

MPEG-4 ALSの性能・応用と関連する標準化活動

音響信号の可逆圧縮符号化方式MPEG-4 ALSの概要とそのアプリケーションを紹介します。MPEG-4 ALSを応用したアーカイブツールを提案し、ZIP方式よりも圧縮性能・処理時間の点で優れていることを示します。MPEG-Aプロ用アーカイブフォーマットなど、関連する標準化活動についても紹介します。

はらだ のぼる もりや たけひろ
原田 登 / 守谷 健弘
 かまもと ゆたか
鎌本 優

NTTコミュニケーション科学基礎研究所

はじめに

CDと同等以上の高品質なフォーマットで音楽を楽しみたいという要求を背景に、音響信号の可逆圧縮符号化ツールであるMPEG-4 Audio Lossless Coding (ALS) が標準化され、2006年3月に出版されました⁽¹⁾。高分解能で録音された音響信号の場合には、高周波成分が欠落し、非可逆の符号化方式で圧縮することは適さないため、可逆の圧縮符号化方式を用いる必要があります。

そのような音源データのインターネット配信（例えばストリーミングやオンライン・ミュージックストアでのダウンロード販売）、携帯ミュージックプレーヤのためのファイルフォーマットや、Blu-rayやHD-DVD等のディスクフォーマットのための圧縮方式として可逆符号化方式を用いることができます。

一方、歴史的録音音源のマスターテープや原版金型は、再生に適さない状態にまで劣化したり、再生機器自体がなくなってしまうなどの問題に直面しています。アナログレコードに記録された音響信号は、可能な限り高いサンプリング周波数で、量子化ビット

数も高く設定してデジタル信号に変換する必要があります。このような信号は、可逆符号化方式で圧縮することが望ましいのです。

加えて、近年の音楽製作現場では、録音されたオーディオ情報だけでなく、メタ情報ファイルやプラグイン・バイナリ、楽譜の情報、画像情報など、音響信号以外のファイルを含めて階層的なフォルダ構造に保持するような場合が多く、このようなファイルの集まりを、レコーディングプロジェクトファイルと呼びます。高品質オーディオコン

テンツの普及に伴い、編集作業に用いる中間データの総量は増加する一方で、3分の曲、1曲に必要なデータサイズを比較した例を図1に示します。例えば、サンプリング周波数192 kHz、24 bitで量子化された音響信号は、48 kHzサンプリングのものに比べて4倍のデータサイズとなります。

録音されたコンテンツの保存・蓄積、伝送・配信のために、標準化されたプロ用途のアーカイブフォーマットが必要とされています。複数のファイルを1つにまとめて圧縮するような用途で

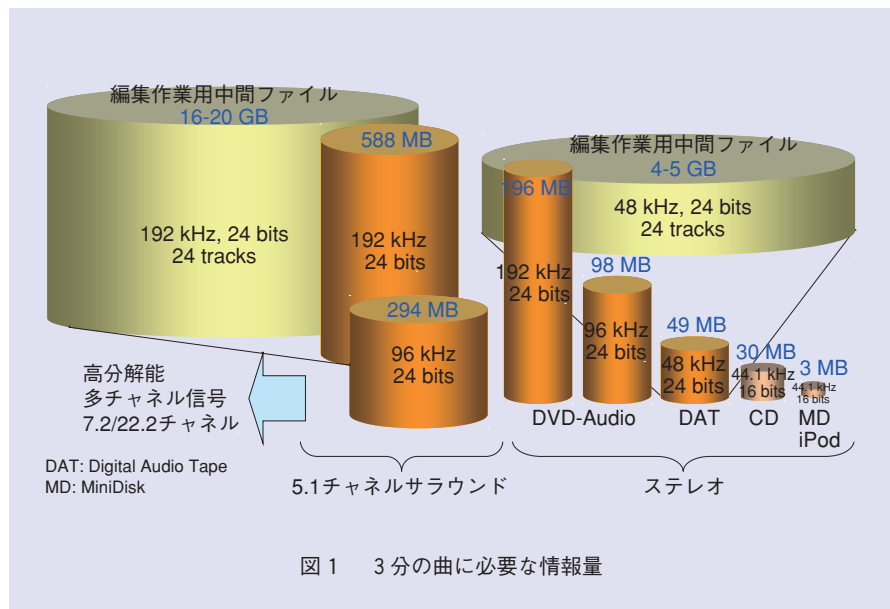


図1 3分の曲に必要な情報量

は、GnuZIPやWinZIPといったツールが用いられる場合が多いのです。しかし、音響信号データに対して圧縮性能がよくないという問題があります。これに対してMPEG-4 ALSを圧縮符号化に用いることで、効率のよいアーカイブツールを実現することができることを示します。

