MPEG-4 ALS: Performance, Applications, and Related Standardization Activities

Noboru Harada†, Takehiro Moriya,
and Yutaka Kamamoto

Abstract

MPEG published a new lossless audio coding scheme called MPEG-4 Audio Lossless Coding (ALS) in March 2006. This paper overviews MPEG-4 ALS and its possible applications. In addition, it introduces an audio archiving tool that uses MPEG-4 ALS as the encoding engine. This archiving tool offers excellent compression performance and functionality for handling several audio files as one archived file. Test results show that its compression performance, in terms of compression ratio and required processing time, is much better than that of ZIP when the input data is audio files. The tool's application format has been proposed to MPEG-A and is now being discussed as one of the multimedia application formats (MAFs) under the name MPEG-A Professional Archival MAF. Other recent standardization activities are also briefly mentioned.

1. Introduction

With the publication of a new lossless audio coding scheme called MPEG-4 Audio Lossless Coding (ALS) in March 2006 [1], the MPEG standard now has efficient lossless encoding tools for audio data files. There is a strong demand for music in formats that provide superior audio quality such as CD (compact disc) quality. If the audio signal is recorded with high resolution, compressing the data using a lossy encoding scheme is of no use because some information at higher frequencies will be lost. A lossless encoding scheme must be used to compress high-resolution audio data. Such a lossless compression format can be used for Internet distribution of audio files (e.g., streaming or downloading of music from online music stores), as an audio format for portable music players, and for other high-resolution disc formats such as Blu-ray and HD-DVD.

On the other hand, master tapes or original pressings of historical analog recordings are in the process of deteriorating, and it is becoming impossible to listen to them in their original form because very few of the players are still available and some no longer exist. Analog recorded audio signals should be digitized with the highest possible sampling ratio and the highest quantized bit depth. Such signals should be compressed by a standardized lossless compression scheme.

In addition, in recent years, it has become popular to store professional-level recordings as a set of files that contain not only audio tracks but also meta-information files and other non-audio files, such as plug-in binaries, notes, and picture images, in a hierarchical folder structure. Such a set of files is called a recording project. With the increasing popularity of high-resolution audio contents, the total size of the intermediate files for an audio recording project is getting larger. For example, the total size of the audio data sampled at 192 kHz with 24-bit resolution is four times larger than data sampled at 48 kHz. The dependence of audio data size for three minutes of music on sampling frequency and bit resolution is shown in **Fig. 1**. A standardized professional archival format is strongly demanded for the storage, delivery, and preservation of recorded digital contents.

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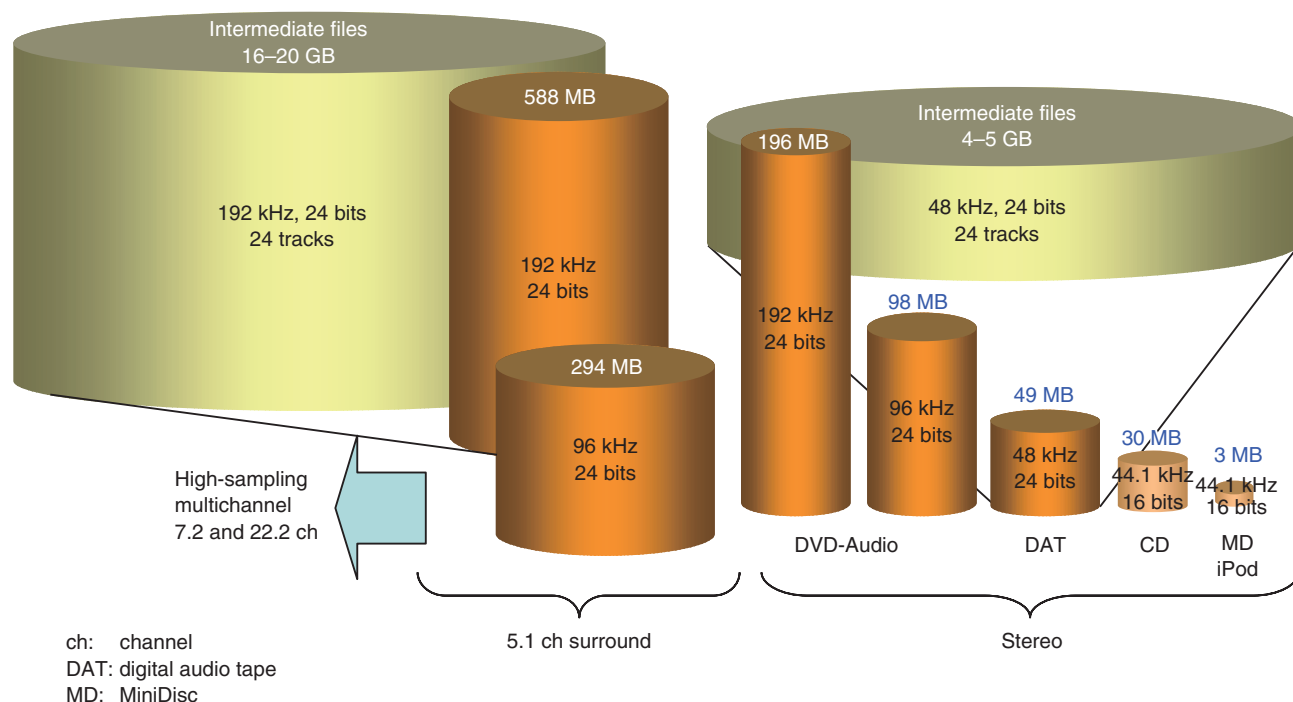 Atsugi-shi, 243-0198 Japan
  Email: harada.noboru@lab.ntt.co.jp

Fig. 1.   Comparison of data sizes for three minutes of music.

