

Implementation and Evaluation of 10.2 channel Microphone for UHDTV Audio

Daeyoung Jang, Jae-hyoun Yoo, Tae Jin Lee

Electronics and Telecommunications Research Institute of Korea(ETRI),
Audio Lab.,
P.O.Box 34219, 218 Gajeong-no, Yuseong-Gu, Daejeon, Korea
{dyjang, jh0079, tjlee}@etri.re.kr

ABSTRACT

As broadcasting environments change rapidly to digital, user requirements for next-generation services that surpass the current HDTV service quality become more demanding. The next-generation of broadcasting services will change from HD to UHD and from 5.1 channel audio to more than 10 audio channels, including a height channel for a high quality realistic broadcasting service. In accordance with the estimated trends of future broadcasting services, we propose a 10.2 channel audio format for a Korean UHDTV broadcasting service. It can create almost similar spatial sound images as 22.2 channel audio with half the number of speakers. In this paper, we propose a 10.2 channel audio acquisition system for the creation of UHDTV content, and measurements and preliminary evaluation are carried out to determine whether the performance is acceptable for broadcasting.

1. INTRODUCTION

Most advanced countries have completely changed their broadcasting services from analog to digital. Other countries are also preparing digital broadcasting services due to its transmission channel efficiency. In particular, ultra-high definition (UHD) video services have been launching for cinema and broadcasting areas since 2014. Immersive audio services are also required to provide spatial synchronization with a widened video display above 100 inches for a UHD video service. Additionally, both horizontal and vertical sound expression are necessary to provide spatial envelopment for a reasonably immersive audio service. Usually, many reports have insisted that an audio format with more than 10 channels is acceptable for immersive sound representation [1–4].

ETRI proposed 10.2 channel audio as an immersive sound format for a UHDTV broadcasting service in Korea in 2011 [5]. This format consists of seven channels of horizontal loud-speakers, three channels of ceiling loud-speakers, and two woofers. The seven-channel horizontal loud-speaker layout is the same as the Dolby/DTS 7.1 channel format for cinema.

This paper describes a 10-channel spherical microphone that can be used to record a 10.2 channel audio signal for broadcasting. First, section 2 illustrates a 10.2 channel format for UHDTV in Korea. Section 3 depicts the design and implementation of a 10-channel microphone system and section 4 describes characteristics of the implemented 10-channel microphone; finally, section 5 makes a conclusion about future work for further plans regarding evaluations and verifications.

2. 10.2 CHANNEL FORMAT FOR UHDTV

2.1. Layout of 10.2 channel audio format

The ideal layout of a 10.2 channel audio format as a beyond 5.1 channel realistic sound representation for UHDTV is depicted in Figure 1. The 10.2 channel audio format contains three channels of horizontal front loudspeakers, two side surround loudspeakers, two back surround loudspeakers, three ceiling loudspeakers, and two Low Frequency Effect (LFE) channels.

The horizontal front and side surround loudspeakers are compatible with a 5.1 channel loudspeakers setup, and the back surround loudspeakers are compatible with Dolby/DTS 7.1 channel loudspeakers for cinema sound. These two channels of back surround loudspeakers are important for reducing front–back confusion, the phenomenon in which listeners confuse whether a sound source is in front or behind them, because the Inter-aural Time Difference (ITD) and Inter-aural Level Difference (ILD) are similar.

An additional two front ceiling channels are located in the upper position of the screen at the adjacent direction of horizontal front left and right loudspeakers. These channels can represent ceiling sounds such as airplanes, lightning, etc. Another ceiling channel is located at a very high position relative to the listeners head or somewhat behind that. To support compatibility with a NHK 22.2 channel sound system and the establishment of a ceiling loudspeaker in a common listening room, it is recommended that the back ceiling loudspeaker is located between the vertical 90° and 135° position from the listener.