次の項では国際標準であるMPEG-4 ALSの概要と、可能な応用について示し、続いてMPEG-4 ALSを応用したアーカイブツールとその性能、および、適用分野、MPEG-Aで標準化中のProfessional Archival Application Format、その他の関連する標準化動向について紹介し、最後に結論を示します。

MPEG-4 ALS

■概要

MPEG-4 ALSはMPEG-4 オーディオ符号化標準の1つで、可逆の圧縮符号化方式です^{(1)~(5)}。ALSの符号化・復号化アルゴリズムは、線形予測符号化を基礎とし、長期予測、多チャンネル符号化、浮動小数点信号符号化^{(6)~(8)}などのツール群を加えることによって低演算量で高い圧縮性能を実現しています。ALSは非常に低演算量のモードから最大の圧縮性能を得るモードまで、圧縮性能と処理量の間の特長を考慮して要求条件に合わせて符号化パラメータを選択することが可能です。

機能面では、次のような特徴があり

ます。

- ・ Sony Wave64ファイルやBroadcast Wave Format (BWF) 等を含む、さまざまな種類の入力フォーマットに対応
- ・ 量子化ビット数最大32ビットまでのPCMファイル入力に対応
- ・ 5.1チャンネル、7.2チャンネル、22.2チャンネルサラウンドを含む最大65 536チャンネルまでのマルチチャンネル信号入力に対応
- ・ IEEE 754 32ビット浮動小数点形式の入力信号に対応
- ・ ランダム・アクセス
- ・ MP4ファイルフォーマット⁽¹¹⁾を用いて、映像符号化方式と組み合わせ利用可能

MPEG-4 ALSで対応可能な入力フォーマットを表に示します。

■アプリケーション

音響信号の可逆圧縮符号化方式MPEG-4 ALSは、プロ用途から一般用途まで、次に示すようにさまざまなレベルの応用が考えられます。

- ・ オーディオファイルのインターネット配信（ストリーミング、オンラインミュージックストア、ダウンロード）
- ・ 高品質ディスク・フォーマット
- ・ 携帯ミュージック・プレーヤ
- ・ アーカイブシステム（放送・スタジオ・レコード、デジタル配信・伝送）
- ・ スタジオ編集（保存・蓄積、伝

送、協調作業）

これらの用途においてMPEG-4 ALSを用いて音響信号を可逆圧縮符号化することができます。特にアーカイブシステムやスタジオ編集用途では、後述するMPEG-4 ALSを応用したアーカイブフォーマットを用いることができます。

■性能評価

MPEG-4 ALSのビットストリームおよび復号手順については標準により規定されていますが、復号の互換性が保たれる範囲で符号化アルゴリズムを改善し、圧縮性能や符号化・復号化処理に要する時間等の性能を改善することができます。

我々は、これまでにMPEG-4 ALSの符号化性能を改善するいくつかの高速アルゴリズムを提案し^{(9)・(10)}、それらを用いて符号化・復号化処理速度や圧縮率を改善したNTT独自開発の符号化・復号化ツールを開発しています。

MPEG-4 ALSの性能を示すために、MPEG-4 ALS RM18として公開されている参照ソフトウェア⁽¹²⁾ およびNTT独自高速化版の“MPEG-4 ALS fast”の2種類のMPEG-4 ALSの実装を、FLAC (Version 1.1.2)⁽¹³⁾、Monkey's Audio (MAC 4.01)⁽¹⁴⁾、OptimFrog (OptimFROG v4.520b1)⁽¹⁵⁾と比較した試験を行いました。

試験は、2.39 GHzのAMD Opteron processor 250、RAM 2 GBを用いて行い、試験用データとしてはPanasonic Corpより提供されたMPEG-4ロスレス符号化試験用データセットを用いました。

サンプリング周波数48 kHz、量子化ビット数16 bitのデータおよびサンプリング周波数48 kHz、96 kHz、192 kHz、量子化ビット数24 bitで録音された各30秒、計51個のステレオフィ

表 MPEG-4 ALSが対応する入力フォーマット

サンプリング周波数	192 kHzで試験済 (384 kHz以上の高サンプリング周波数にも対応可能)
量子化ビット数	PCM (最大32 bit)、IEEE754 32 bit浮動小数点信号
チャンネル数	最大65 536チャンネル
ファイルフォーマット	Raw, WAVE, AIFF, BWF, Sony Wave64, BWF with RF64

ルを用いて試験を行いました。トータルのデータサイズは682 562 430バイトです。

試験の結果を図2および図3に示します。横軸は30秒のデータを符号化・復号化するのにかった時間を表し、左側の小さな値ほど処理が高速であることを表します。縦軸は圧縮後のサイズが元のサイズの何パーセントになったかを次式で計算した値で、小さいほど圧縮後のサイズが小さくなり性能が良