For archival application, users often want to compress audio data files together with the whole folder structure. Gnu Zip [2] or WinZip [3] is widely used for this purpose, but, unfortunately, the compression performance is not high enough for raw audio data files. Here, as a way to handle several audio files as one archived file, we describe an audio archiving tool that uses MPEG-4 ALS for the encoding engine.

Section 2 overviews MPEG-4 ALS and its possible applications. Section 3 introduces our MPEG-4 ALS-based archiving tool, describes its possible applications, and presents experimental test results. Section 4 covers MPEG's latest standardization activities concerning MPEG-A Professional Archival MAF and section 5 mentions other standardization activities. Finally, section 6 makes some concluding remarks.

2.   MPEG-4 ALS

2.1   Overview

MPEG-4 ALS is an extension of the MPEG-4 audio coding family for lossless compression of audio data [1], [4]–[14]. The ALS core codec is based on forward-adaptive linear prediction, which offers remarkable compression performance with low complexity. Additional features include long-term prediction,

multichannel coding, and compression of floating-point audio material [15], [16]. ALS also offers good flexibility in terms of the compression-complexity tradeoff, ranging from very low-complexity implementations to maximum compression modes, so it is adaptable to different requirements. Its many other features include:

-General support for virtually any uncompressed digital audio format, including the Sony Wave64 file format and Broadcast Wave Format (BWF) [19]–[21].

-Support for linear PCM (pulse code modulation) resolutions of up to 32 bits at arbitrary sampling rates.

-Multichannel/multitrack support for up to 65,536 channels, including 5.1-, 7.1-, and 22.2 channel surround sound.

-Support for audio data in the IEEE 754 32-bit floating-point audio format

-Quick random access to encoded data

-Support for the MP4 file format, which allows multiplexing and synchronization with video [22].

Input formats supported by MPEG-4 ALS are listed in **Table 1**.

Table 1.   Input formats supported by MPEG-4 ALS.

Sampling frequency	Currently tested up to 192 kHz (higher frequencies, such as 384 kHz, can be handled with the current specifications.)
Bit depth	PCM (up to 32 bit), IEEE754 32 bit floating-point
No. of channels	Up to 65,536
File formats	Raw, WAVE, AIFF, BWF, Sony Wave64, BWF with RF64

Note: Container format for the encoded bitstream is MP4ff.

2.2   Possible applications

Lossless audio coding in general and MPEG-4 ALS in particular have many applications, at both the professional and consumer levels. These include:

- Internet distribution of audio files (streaming, online music stores, and downloading)
- High-resolution disc format
- Portable music players
- Archival systems (for broadcasting, studios, record labels, and digital transfer)
- Studio operations (for storage, collaborative working, digital backup, and digital transfer)

In these applications, MPEG-4 ALS is used as an encoding engine for the lossless compression of the audio data. For the archival systems and studio operations, the archiving tool based on MPEG-4 ALS described in section 3 will be more suitable.

2.3   Performance

The specifications for the coded bitstream of MPEG-4 ALS and its decoding scheme have been established, but there is still some room for improving coding efficiencies with respect to compression performance and encoding and decoding times, which should be done without losing the decoding compatibility with the standard.

We have proposed several fast encoding algorithms for MPEG-4 ALS [23]–[24] and developed an efficient encoding and decoding tool. The improved encoding and decoding software developed by NTT for MPEG-4 ALS, called "MPEG-4 ALS fast", can encode and decode very efficiently in terms of both processing speed and compression ratio. It has been shown that it can encode and decode much faster than the publicly available reference software for MPEG-4 ALS and faster than any other lossless coding software.

In tests, two implementations of an MPEG-4 ALS codec, MPEG-4 ALS RM18 [25] and MPEG-4ALS fast, were compared with three of the most popular programs for lossless audio compression: FLAC (version 1.1.2) [26], Monkey's Audio (MAC 4.01) [27], and OptimFrog (OptimFROG v4.520b1) [28]. The tests were conducted on a 2.39-GHz AMD Opteron processor 250 with 2 GB of RAM (random access memory). The data for the tests was acquired from the standard set of audio sequences for MPEG-4 lossless coding originally donated by Panasonic Corporation. We used a total of 51 stereo waveform files with sampling rates of 48 kHz and resolution of 16 bits and with sampling rates of 48, 96, and 192 kHz and a resolution of 24 bits. The duration of each waveform was 30 seconds. The total size of the 51 files was 682,562,430 bytes.

The results for the tested encoders and decoders are shown in **Figs. 2** and **3**. The vertical axis shows the compression ratio, defined as

$$C = \frac{Compressed\ file\ size}{Original\ file\ size} \times 100\%,$$

where smaller values mean better compression. The horizontal axis shows the average processing time for encoding or decoding a 30-s file. Smaller values mean faster processing.

The results show that MPEG-4 ALS RM18 and MPEG-4 ALS fast can achieve better performance than other lossless codecs in terms of the best balance of compression performance and encoding/decoding speed, with MPEG-4 ALS fast offering the best performance. The compression ratio was 7 to 8% better than that of FLAC even when the encoding/decoding speed was twice as fast. The results also show that the CPU (central processing unit) load for realtime encoding and decoding using MPEG-4 ALS fast was approximately 1 to 2% that of the Opteron Processor 250 when the fastest operating mode was used.