The LFE channels are two channels at the adjacent position of front left and right loudspeakers, and provide a flatter front sound image. Usually, these two channels have the same signal to provide powerful effect sounds for cinema, but a 90° phase shift from each other provides the envelopment of low frequency sounds for music reproduction.

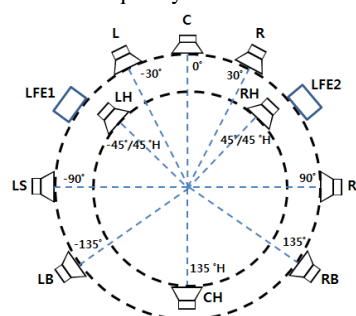


Figure 1: 10.2 channel loudspeaker layout.

2.2. Sound localization performance of 10.2 channel

We compared the 5.1 and 22.2 channel format through objective and subjective evaluation methods to evaluate the sound localization performance of the 10.2 channel audio format [6]. The objective sound localization method used in this paper is the Auditory Process Model (APM) by M. Park [7]. And evaluation test results are followed in subsections.

2.2.1. Objective test

APM is a mathematical model of a human's three-dimensional sound perception process that consists of a peripheral process to simulate neural transduction, a binaural process to simulate ITD, ILD characteristics, and central processes to determine sound localization. In these processes, APM can get similar results to subjective sound localization tests.

The objective horizontal sound localization performance of the right side of a 10.2 channel system was calculated using APM, and those of 5.1 and 22.2 channel systems were also calculated for comparison. The layouts of the loudspeakers of the 5.1, 10.2, and 22.2 channel formats under test were depicted in Figure 2.

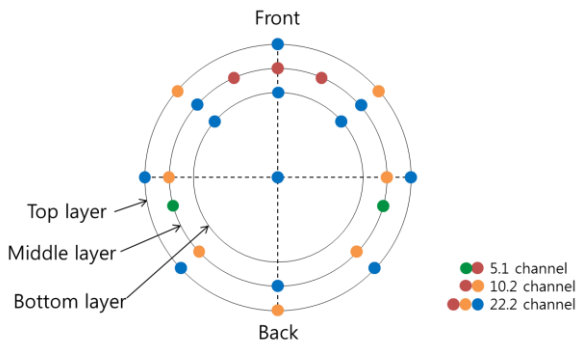


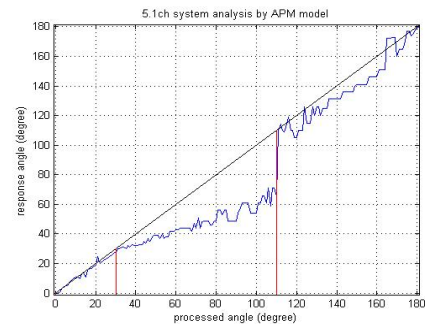
Figure 2: The loudspeaker layouts for tests.

The side-right surround loudspeaker of the 5.1 channel format is located at 110° from the center channel position. The subset of 22.2 channel loudspeakers are also used in the 10.2 channel format, because the 10.2 and 22.2 channel formats are compatible in the loudspeaker position. Sound source are localized by panning with adjacent loudspeaker pairs to obtain the APM parameters for every direction on the right side of the listener. Then, a test signal is reproduced and APM was calculated with stereo signals that were acquired at the sweet spot.

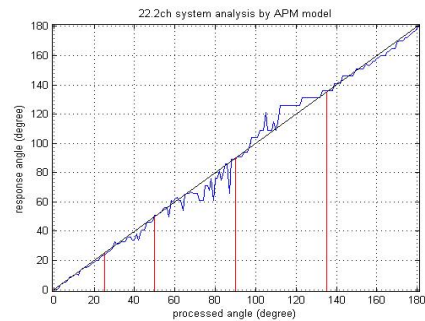
Figure 3 shows the results of APM for the horizontal right side channels of 5.1(a), 22.2(b), and 10.2(c) channels. In the graph of Figure 3, the horizontal axis represents the reproduced sound image and the vertical axis represents the sound image calculated by APM. Therefore, the sound localization performance is better when the curve is more coincident with the diagonal line. In Figure 3(a), the 5.1 channel system has degraded the sound localization performance for 30–110° of the side surround and 110–180° of the back surround.