いことを表します。

$$P = \frac{\text{CompressedFileSize}}{\text{OriginalFileSize}} \cdot 100(\%)$$

試験結果より、MPEG-4 ALS RM18およびMPEG-4 ALS fastは圧縮率と処理速度のバランスの観点から、他の符号化方式よりも優れていることが分かります。例えば、符号化・復号化処理に必要な時間が半分であるにもかかわらず、MPEG-4 ALSの圧縮率

はFLACに対して7～8%良好であり、MPEG-4 ALS fastのもっとも高速なモードでは、リアルタイムの符号化、復号化処理に必要な演算量はOpteron Processor 250 CPUの1～2%程度です。

ALSを応用したアーカイブツール

■アプリケーション

- (1) 音楽録音データのバックアップおよび伝送

The Recording Academy P&E WingのNashvilleメンバで構成される Delivery Specifications Committeeはマスタ録音の配信のための指針 (the Delivery Recommendations for Master Recordings document)⁽¹⁶⁾ を作成しました。この指針内で、録音されたデータはBroadcast Wave Flie (BWF) フォーマットを用いて保存されることが推奨されています。BWFファイルに加えて、トラッキングシートや設定メモ、マイク配置情報、レコーディング情報、楽譜や歌詞、オーケストラ配置やパートミキシング情報、その他関連するあらゆる情報がデジタル化されてパッケージ内に保持されています。

このようなファイル群は、関連する情報をまとめてフォルダの中に格納されるのが自然です。複数のファイルをまとめてフォルダごとアーカイブするツールはこのようなパッケージを扱うのに適しています。1つのアーカイブファイルとしてまとめて扱うことによって、コピーや転送の失敗を防ぐことができます。FTPでネットワーク転送する際にも複数のファイルを個々に転送するよりも、まとめて1つのファイルとして扱うほうが安全です。概念図を図4および図5に示します。

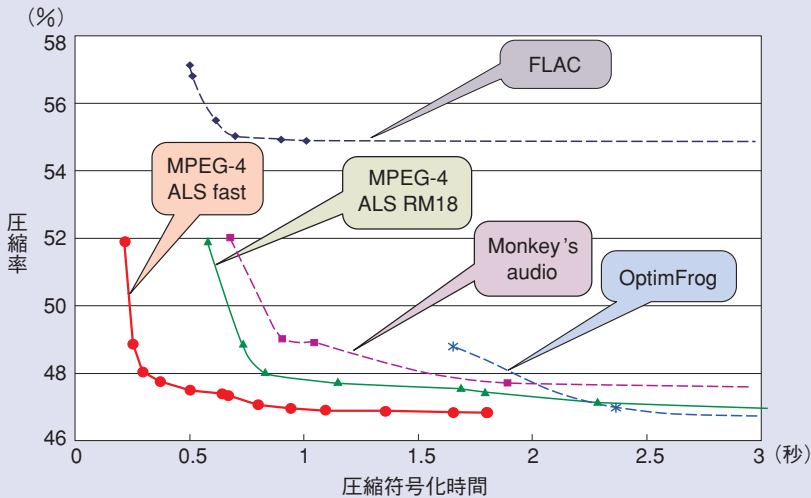


図2 圧縮符号化処理性能

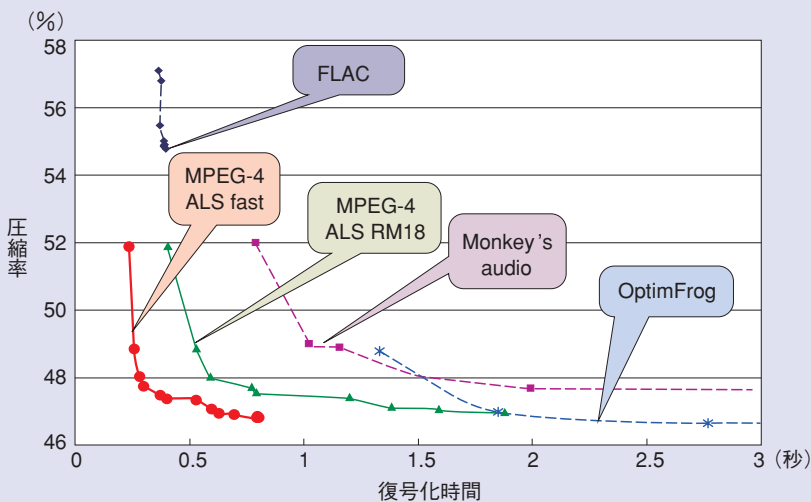


図3 復号化処理性能

(2) サウンド編集ツールの中間データの圧縮およびアーカイブ保存

DigiDesign社のProToolsやSteinberg社のNuendo等を含め、多くのプロ用途の波形編集ツールが生成された編集ファイルをフォルダの中に格納して保存する形式となっています。録音された波形データは、オーディオ・トラックごとにwavやaiffなどの決められたファイルフォーマットでフォルダ内に保存されます。フォルダ内には複数の波形データファイルやその他のメタ情報ファイルが含まれることが多く、1つの曲を構成するのに必要な複数のファイルをフォルダ内に保存したものは「プロジェクトファイル」と呼