3.   Archiving tool based on MPEG-4 ALS

3.1   Possible applications

(1) Digital backup and direct delivery of recorded music projects

Nashville members of the P&E Wing of The Recording Academy have formed a Delivery Specifications Committee, which has created the Delivery Recommendations for Master Recordings document
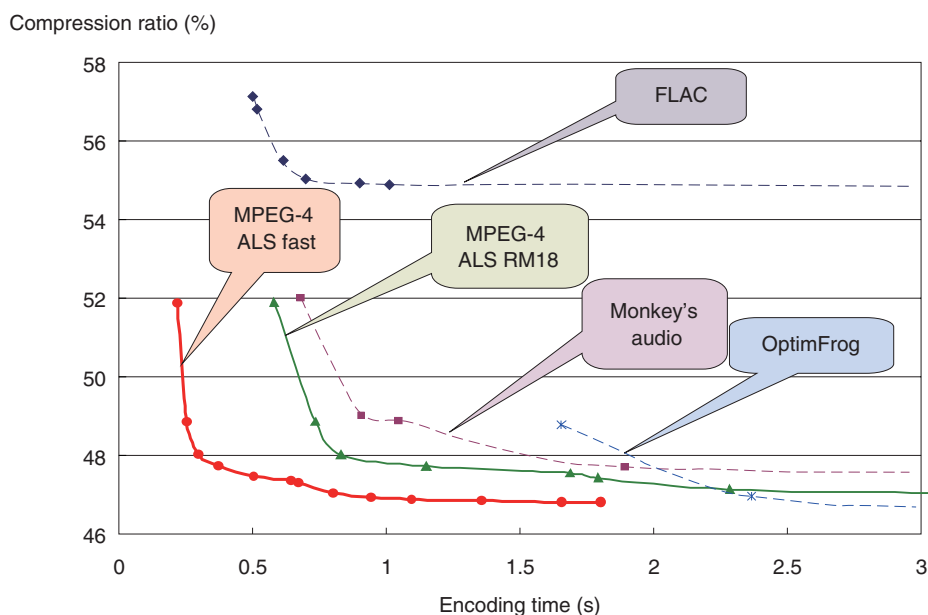
Compression ratio (%)

Fig. 2.   Performance of the encoder.
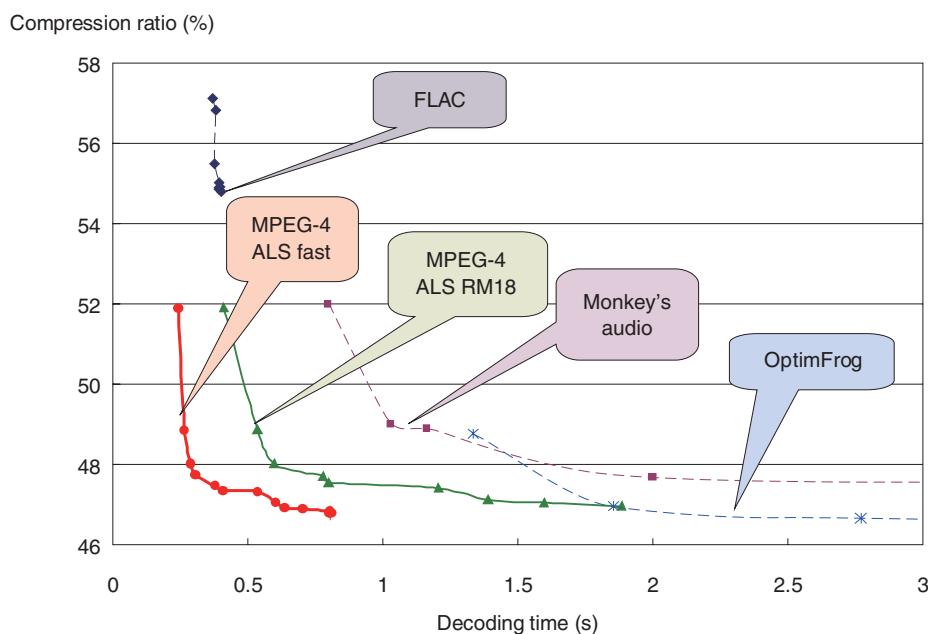
Compression ratio (%)

Fig. 3.   Performance of the decoder.

[29]. In this document, the committee states that it expects direct delivery (via secure connection on the Internet) to be commonplace in the future and recommends uploading files to very-large-scale digital libraries. The preferred delivery of a recorded music project would include Broadcast Wave Files (BWF) [20]–[22] of every multitrack and two-track element,

