Scope of Research on High-quality Audio Signal Processing and Coding

Takehiro Moriya[†], Noboru Harada, and Yutaka Kamamoto

Abstract

Current and future research activities of the Moriya Research Laboratory are introduced. To date, various compression coding technologies for speech and audio have been used for convenient and economical communication systems. However, compression makes the sound quality more band-limited and contaminated with unnatural distortion. We are seeking to construct more comfortable and more convenient communications systems by making full use of the broadband network environment. To achieve this goal, we are focusing on the development of lossless compression coding and exploring new concepts in quality through the use of newly developed devices and our deepening understanding of human perception.

1. Introduction

In the evolution of communications systems, compression has been essential because it allows users to share the limited capacity of communications channels and storage spaces. Various types of compression coding for speech and audio have been developed, and these have found important applications in cellular phone systems, music delivery over networks, and portable players. However, most of the speech coding and audio coding standards in ISO/IEC MPEG* (International Organization for Standardization and International Electrotechnical Commission Moving Picture Experts Group), such as MP3 (MPEG-1 audio layer 3) and AAC (advanced audio coder), achieve a high compression ratio at the sacrifice of minor waveform distortion and band limitation at the decoder.

In the near future, we can expect to enjoy a broadband network at a reasonable price, so greater bandwidth will be available. If we can make use of this rich information environment, we should be able to enhance perceptual quality dramatically. At present, we face the challenge of shifting from the need for efficient compression to the desire for excellent quality. To meet this challenge, we must find ways to exploit the rich bandwidth for higher quality, greater convenience, and more comfort in communications. In this paper, we introduce our current research activities on lossless coding and describe the future challenges for high quality.

2. Current activities on lossless coding

Along with the evolution of the broadband network and digital audio equipment, information rates for delivery and storage have increased rapidly owing to the demands for high-quality audio signal (high sampling rates, high word resolution, and multichannel capability). In the broadband environment, we do not want to lose any quality as a result of data compres-

† NTT Communication Science Laboratories
  Atsugi-shi, 243-0198 Japan
  Email: moriya.takehiro@lab.ntt.co.jp

*1 MPEG is the standardization group in ISO/IEC JTC1 SC29 for video and audio coding. MP3 and AAC are very popular audio coders used for various applications such as the Japanese digital broadcasting system, music delivery over the network, and portable music players.
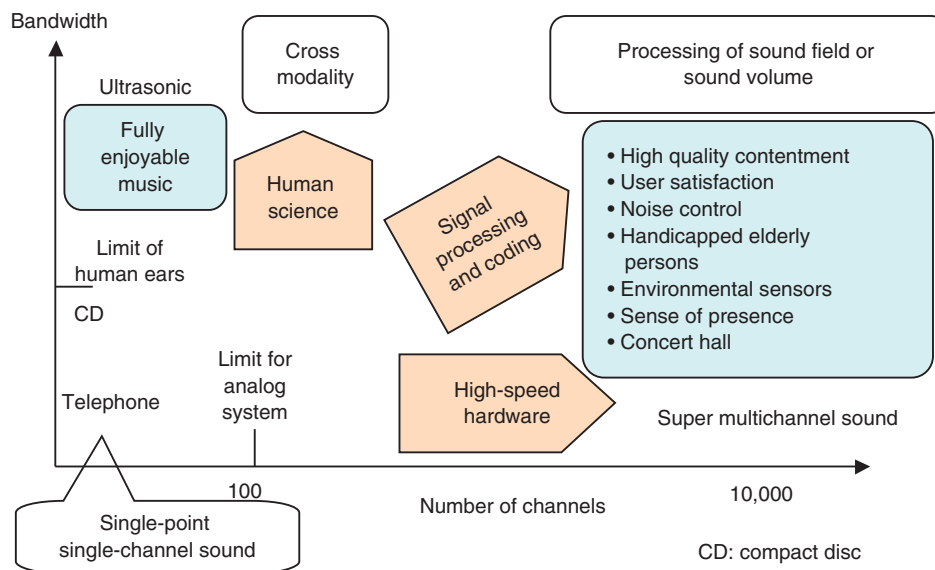http://www.chiariglione.org/mpeg/

Fig. 1.   Extension of research field.

sion. However, as long as the original quality remains unchanged and the processing cost is low, compression will always be useful because the information rates might exceed the available transmission speed or storage capacity.

In this sense, our first endeavor for high-quality coding was the development of a lossless coding scheme that assures perfect reconstruction of the original waveform. This is essential for economically storing or transmitting high-quality signals without any degradation. For interoperability of various applications throughout the world and over time, international standardization is extremely useful. We have continually contributed to the establishment of a lossless coding standard in the MPEG community since 2002. The standard (MPEG-4 ALS) [1] was published in 2006 as part of ISO/IEC 14496-3. Even after the publication of this standard, we continued to make efforts for further improvement of the encoder and for commercialization.

A brief introduction to the lossless coding standard and its basic technologies was given in a letter last year [2]. The second paper in this set of Selected Papers covers our recent advances with the encoder algorithm and an optimized software implementation (speed and compression). The third paper describes application examples and additional standardization activities.

We will continue our efforts to further compress audio signals without losing quality. Compression technology is sometimes dependent on the analysis method or model estimation. Efficient model estimation is also useful for recognition and search tasks. Maintaining a high level of compression technology is also essential for other types of signal processing.