べれます。

編集作業中のプロジェクトファイルのセットをまとめて1つのアーカイブファイルに圧縮するツールは非常に便利です。例えば、ある時点の作業ファイルをまとめてバックアップしておくことで、紛失に備えると同時に、日々の作業のバージョン管理を容易に行うことができます。フォルダ内に保存されるファイルには、音響信号以外のファイルも含まれ、すでにMP3やAACのような非可逆の圧縮符号化方式で符号化されているものもあります。このような場合には、ファイルは単にアーカイブファイルの中に加えられます。

(3) 保存・アーカイブ

必要なストレージの容量を節約することができるため、ファイルの可逆圧縮は一般的に用いられるようになってきました。コンテンツに関連するすべてのファイルは1つのアーカイブにまとめて保存されます。ファイル間には密接な関連がある場合もあるし、そうでない場合もあり、このようなファイル群を保持するのにも、フォルダ階層構造をまるごと圧縮するアーカイブツールを用いるのが良いわけです。

(4) 音響信号以外の信号に対するパッケージファイルフォーマット

音響信号に加え、MPEG-4 ALSはEFGやMEG等の音響信号以外の多チャンネル信号も効率よく圧縮することができます^{(6)・(8)}。例えば、512チャンネルのMEGデータを下のサイズの15~40%のサイズに圧縮することができます。提案するアーカイブツールを用いて、上記のような医療波形信号と、関連するメタ情報を一緒に1つのアーカイブファイルにまとめることができます。