without processing or automation. All of the audio tracks should be *flattened*, which means converting small recorded segments into a contiguous sound file padded with silence between the segments, and migrated to the BWF format with a maximum of one channel per BWF. In addition to BWFs, digitized versions of usual documentation (tracking sheets, engi-
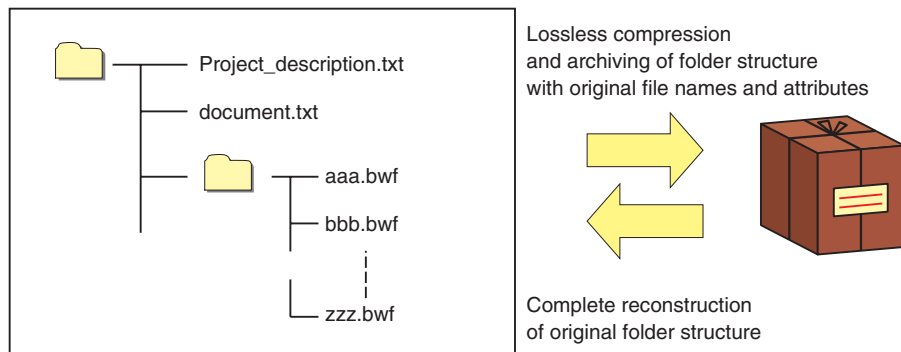
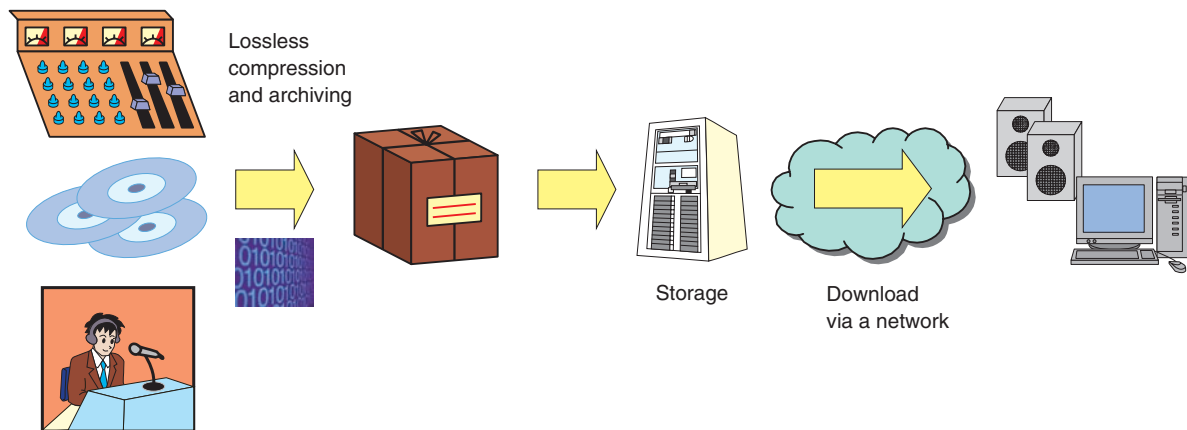Fig. 4.   Example of a folder image and an archive file.

Fig. 5.   Example of audio archiving tool applications.

neer's notes, setup notes, sketches of microphone placement, recording map documents, lyrics, charts, orchestral arrangements, and parts-and-mix documentation) and any other data pertinent to the recording project should be included in the delivery files.

Folder-structure-based archiving is suitable for the direct delivery of a recorded music project. It is natural that all files related to the recorded music project should be kept in an appropriate folder. The files in the folder should be compressed and archived together in a single file (see **Fig. 4**). This would offer a high degree of safety because it can prevent users from miscopying a file during the delivery process. In addition, when uploading or downloading a project by FTP (file transfer protocol), it is much easier to transmit a single archive file than it is to transmit many individual BWF files (see **Fig. 5**).

(2) Compression and archiving of intermediate data generated by sound editing tools

Many professional sound editing tools, such as Pro-

Tools from DigiDesign [30] or Nuendo from Steinberg [31], keep work files in folders. Raw waveform data from separate tracks is stored in separate files with a specific file format, such as .wav or .aiff, and those files are kept in one or more folders. The set of files in the folder for a song consisting of the waveform files of all tracks is called a project. A project contains the project information file and individual audio track files.

During editing operations, compression and archiving tools that can compress all files in a folder with the folder structure into a single file would be very advantageous for this application. Users could keep a snapshot of a version of edited files in an archive file, so that they could roll back to the previous version of the edits if desired. Sometimes, a target file may be a non-audio data file or may already be encoded with a lossy compression tool, such as MP3/AAC. In that case, the file should simply be added to the archive file as is.
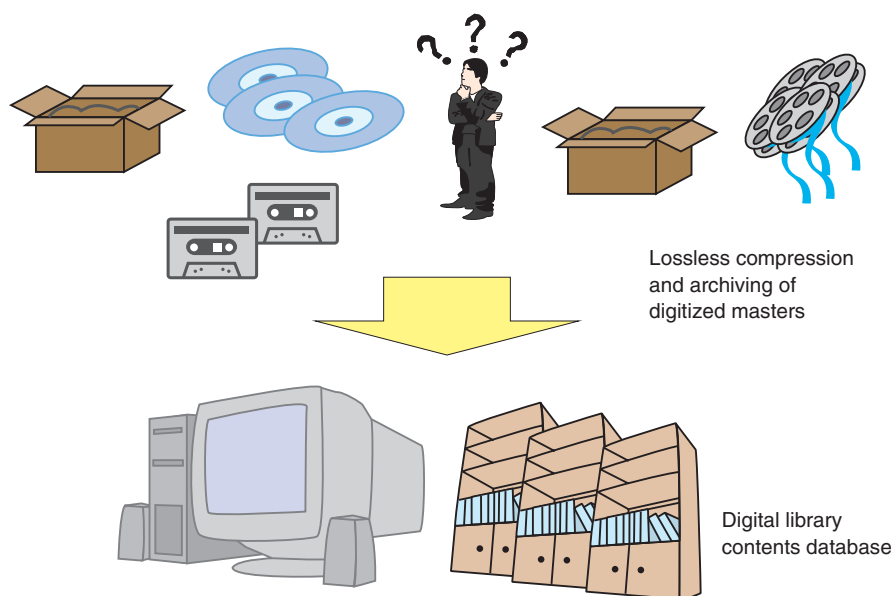
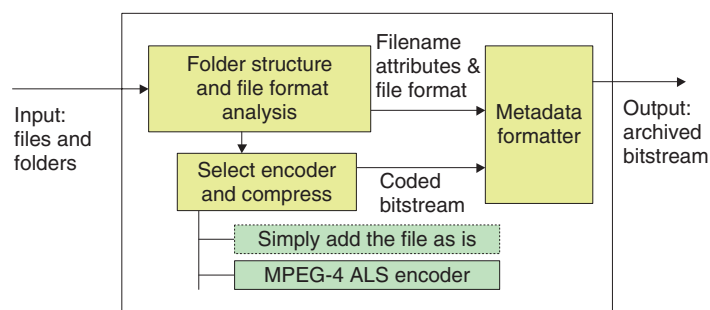Fig. 6.   Another example of audio archiving tool applications.