In Figure 3(b), the 22.2 channel audio system shows very good sound localization performance because it uses enough loudspeakers for the horizontal surround sound image.

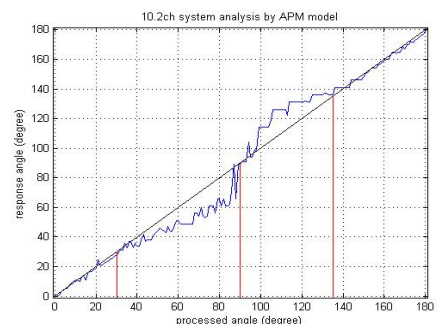
Compare Figure 3(c) to 3(a) and 3(b); the 10.2 channel audio system shows reasonable sound localization performance, meaning that the 10.2 channel audio system has better sound localization than the 5.1 channel audio system. However, the 10.2 channel audio system has a somewhat poorer sound localization performance than the 22.2 channel audio system, but a quite stable sound localization performance was obtained compared to the 5.1 channel audio system.



(a) 5.1 channel



(b) 22.2 channel



(c) 10.2 channel

Figure 3: Compare of sound localization by APM

2.2.2. Subjective test

Subjective testing with a 22.2 channel audio system was also conducted to confirm the sound localization performance of the 10.2 channel audio system. A 22.2 channel audio system was installed in the listening room, and the 10.2 channel audio system was implemented as a subset of the 22.2 channel audio system.

The subjects used for this listening test are nine spatial audio experts and three non-experts. In the test, the subjects heard several localized sound sources rendered by the 10.2 and 22.2 channel audio systems.

Table 1: Direction of stimuli for sound localization test.

	Elevation Angle (°)	Azimuth Angle (°)
A	70	330
B	90	15
C	80	80
D	60	260
E	70	165

Five stimuli of directional sounds for sound localization tests are defined in Table 1. Each sound source was rendered using the Vector Base Amplitude Panning (VBAP) method designed by V. Pulkki [8]. The subjects evaluated the sound localization performance of the 10.2 and 22.2 channel audio systems by listening to the submitted sound sources and pointing in the perceived sound direction with a laser pointer.

Figure 4 shows the subjective test results for the sound localization performance of the 10.2 and 22.2 channel audio systems. The 22.2 channel audio system had a 7.8° average difference for the perceived sound position, and the 10.2 channel audio system had a 10.6° average difference for the sound position. This 2.8° difference is an angle at which humans cannot usually perceive a difference in direction with their auditory system. Therefore, the results show that the sound localization performance of the 10.2 channel audio system is not degraded compared to that of the 22.2 channel audio system.

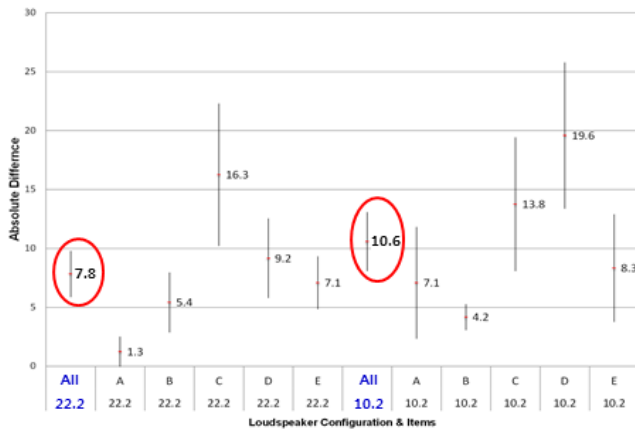
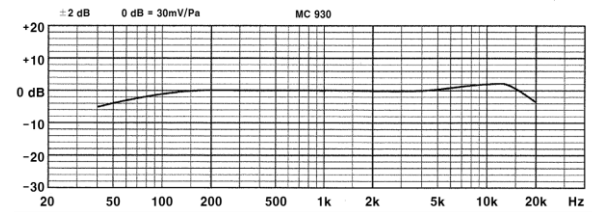