■アーカイブツール概要

提案するオーディオ・アーカイブツールの概要を図6に示します。入力されたファイルがPCM音響信号を含むと

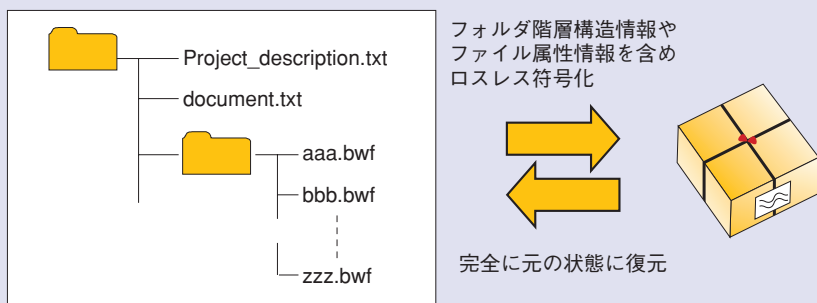


図4 フォルダとアーカイブ・パッケージの例

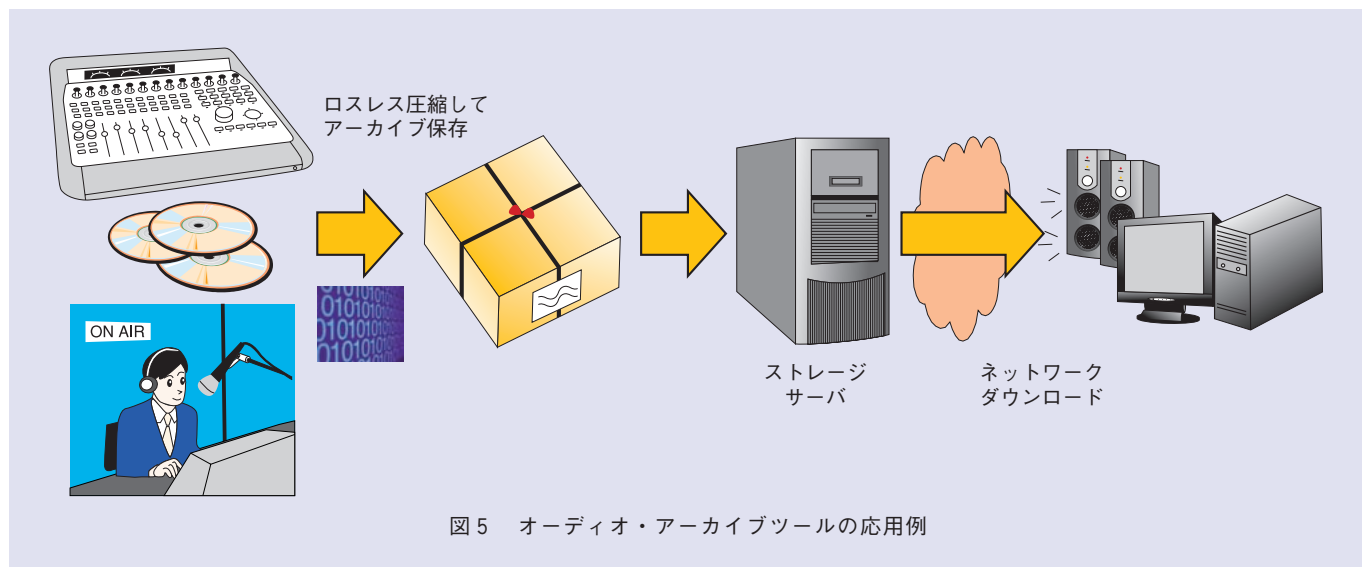


図5 オーディオ・アーカイブツールの応用例

きには、そのファイルはMPEG-4 ALSで可逆圧縮されます。例えば、wavファイルやaiffファイルはMPEG-4 ALSで符号化され、圧縮後のビットストリームが、読み込み、書き込みの可否やファイル名などの属性情報とともにアーカイブファイルに加えられます。もしもファイルがAACのような非可逆の符号化方式ですでに圧縮済みの音響データである場合や、音響データファイル以外のファイルである場合には、そのファイルはこれ以上圧縮処理を施されることなく属性情報とともにアーカイブファイルに加えられます。

■実験結果

提案するアーカイブツールの圧縮性能をPentium M (1.6 GHz) CPUのPCを用いて評価しました。試験用のデータは、波形編集用ツールを用いて録音されたプロジェクトファイルで、1つのフォルダの中に、82個のオーディオファイルと、40個の画像ファイル、2個のプロジェクト情報ファイル(バイナリファイル)を含む計242 805 050バイトを用いました。このデータを提案法とWinZIPを用いて圧縮した際の符号化処理時間、復号化処理時間と圧縮率を図7に示します。

縦軸は圧縮後のサイズが元のサイズの何パーセントになったかを前述の式で計算した値で、小さいほど圧縮後のサイズが小さくなり性能が良いことを表します。横軸は1MBのデータを圧縮符号化・復号化処理するのにかかった平均時間であり、値が小さな処理が高速であることを表しています。図7に示すように、提案するアーカイブツールはオーディオデータに関してWinZIPよりも圧縮率の点で優れているにもかかわらず処理速度も幾分速いのです。

MPEG-A プロ用アーカイブフォーマットの標準化

■MPEG-A：マルチメディア・アプリケーション・フォーマット

MPEGで新たに標準化が始まったマルチメディア・アプリケーション・フォーマット(MAF)は、特定の用途向けではあるが一般的に広く用いられるフォーマットとして複数のMPEG標準をまとめ構成した应用フォーマットです。

例えば、伝送、管理、編集、再生などに適したかたちにメタ情報と時系列メディア情報をまとめたフォーマットがMAFとして定義されています。MAFはISO/IEC 23000シリーズとして標準化されており、MPEG-Aとも呼ばれています(Aはapplicationに由来する)。

■Professional Archival Application Format

オーディオ・トラックのデータと一緒に含まれるメタ情報ファイルや画像ファイルをフォルダ階層構造情報とともに保存するフォーマットを提案しました⁽¹⁷⁾。オーディオデータの圧縮にはMPEG-4 ALSを用いることができます。