Fig. 7.   Overview of an encoder for the audio archiving tool.

(3) Preservation/archiving

Lossless compression of files is becoming very popular because it reduces the demand on storage media for bitwise-exact copies of digitized masters (**Fig. 6**). All files related to the content are archived together in a single archive. The relationships among the files may sometimes be tight, but sometimes not. Folder-structure-based archiving is more suitable for archiving files of this kind.

(4) Packaging file format for non-audio wave-
forms

In addition to audio data, MPEG-4 ALS can also efficiently compress non-audio multichannel signals, such as biomedical (e.g., electroencephalography (EEG) or magnetoencephalography (MEG)) and seismic data [15], [17]. For example, 512-channel MEG data can be compressed to about 15 to 40% of its original size. This archival application would be used for packaging such non-audio waveform data together with other related information.

3.2   Overview of the encoder and decoder for the archiving tool

Overviews of the encoder and decoder for our audio archiving tool are shown in **Figs. 7** and **8**. If the target file contains PCM audio data, it is encoded with the MPEG-4 ALS encoder for lossless compression. Audio data files, such as .wav and .aiff files, are compressed with MPEG-4 ALS and the encoded bitstreams are added to the archive file with file attributes, such as file names, read/write attributes and folder structure information. If the target file is an audio file that has already been encoded with a lossy compression tool, such as AAC (advanced audio coding), or if it is a non-audio file, it is simply added to the archive file as is without any further compres-
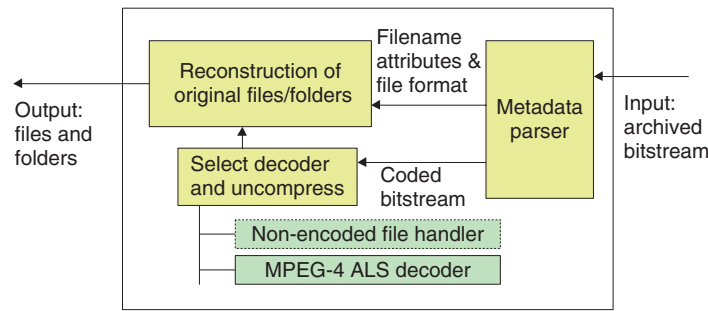
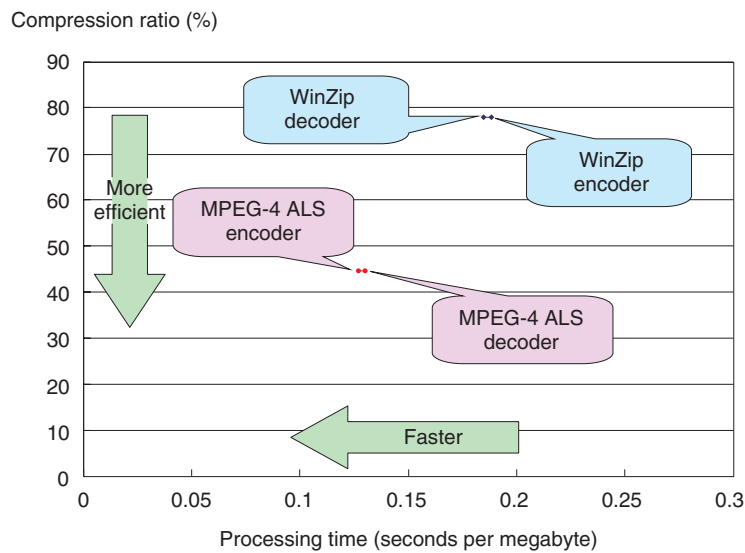Fig. 8.   Overview of a decoder for the audio archiving tool.

Fig. 9.   Compression performance of our archiving tool.

sion.

3.3   Experimental results

The compression performance of our archiving tool was tested with a set of files totaling 242,805,050 bytes. A folder containing the files was compressed and archived with our archiving tool and with Win-ZIP. The folder contained 82 audio files, 40 image files, and two project information files (binary files) of the audio editing tool. The test was conducted on a personal computer with a Pentium M (1.6 GHz) CPU.

The encoding and decoding times of our tool and those of WinZIP for compressing and extracting the song files are shown in **Fig. 9**. The vertical axis shows the compression ratio, defined as

$$C = \frac{Compressed\ data\ size}{Original\ data\ size} \times 100\%,$$

where smaller values mean better compression. The horizontal axis shows the average processing time for encoding or decoding a unit input data size of 1 MB. Smaller values of the average processing time mean faster processing. As shown in Fig. 9, the compression performance of our archiving tool was much better than that of WinZIP when the input data was audio files. Even though the compression ratio of our archiving tool for audio data was much higher than that of WinZIP, its encoding and decoding speed was slightly higher.