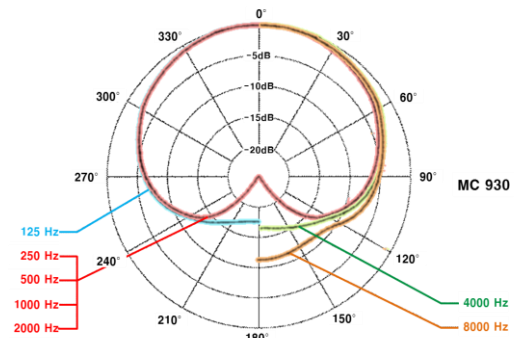
Figure 4: Subjective sound localization test results for the 10.2 and 22.2 channel audio systems.

3. IMPLEMENTATION OF A 10.2 CHANNEL MICROPHONE

The microphone units used in 10.2 channel microphone systems are cardioid directional microphones (Beyerdynamic MC930). The microphone unit has very flat frequency response, as shown in Figure 5(a), and has a stable directivity pattern for a wide frequency range, as shown in Figure 5(b).



(a) Frequency response

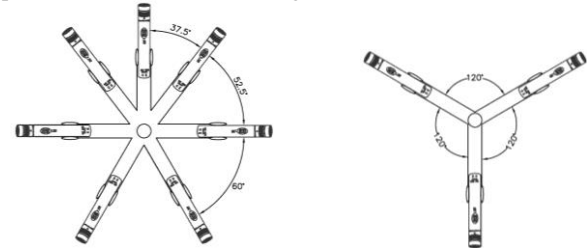


(b) Directivity patterns

Figure 5: Characteristics of the MC-930 (Beyerdynamic).

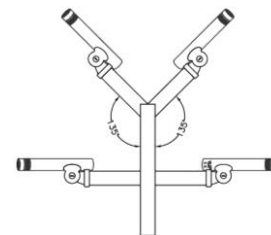
A 10-channel one point spherical microphone set was designed, as shown in Figure 6. The actual shape of the implemented microphone system is shown in Figure 7. The diameter of the microphone system is designed at about 55 cm, providing sufficient inter-channel signal separation while still being easy to carry.

Each microphone holder is a floating type for shock-and-vibration-free recording, with elastic string nets. The front left and right microphones have a wider angle of 37.5° from the center microphone for a lower correlation. In addition, the LFE channel can be generated by additional post-processing filtering and mixing.



(a) Lower channels layout

(b) Height channels layout



(c) Vertical section view

Figure 6: Diagram for the 10-channel microphone array.



Figure 7: Picture of an implemented 10-channel microphone.

4. MEASUREMENTS AND FIELD RECORDING FOR LISTENING

4.1. Measurements of directivity characteristics

The directivity of the implemented microphone system was measured in the listening room by the frontal sound source. Because the 10-channel microphone system has a symmetric shape for microphone directions, only the left side microphones were measured. Figure 8(a) shows the relative frequency responses for the frontal sound stimulus of white noise. The front center channel has the highest gain for almost all of the frequency range, and about 3 dB of the relative higher gain with front left channel.

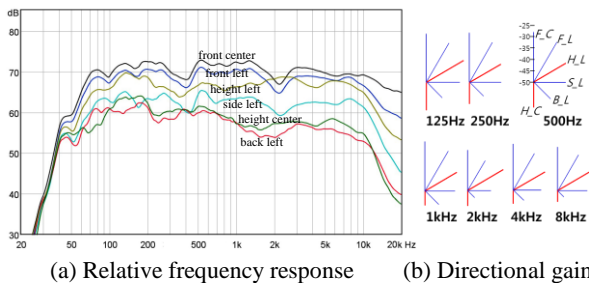


Figure 8: Directivity characteristics of the left half-sphere of the 10-channel microphones for a frontal sound source.