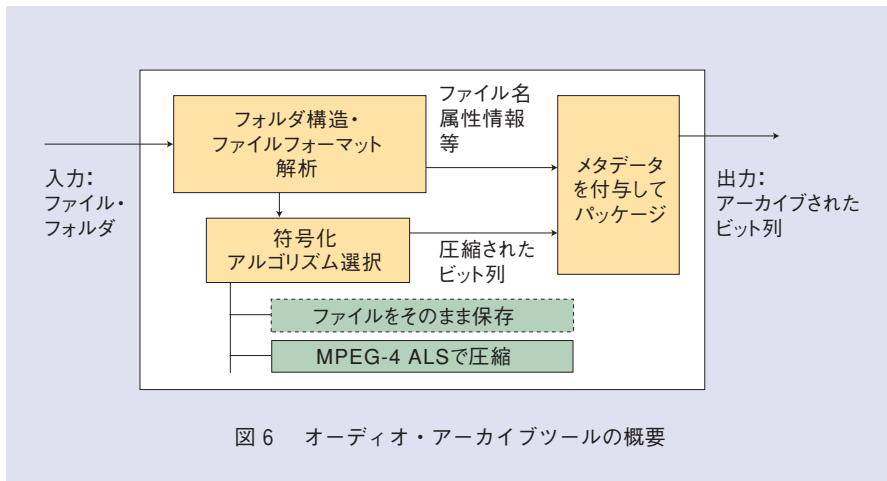


図6 オーディオ・アーカイブツールの概要

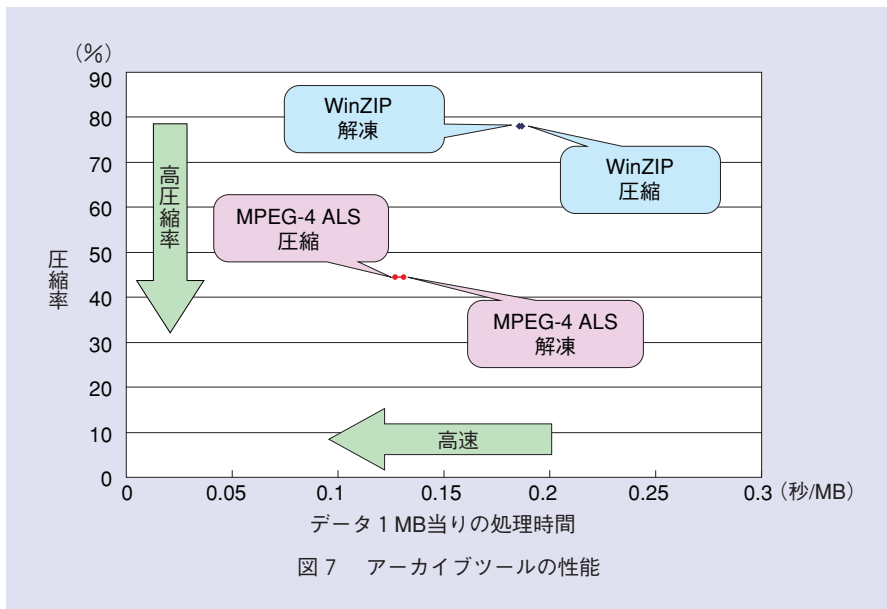


図7 アーカイブツールの性能

提案フォーマットは、MPEG-A Professional Archival Application FormatとしてMPEGで標準化中です。MPEG標準ツールを応用した本フォーマットによって、デジタルコンテンツの長期保存性、アクセスのしやすさ、再生可能性等を考慮して、デジタル時代の文化遺産およびアナログ時代のコンテンツを効率よく保存することが可能となります。

■要求仕様

当初の提案は音響データファイルの圧縮保存でしたが、適用領域をビデオ、画像を含むその他のデータにも拡大して、より一般化したツールとするため、MPEGでプロ用アーカイブフォーマットの要求条件を公募しました⁽¹⁸⁾。現時点の仕様に基づくプロ用アーカイブ

ツールの概要を図8に示します。

MPEG-4 ALSに関連するその他の標準化動向

MPEG-4 ALSの応用範囲を拡大するために、図9に示すように関連したいくつかの標準化作業が進められています。IEC TC100では、IEC 61937-10 (Digital audio -Interface for nonlinear PCM encoded audio bitstreams applying IEC 60958 Part 10 Nonlinear PCM bitstreams according to the MPEG-4 ALS format) として、MPEG-4 ALSで圧縮したビットストリームをS/PDIFとして知られるIEC 60958インターフェースで伝送するための規格を策定中です。これによって圧縮した状

態のままデコードすることなく機器間を伝送することが可能となります。

結 論

国際標準であるMPEG-4 ALSのアプリケーションとNTTによる高速な符号化・復号化ツールの実装およびその性能について紹介しました。また、MPEG-4 ALSを符号化エンジンとして応用したアーカイブツールを提案し、このツールはオーディオファイルに対してZIPよりも圧縮性能において優れていることを実験により示しました。このアーカイブフォーマットはMPEG-Aに提案され、MAFの1つとして標準化作業中です。MPEG-4 ALSと関連する標準はプロ用途、コンシューマ用途を問わず有益なツールとして広く利