4.   Standardization of MPEG-A Professional Archival MAF

4.1   MPEG-A: MAFs

MPEG's multimedia application formats (MAFs) [32] provide a framework for integrating elements from several MPEG standards into a single specifica-
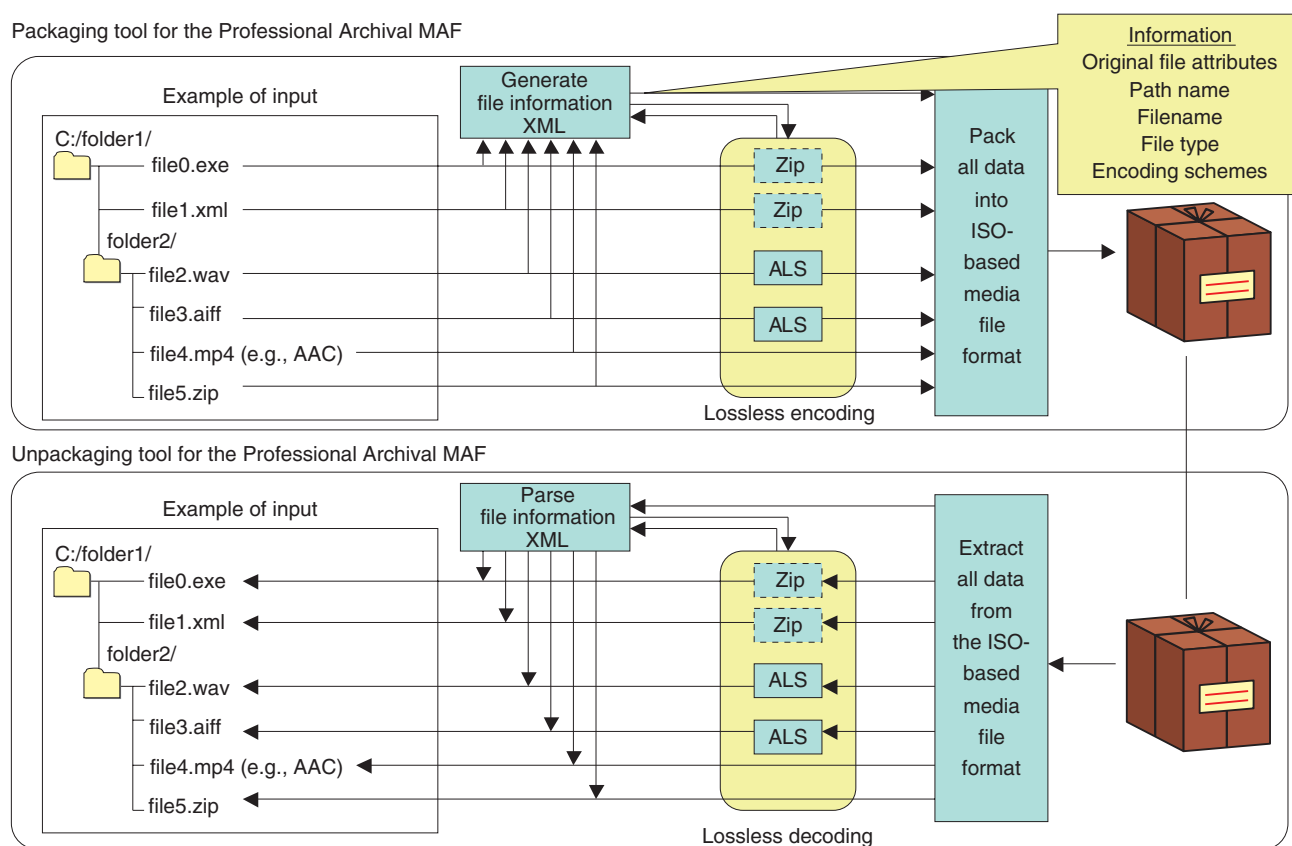
Packaging tool for the Professional Archival MAF

Example of input

Generate file information XML

Information
Original file attributes
Path name
Filename
File type
Encoding schemes

C:/folder1/
file0.exe
file1.xml
folder2/
file2.wav
file3.aiff
file4.mp4 (e.g., AAC)
file5.zip

Zip
Zip
ALS
ALS

Pack all data into ISO-based media file format

Lossless encoding

Unpackaging tool for the Professional Archival MAF

Example of input

Parse file information XML

C:/folder1/
file0.exe
file1.xml
folder2/
file2.wav
file3.aiff
file4.mp4 (e.g., AAC)
file5.zip

Zip
Zip
ALS
ALS

Extract all data from the ISO-based media file format

Lossless decoding

Fig. 10.   Example packaging and unpackaging tools.

tion that is suitable for specific but widely usable applications. For example, MAFs specify how to combine metadata with timed media information for a presentation in a well-defined format that facilitates interchange, management, editing, and presentation of the media. The presentation may be 'local' to the system or may be accessible via a network or other stream delivery mechanism. Selected MAFs are expected to become parts of the ISO/IEC 23000 specification. The 23000 series is also called MPEG-A (where "A" stands for application).

