Preface to the Special Issue on High-reality Audio: From High-fidelity Audio to High-reality Audio

Akio Ando*
Guest Editor-in-Chief
3174–8 Machida, Tokyo, 195–0062 Japan

In the mid-20th century, many youths longed for audio. There was a television program, “The Golden Ring,” broadcast in May 1965 by the British Broadcasting Corporation that described people who gave their all to audio, as well as the enthusiasm of revival from the devastation of World War II. The recording style described in the program became a thing of the past, while low-cost live recording is the mainstream today. In that era, one of the targets of audio was high fidelity. This word, “fidelity,” meant fidelity to the tone or timbre of the original sound. Audio has since evolved from the traditional two-channel stereo towards three-dimensional (3D) multi-channels by increasing the spatial dimensions. From the viewpoint of current 3D audio, “fidelity” may be considered to mean fidelity to the whole sound space. However, this is not easy to achieve even if state-of-the-art audio technologies are used. Therefore, we hope to propose high-reality audio aiming at reproducing the impression of the sound space.

Recently, papers about “spatial audio” can often be seen in the academic journals of acoustics. Among them, some papers show only computer simulation results of evaluating their proposed methods. Such studies might be outside the scope of high-reality audio. The quality of sound cannot be evaluated until a listener with a normal hearing ability listens to the sound. In the study of high-reality audio, the sound quality must be evaluated subjectively, because the impression of the reproduced sound is the most important.

High-reality audio is an academic area concerned with capturing high-quality sound events such as musical performances, transmitting or recording them, and reproducing them. In the capturing of sound, it is important to consider the properties of sound sources, such as musical instruments, and those of electroacoustical devices, such as microphones, as well as the room acoustics. In the transmission of sound, the signal processing technologies including audio coding are necessary. In recent 3D audio, sound field control technology is often used. Moreover, in the evaluation of sound impression, psycho-acoustical study should be involved. Therefore, high-reality audio is considered to be an interdisciplinary area crossing electro-acoustics, architectural acoustics, musical acoustics, psychoacoustics, signal processing, and more. Reflecting such characteristics of high-reality audio, this special issue has three invited papers regarding recording of musical sound, coding of the recorded sound, and subjective evaluation of downmixed sound in the reproduction.

There are two typical microphone placements for a two-channel stereo, namely, a coincident microphone (i.e., with capsules placed immediately above one another) and a spaced microphone. The former can capture the direct sound with clear localization and the latter can capture the sound stage with rich ambience. On the other hand, higher-order Ambisonics uses a spherical array microphone that has a property similar to that of a coincident microphone, that is, it may not capture the rich ambient sound. The first paper, “Spaced A-B placements of higher-order Ambisonics microphone arrays: Techniques for recording and balancing direct and ambient sound” written by F. Grond, J. Kelly, and W. Woszczyk, addresses a new method that uses two spaced Ambisonic microphone arrays. The two arrays capture the sounds coming from the left and right hemispheres. They reported that the sound captured by the new method can realize a better presentation of ambient sound.

The captured 3D sound is usually transmitted, broadcast or recorded after post-production. In broadcasting in Japan, MPEG technology is used for the compression of sound. The second invited paper, “MPEG-H 3D Audio: Immersive Audio Coding” written by J. Herre and S. R. Quackenbush, addresses a new coding method standardized in MPEG. This new paradigm provides tools to code three types of audio format, namely, conventional channel-based audio, object-based audio, and higher-order Ambisonics. In this paper, the architecture of the HPEG-H decoder including rendering methods for the reproduction is

*e-mail: andio@a.memail.jp

1MPEG is the abbreviation of Moving Picture Experts Group, which is the working group of ISO/IEC in charge of the development of international standards for coding of moving pictures, audio, and their combination.
described. Subjective evaluation revealed that this new method can code 3D audio signals with a lower bitrate while maintaining the sound quality.

A state-of-the-art audio system has a large number of channels, for example, 24 channels in 22.2 multichannels. In the reproduction of sound, the downmix method is generally used because it may be difficult to set up such a number of loudspeakers. The third invited paper, “Subjective evaluation of three-dimensional, surround and stereo loudspeaker reproductions using classical music recordings” written by C. Eaton and H. Lee, shows how the impression of 3D audio sound is changed by downmixing. Eaton and Lee downmixed 22-channel (22.2 channels without two low-frequency effect (LFE) channels) musical sound into nine-channel (i.e., five middle-layer and four upper-layer channels), five-channel (5.1 channels without the LFE channel), and two-channel stereo. The original 22-channel sound and downmixed sounds were compared by multiple comparison listening tests in terms of Listener Envelopment, Presence, Overall Tonal Quality, and Overall Listening Experience using four recordings of classical music. As a result, the perceived degradation from 22 channels to nine channels depended on the recording, and five-channel sound was not significantly different from nine-channel sound.

This special issue also includes three Letters. I hope this issue will have a stimulating impact on and provide new insight to a wide range of readers in acoustical fields.

Acknowledgements

I would like to express my sincere thanks to all of my colleagues who were involved in publishing this issue, especially the anonymous reviewers who discussed papers constructively during the review process, as well as the following editors for their generous contributions: Mitsu-nori MIZUMACHI, Takehiro SUGIMOTO, Takanori NISHINO, Toshiharu HORIUCHI, Toru KAMEKAWA and Atsushi MARUI.