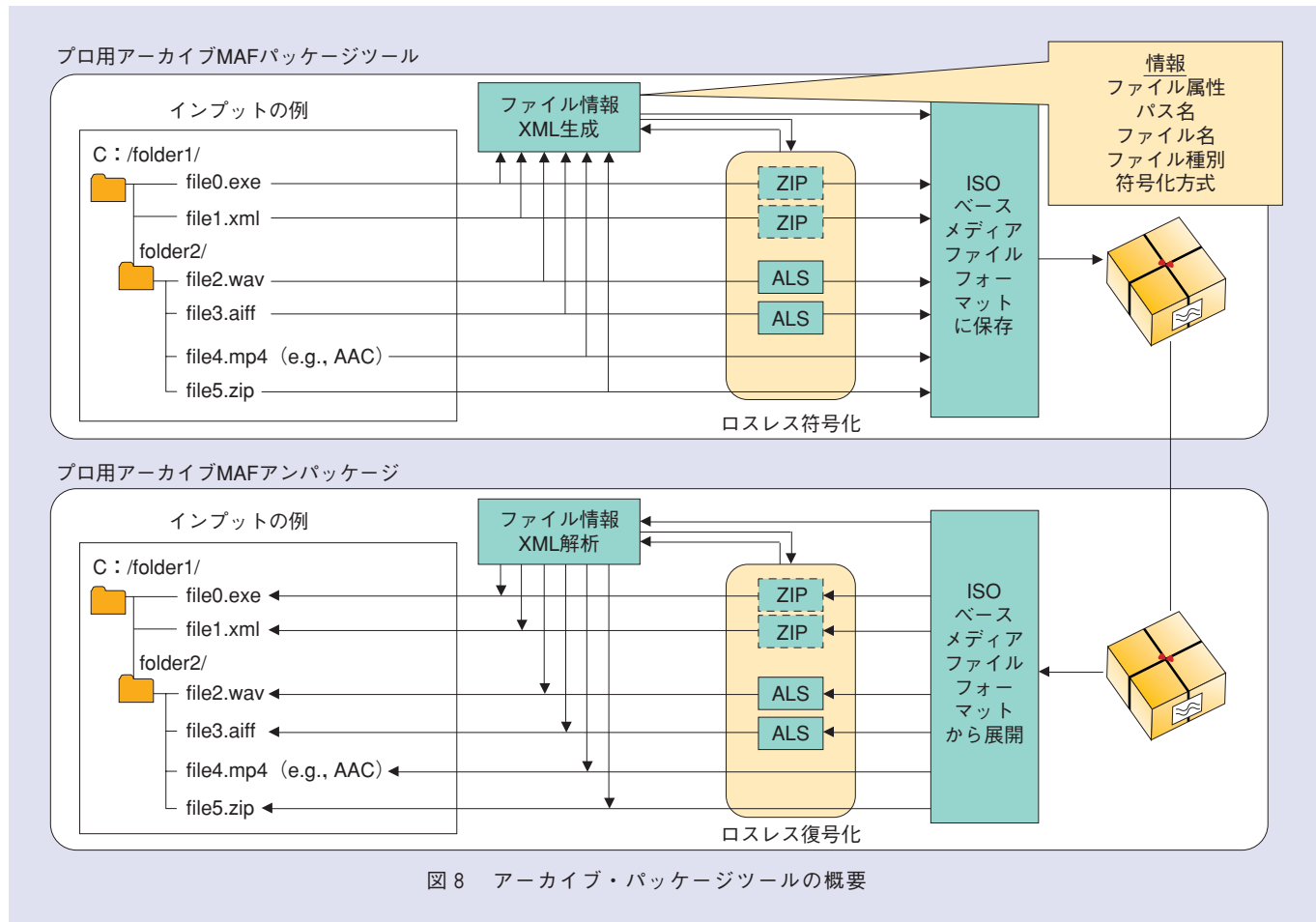
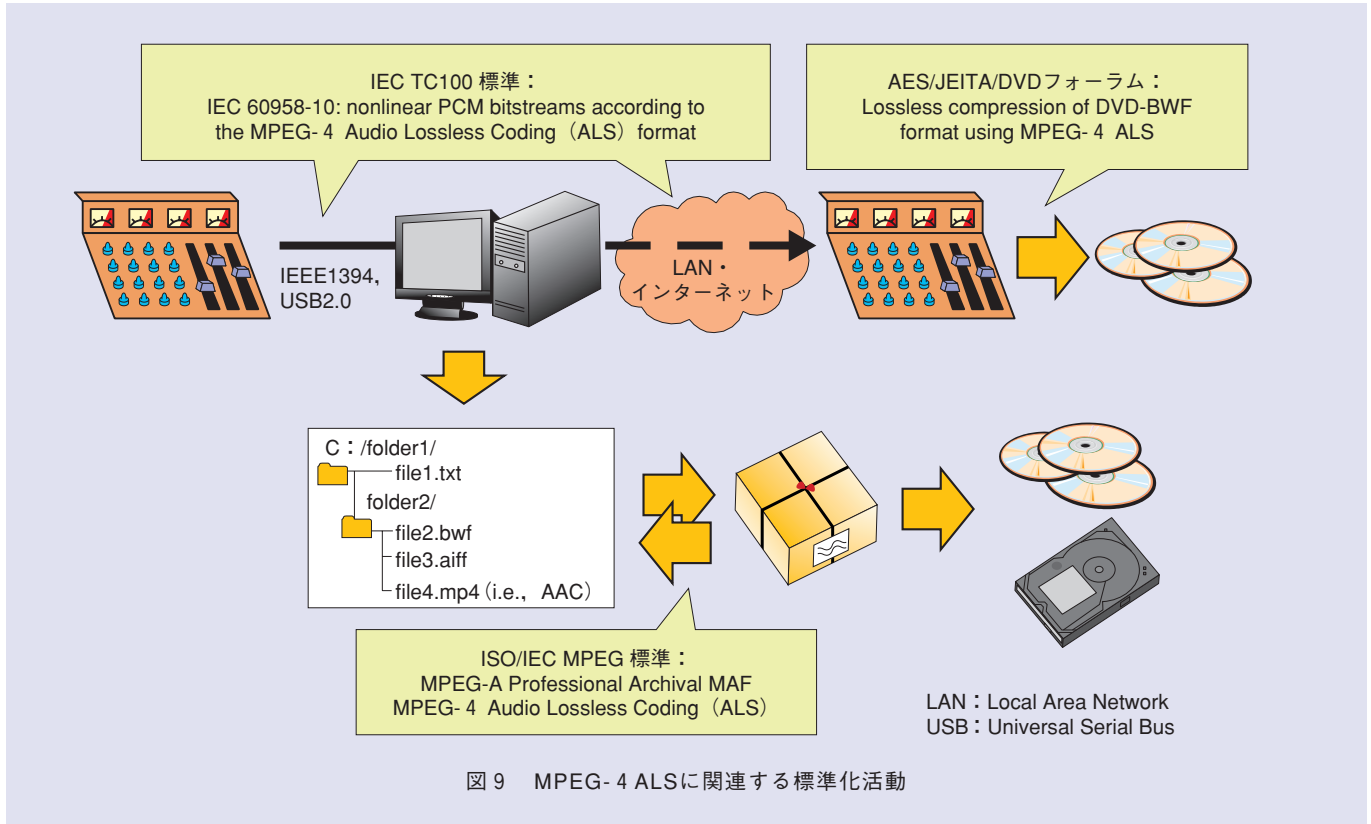