4.2   Professional Archival MAF

To archive files containing audio tracks along with meta-information files and picture images in a folder with a hieratical folder structure, we have designed an archiving file format that makes use of MPEG-4 ALS as the encoding engine and proposed it to the standardization committee. We are now working to develop the proposed format into an MAF standard called the MPEG-A Professional Archival MAF. This is intended to be a standardized solution that enables sustainability, accessibility, and playability of digital content, making full use of MPEG tools for maximum interoperability. As such, it aims to preserve our digital heritage and also provide for improved preservation of pre-digital content.

4.3   Technical requirements

In order to generalize the requirements and expand the scope to include audio, video, picture images, and other multimedia data, MPEG issued a call for requirements for the Professional Archival MAF [33]. Examples of packaging and unpackaging tools for the Professional Archival MAF based on the current specification are shown in **Fig. 10**.

5.   Other standardization activities related to MPEG-4 ALS

Several ongoing standardization activities related to MPEG-4 ALS are seeking to enhance its utility and appeal. An overview of these activities is shown in **Fig. 11**.

At IEC TC100, we proposed a new work item for IEC 61937, called "Digital audio—Interface for non-
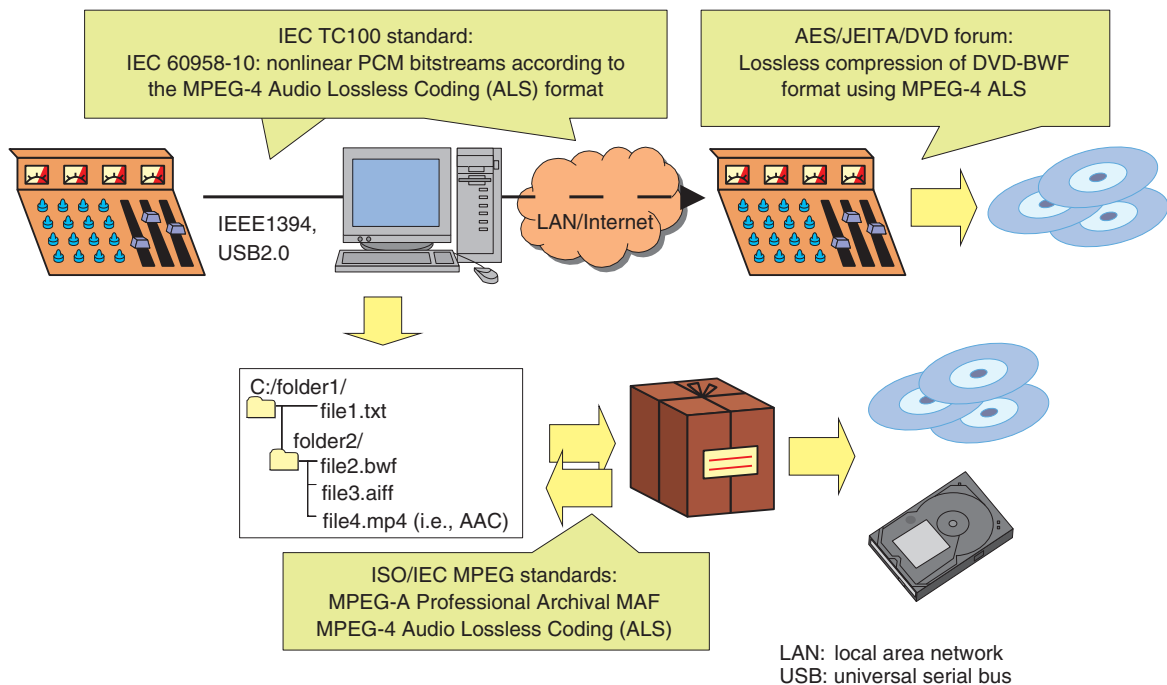
Fig. 11.   Standardization activities related to MPEG-4 ALS.

linear PCM-encoded audio bitstreams applying IEC 60958 Part 10: Nonlinear PCM bitstreams according to the MPEG-4 ALS format". The purpose of this specification is to transmit MPEG-4 ALS encoded bitstreams via the IEC 60958 interface, which is well known as the S/PDIF (Sony/Philips digital interconnect format) digital interface specification. There is a strong need for the capability to store, preserve, and transmit high-resolution audio signals in a cost efficient way using a standardized lossless compression scheme. The development an interface standard for the transmission of MPEG-4 ALS encoded bitstreams among various professional or high-end consumer equipment is expected.

6.   Conclusion

We have shown several possible applications of MPEG-4 ALS and described the performance of our latest implementation encoder and decoder for it. We described a new tool based on MPEG-4 ALS for archiving audio data. Test results show that its compression performance is much better than that of ZIP. The tool's application format has been proposed to MPEG-A and is now being discussed as one of the MAFs under development. MPEG 4 ALS and its related standards are expected to be very useful tools in both professional and consumer applications.