3.   Meeting the challenge of comfortable communication

3.1   Extension of research field

Provided that reliable, high-speed transmission is available, there are various possibilities for achieving more comfortable and higher-quality communications. The final goal is to achieve high-quality human-to-human communications in various environments. An example would be recording the whole audio environment of a live concert, which would enable transmission to distant places immediately or at future times with full fidelity.

We need to extend our research field toward improving the quality and comfort of communications as shown in **Fig. 1**. At present, a single-point-source single-channel band-limited audio signal is used for most communications systems. We want to extend the way that this signal, or information, is used in two ways. One is for human interaction. To explore comfortable communication and the sensation of real presence in music, we need to understand the characteristics of human perception. The other is to significantly increase the number of signal channels (super multichannel capability). There is a huge amount of information hidden in the sound field of a room. We

can make use of new devices and hardware tools to facilitate cost-effective multichannel communications systems.

3.2 Human interaction

To deliver fully enjoyable high-quality music signals to humans, we need a comprehensive understanding of the brain and human behavior. A full understanding of human reactions to ultrasonic signals might enhance our contentment with communication or entertainment, for example [3]. In addition, combining some other modality of sensation with auditory perception (cross modality) might be useful for enhancing the total experience of communication or entertainment.

3.3 Super multichannel sound

The interface to the real environment can be enhanced by introducing a massive number of channels for sound-field control. For this purpose, we need economical high-speed hardware as well as processing and control software. It is impossible to increase the number of channels beyond a few hundred with conventional parallel cable distribution of signals from microphone and loudspeaker arrays. A very promising solution is to use the rapidly developing technologies for high-speed transmission and multiplexing through optical fiber and in small devices. One interesting example is an array of microphones multiplexed in an optical fiber [4]. If a super multichannel sound system can be achieved, it will find general use in various applications such as noise control and environmental sensors.

4. Standardization and alliances

Information systems have become highly complex with huge variations from one system to another. Generally speaking, it is therefore very important for users, manufacturers, and service providers to work together to establish international standards for interoperability and long-term maintenance. Communications systems can be useful only if users or potential users are attracted to them for their convenience and reasonable cost. For this purpose, we will make the necessary efforts to establish standards and support commercialization. For success in these activities, we will need global collaborations or alliances among organizations and companies, keeping in mind the idea that technologies are good only when users can enjoy them.

5. Conclusion

The research activities at the Moriya Research Laboratory include the development of lossless coding and future exploration of human interaction and super multichannel signal processing. All are aimed at the creation of high-quality comfortable communications systems that make use of the rich information available through broadband networks. Our work will be carried out under flexible collaborations with other NTT laboratories in the fields of innovative communication devices and human sciences. In addition, we will continue to promote standardization and alliances, which are important for these new technologies. We hope these technologies will also contribute to other research fields besides the acoustical signal processing field.

References

[1] ISO/IEC 14496-3:2005/Amd.2:2006, Information technology—Coding of audio-visual objects—Part 3: Audio, Amendment 2: Audio Lossless Coding (ALS), new audio profiles and BSAC extensions, edition 2006-03-15.
[2] T. Moriya, N. Harada, Y. Kamamoto, and H. Sekigawa, "MPEG-4 ALS—International Standard for Lossless Audio Coding," NTT Technical Review, Vol. 4, No. 8, pp. 40–45, 2006.
[3] T. Oohashi, E. Nishina, M. Honda, Y. Yonekura, Y. Fuwamoto, N. Kawai, T. Maekawa, S. Nakamura, H. Fukuyama, and H. Shibasaki, "Inaudible High-Frequency Sounds Affect Brain Activity: Hypersonic (*sic*) Effect," Journal of Neurophysiology, Vol. 83, pp. 3548–3558, 2000.
[4] T. Fujise, K. Nakamura, and S. Ueha, "Demodulation of Acoustic Signals in Fiber Bragg Grating Ultrasonic Sensors Using Arrayed Waveguide Gratings," Jpn. J. of Appl. Phys., Vol. 45, No. 5B, pp. 4577–4579, 2006.

Takehiro Moriya
Research Fellow, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received the B.S., M.S., and Ph.D. degrees all in applied mathematics and instrumentation physics from the University of Tokyo, Tokyo, in 1978, 1980, and 1989, respectively. Since joining the Musashino Electrical Communication Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT) in 1980, he has been engaged in research on and the standardization of speech and audio coding. In 1989, he stayed at AT&T Bell Laboratories as a guest researcher. He is a member of the Acoustical Society of Japan (ASJ), the Information Processing Society of Japan (IPSJ), and the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan and a fellow of IEEE.

Yutaka Kamamoto
Researcher, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received the B.S. degree in applied physics from Keio University, Kanagawa, in 2003 and the M.S. degree in information physics and computing from the University of Tokyo, Tokyo, in 2005. Since joining NTT Communication Science Laboratories in 2005, he has been studying signal processing and information theory. He is a member of ASJ, IPSJ, the Society of Information Theory and its Applications, IEICE, and IEEE.

Noboru Harada
Research Scientist, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received the B.S. and M.S. degrees from the Department of Computer Science and Systems Engineering of Kyushu Institute of Technology, Fukuoka, in 1995 and 1997, respectively. He joined NTT Human Interface Laboratories in 1997. His main research area has been lossless audio coding and high-efficiency coding of speech and audio. He is a member of ASJ, IEICE, the Audio Engineering Society, and IEEE.