Figure 8(b) shows the directional gains for several typical frequencies. These directional gains represent the directivity characteristics of a 10-channel microphone system. The measurement results show that the implemented microphone system has very stable directivity characteristics throughout the whole frequency range.

4.2. Field recording and listening

Field recording was conducted using a 10-channel microphone system in the Korean traditional music hall (Figure 9(a)) and concert hall (Figure 9(b)). The Korean traditional music hall is an almost rectangular shaped room and has a somewhat reverberant effect. The recorded music pieces were interior music “cheon nyon manse,” Korean string instrument “hae geum solo” and the “song of Korean poetry.”

Several recording systems were used in the concert hall for comparison. Several stereo microphone setups and a 5.1 channel microphone were installed in the room, as shown in Figure 8(b). The recorded music contents included “Hungarian dances No. 1” by Brahms, “voice of spring” by Johann Strauß II, and “violin concerto” by Mendelssohn.



Figure 9: 10-channel field recording.

Recorded music contents were evaluated by informal listening tests performed by recording engineers and architectural acoustics experts (Figure 10). When adjusting the reproduction loudness levels of recorded music content similar to that of the recording place, we could feel similar envelopment and sound images. In addition, the frontal sound images of 10-channel sounds are more stable than those of 5.1 channel sounds.



Figure 10: Listening room for the field recording sounds.

5. CONCLUSION

This paper proposes a 10-channel microphone system for acquiring 10.2-channel high quality spatial sounds for Korean UHDTV broadcasting. A 10-channel microphone system was designed, implemented, and verified using directivity measurements and listening to field-recorded 10-channel audio content. The results of directivity measurements show that characteristics of a 10-channel microphone system are reasonable for spatial sound recording. It is identified through additional informal listening tests that an implemented 10-channel microphone can be used for the live recording of music programs for broadcast.

We have additional future plans to verify the performance of our 10-channel microphone system for use in broadcasting contents production. First, formal listening tests for 10-channel microphone systems will be conducted to confirm the performance of microphone systems for UHDTV broadcasting. Then, additional field recording will be followed for several programs such as sports, street scenes for live news, and wild life documentaries. Furthermore, we have to generate LFE channel signals from 10-channel signals recorded with a low-pass filter, mixing, etc.

6. ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (B0101-16-0295, Development of UHD Realistic Broadcasting, Digital Cinema, and Digital Signage Convergence Service Technology)

7. REFERENCES

- [1] Kimio Hamasaki et al., "Development of a 22.2 Multichannel Sound System," NHK STRL Broadcast Technology, No. 25, Winter 2006.
- [2] Kazuho ONO et al., "Portable spherical microphone for Super Hi-Vision 22.2 multichannel audio," AES 135th Convention, Oct. 2013.
- [3] Report ITU-R BS.2159-4, "Multichannel sound technology in home and broadcasting applications," ITU-R BS.2159-4, May. 2012.
- [4] White paper, "Dolby® ATMOS™ Next-Generation Audio for Cinema," Dolby Laboratories, Inc., 2013
- [5] TTA.KO-07.0098, Audio Signals for UHD Digital TV, 2011.
- [6] Taejin Lee et al., "Multichannel Audio Reproduction Technology based on 10.2 ch for UHDTV," JBE Vol. 17, No. 5, September 2012.
- [7] Munhum Park, Phillip A. Nelson and Kyeongok Kang, "A Model of Sound Localisation Applied to the Evaluation of Systems for Stereophony," ACTA Acoustica, Vol. 94, pp. 825~839, 2008.
- [8] Ville Pulkky, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," J. Audio Eng. Soc., Vol. 45, No. 6, June 1997.