References

[1] ISO/IEC 14496-3:2005/Amd.2:2006, Information technology—Coding of audio-visual objects—Part 3: Audio, 3rd Ed. Amendment 2: Audio Lossless Coding (ALS), new audio profiles and BSAC extensions, 2006.
[2] "GNU Zip," http://www.gzip.org
[3] "WinZip," http://www.winzip.com
[4] "Call for proposals on MPEG-4 lossless audio coding," ISO/IEC JTC 1/SC29/WG11 N5040, Klagenfurt, Austria, 2002.
[5] "Revised call for proposals on MPEG-4 lossless audio coding," ISO/IEC JTC 1/SC29/WG11 N5208, Shanghai, China, 2002.
[6] ISO/IEC 14496-3:2005, Information technology—Coding of audio-visual objects—Part 3: Audio, 2005.
[7] T. Moriya, D. Yang, and T. Liebchen, "Extended Linear Prediction Tools for Lossless Audio Coding," Proc. of ICASSP 2004, pp. III-1008–1011, 2004.
[8] T. Liebchen and Y. A. Reznik, "MPEG-4 ALS: an emerging standard for lossless audio coding," Proc. of Data Compression Conference (DCC) 2004, pp. 439–448, Utah, USA, 2004.
[9] T. Liebchen, "An introduction to MPEG-4 audio lossless coding," Proc. of IEEE ICASSP, Montreal, USA, 2004.
[10] T. Moriya, T. Liebchen, Y. A. Reznik, and D. Yang, "MPEG-4 audio lossless coding," in Preprint AES 116th Convention, #6047, Berlin, Germany, 2004.
[11] T. Liebchen and Y. A. Reznik, "Improved forward-adaptive prediction for MPEG-4 audio lossless coding," in Preprint AES 118th Convention, #6449, Barcelona, Spain, 2005.
[12] N. Harada, T. Moriya, H. Sekigawa, and K. Shirayanagi, "Lossless compression of IEEE floating-point audio using approximate common factor coding and masked-LZ compression," in Preprint AES 118th Convention, #6352, Barcelona, Spain, 2005.
[13] T. Moriya, N. Harada, and Y. Kamamoto, "An enhanced encoder for the MPEG-4 ALS lossless coding standard," in Preprint AES 121st Convention, #6869, San Francisco, USA, 2006.
[14] T. Liebchen, T. Moriya, N. Harada, Y. Kamamoto, and Y. A. Reznik,

"The MPEG-4 audio lossless coding (ALS) standard—technology and applications," in Preprint AES 119th Convention, #6589, NY, USA, 2005.

[15] Y. Kamamoto, T. Moriya, T. Nishimoto, and S. Sagayama, "Lossless compression of multi-channel signals based on inter-channel correlation," IPSJ Trans., Vol. 46, No. 5, pp. 1118–1128, 2005 (in Japanese).

[16] N. Harada, T. Moriya, H. Sekigawa, K. Shirayanagi, and Y. Kamamoto, "Lossless Compression of IEEE754 Floating-point Signal in ISO/IEC MPEG-4 Audio Lossless Coding (ALS)," IEICE Trans. on Communications, Vol. J89-B, No. 2, pp. 204–213, 2006 (in Japanese).

[17] Y. Kamamoto, T. Moriya, N. Harada, N. Nishimoto, and S. Sagayama, "Intra- and Inter-Channel Long-Term Prediction in ISO/IEC MPEG-4 Audio Lossless Coding (ALS)," IEICE Trans. on Communications, Vol. J89-B, No. 2, pp. 214–222, 2006 (in Japanese).

[18] T. Moriya, N. Harada, and Y. Kamamoto, "Performance-complexity tradeoffs of the MPEG-4 ALS lossless coding standard," Proc. IEEE 40th Asilomar Conference on Signals, Systems and Computers, WA7a-4, pp. 2130–2134, 2006.

[19] "The Broadcast Wave Format: A format for audio data files in broadcasting,"
http://www.ebu.ch/CMSimages/en/tec_text_n22-1997_tcm6-4645.pdf

[20] "New File Format and Methods for Multichannel Sound in Broadcasting,"
http://www.sr.se/utveckling/tu/bwf/prog/RF_64v1_4.pdf

[21] "The latest version of the RF64 draft specification,"
http://www.sr.se/utveckling/tu/bwf/

[22] ISO/IEC 14496-14:2003, Information technology—Coding of audio-visual objects—Part 14: MP4 file format, 2003.

[23] T. Moriya, N. Harada, and Y. Kamamoto, "An enhanced encoder for the MPEG-4 ALS Lossless Coding standard," in Preprint AES 121st Convention, #6809, San Francisco, USA, 2006.

[24] N. Harada, T. Moriya, and Y. Kamamoto, "An audio archiving format based on MPEG-4 Audio Lossless Coding," in Preprint AES 121st Convention, #6895, San Francisco, USA, 2006.

[25] "MPEG-4 ALS Reference Software,"
http://www.nue.tu-berlin.de/forschung/projekte/ lossless/mp4alsRM18.zip

[26] "FLAC—Free Lossless Audio Codec," http://flac.sourceforge.net

[27] "Monkey's Audio," http://www.monkeysaudio.com

[28] "OptimFROG Lossless Audio Compression," http://www.losslessaudio.org

[29] "The Delivery Recommendations for Master Recordings," P&E Wing Delivery Recommendations 030609. 31 revision:
http://www.grammy.com/PDFs/Recording_Academy/Producers_And_Engineers/DeliveryRecs.pdf

[30] "ProTools," http://www.digidesign.com

[31] "Nuendo," http://www.steinberg.net

[32] ISO/IEC JTC1/SC29/WG11/N8781, "MAF Overview," 79th MPEG Meeting, Marrakech, 2007.

[33] ISO/IEC JTC1/SC29/WG11/N9165, "Call for Requirements for Professional Archival Multimedia Application Format," 81st MPEG Meeting, Lausanne, 2007.

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 the Acoustical Society of Japan (ASJ), the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan, the Audio Engineering Society, and IEEE.

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 ASJ, the Information Processing Society of Japan (IPSJ), and IEICE 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.