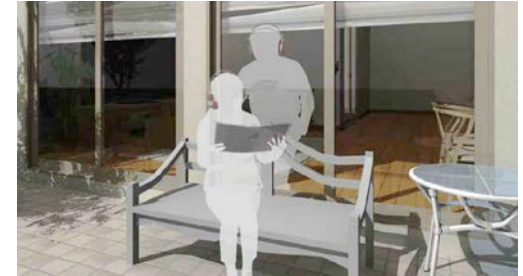
Loudspeakers integrated with display using virtual rendering

- Condition number is one of the factors which indicates the stability of the binaural reproduction.
- Increase of the number of loudspeaker realizes more stable binaural reproduction regardless of loudspeaker configuration

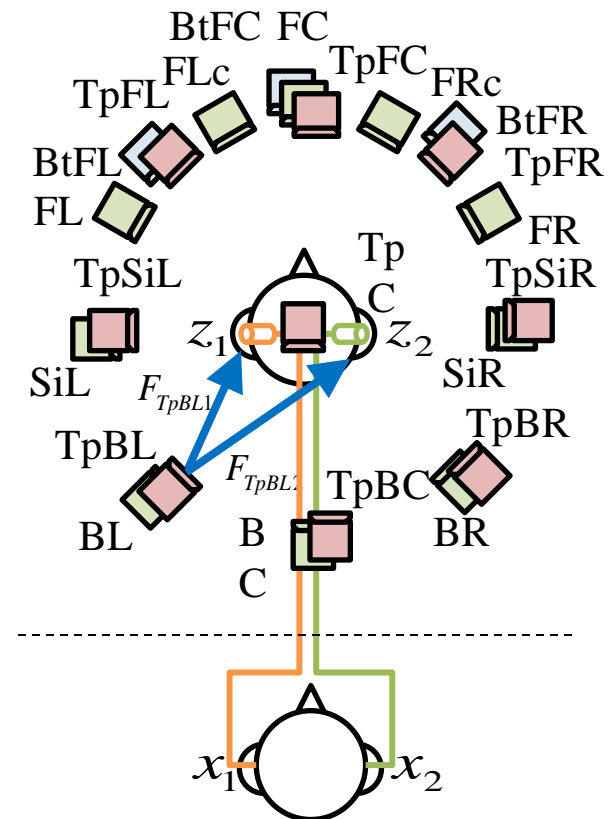


Headphone with virtual rendering (Binaural reproduction)

- ◆ For mobile or personal use, 22.2 multichannel headphone processor was developed.
- ◆ 22.2 multichannel sound field is simulated using HRTFs corresponding to each loudspeaker.
- ◆ The system in which audio engineer's HRTFs are installed has already used in 22.2 multichannel sound production.



22.2 multichannel headphone processor



Conclusion

- 8K SHV broadcasting service
 - is composed of stereo, 5.1ch and 22.2ch,
 - with metadata related to dialogue level control.
 - ➔ toward personalization, especially for the older person.
 - ➔ toward adaptation of the listening environment.
- Reproductions of 22.2 multichannel audio for home use
 - Theatrical environment using 24 loudspeakers or more.
 - Rendering to other channel configurations such as 9.1ch.
 - Loudspeakers integrated with the display using virtual rendering (Binaural reproduction over loudspeakers).
 - Headphone using virtual rendering (Binaural reproduction).

8K
SUPER Hi-VISION

Thank you for your attention

SUPER Hi-VISION