図8 アーカイブ・パッケージツールの概要



用されていくことが期待されています。

■参考文献

(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) T. Moriya, T. Liebchen, Y. A. Reznik and D. Yang, “MPEG-4 audio lossless coding,” in Preprint AES 116th Convention, No. 6047, Berlin, Germany, 2004.

(3) 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.

(4) 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, CA, USA, 2006.

(5) 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.

(6) 鎌本・守谷・西本・嵯峨山：“チャンネル間相関を用いた多チャンネル信号の可逆圧縮符号化,” 情報学論, Vol.46, No.5, pp.1118-1128, 2005.

(7) 原田・守谷・関川・白柳・鎌本：“ISO/IEC MPEG-4 Audio Lossless Coding (ALS) におけるIEEE 754浮動小数点信号の可逆符号化,”

IEICE, Vol.J89-B, No.2, pp.204-213, 2006.

(8) 鎌本・守谷・原田・西本・嵯峨山：“ISO/IEC MPEG-4 Audio Lossless Coding (ALS) におけるチャンネル内とチャンネル間の長期予測,” 信学論, Vol.J89-B, No.2, pp.214-222, 2006.

(9) 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.

(10) 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, CA, USA, 2006.

(11) ISO/IEC 14496-14: 2003, Information technology—Coding of audio-visual objects—Part 14: MP4 file format, 2003.

(12) <http://www.nue.tu-berlin.de/forschung/projekte/lossless/mp4alsRM18.zip>

(13) <http://flac.sourceforge.net>

(14) <http://www.monkeysaudio.com>

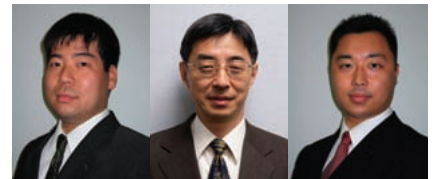
(15) <http://www.losslessaudio.org>

(16) [http://www.grammy.com/PDFs/Recording_Academy/Producers_And_Engineers/Delivery Reccs.pdf](http://www.grammy.com/PDFs/Recording_Academy/Producers_And_Engineers/Delivery_Recs.pdf)

(17) 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, CA, USA, 2006.

(18) ISO/IEC JTC1/SC29/WG11/N9165, “Call for Requirements for Professional Archival Multimedia Application Format,” 81st MPEG

Meeting, Lausanne, 2007.



(左から) 原田 登 / 守谷 健弘 / 鎌本 優

ロスレス符号化を用いることで高品質な音楽がより身近になることが期待されます。研究成果が実際に使われるためには、いろいろな課題がありますが、それらを一つひとつ解決して普及に努めていきたいと思えます。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
守谷特別研究室
TEL 046-240-3676
FAX 046-240-3145
E-mail harada.noboru@lab.ntt.co.jp