

知っておきたいキーワード

MPEG-4 ALS

原田 登†

† NTTコミュニケーション 科学基礎研究所

"MPEG-4 ALS" by Noboru Harada (NTT Communication Science Labs., Atsugi)

キーワード：MPEG-4 Audio, ロスレス符号化, 可逆符号化, 線形予測符号化

MPEG-4 ALSの概要

ネットワークやデジタル機器のブロードバンド化が進展する一方、オーディオ信号のデジタル化、高品質化（高サンプリングレート、高振幅分解能、多チャンネル化）への要求が高まり、蓄積や配信のための情報量が飛躍的に増大しつつあります。これらを背景に、原音を忠実に再現する音響信号の圧縮符号化方式の標準化がMPEGオーディオ委員会にて2002年に開始されました。この標準は、NTT、ドイツのベルリン工科大、アメリカのReal Networks、シンガポールのI2R等の提案技術を組んで、ISO/IEC 14496-3:2007 Amd.2 MPEG-4 Audio Lossless Coding（通称MPEG-4 ALS）として2006年3月に出版されました。

MPEG-4 ALSが対応する入力フォーマット表1に、符号化・復号化の基本構成を図1に示します。MPEG-4 ALSは、波形データを完全に再構成することを保証しつつ、多チャンネル信号や、高サンプリングレートの信号を歪なく

元のサイズの15~70%のサイズに圧縮することができます。この技術は、過去の複数のサンプルの値に予測係数を乗じて足し合わせることで現在のサ


ンプル値を予測する線形予測とばれる技術に基づいています。一般的な線形予測の数式を以下に示します。 

表1 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	192kHzで試験済 (384kHz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ビット), IEEE754 32ビット浮動小数点信号
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, BWF with RF64, Sony Wave64

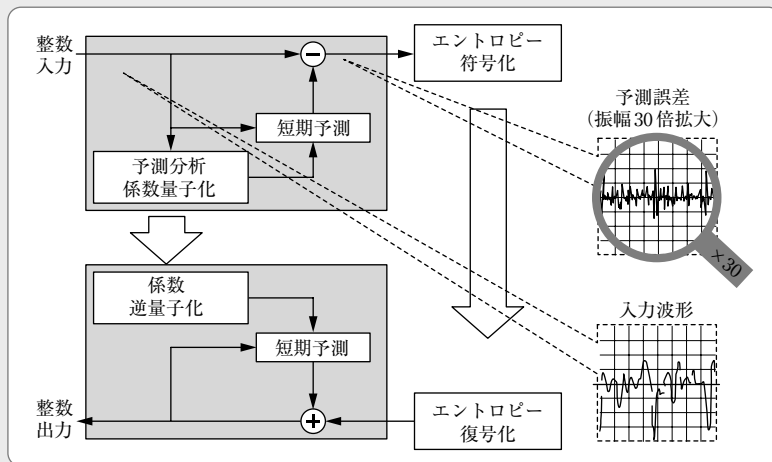


図1 MPEG-4 ALS符号化・復号化の基本構成

$$\begin{aligned}
 e(n) &= x(n) - \hat{x}(n) \\
 &= x(n) - \left(-\sum_{k=1}^P a_k \cdot x(n-k) \right) \\
 &= \sum_{k=0}^P a_k \cdot x(n-k)
 \end{aligned}$$

- $e(n)$: 予測誤差
- $x(n)$: 入力サンプル値
- $\hat{x}(n)$: 予測値
- P : 予測次数
- a_k : 予測係数 ($k=1, 2, \dots, P$)
ただし, $a_0=1$

また、元信号と線形予測誤差信号の振幅の違いを図2に示します。線形予測係数はPARCOR(偏自己相関)係数の形で量子化され、振幅の小さくなった線形予測誤差信号はRice符号化またはBlock-Gilbert Moore符号化(BGMC)を用いてエントロピー符号化

されます。

上記基本構成に加えて、ブロック分割、Multi Channel Coding (MCC)、Long Term Prediction (LTP) ツール等を用いて圧縮性能を向上しています。

また、近似共通因数符号化、マスクLZ符号化ツール等を用いて、IEEE754浮動小数点形式の信号も効率よく圧縮符号化することができます。

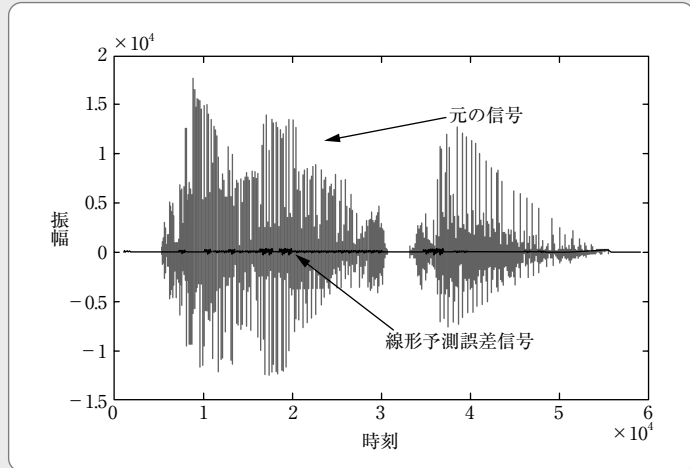


図2 元の信号と線形予測誤差信号の振幅の違い

音響信号のデータサイズ・可逆符号化と非可逆符号化の違い

図3は、3分の曲1曲に必要な情報量を表しています。例えば、CDではサンプリング周波数44.1kHz、量子化ビット数16ビット、ステレオの信号が用いられ、3分の曲1曲分のデータサイズは49MB程度になります。音楽情報の高サンプリングレート化、多チャンネル化に伴いデータサイズは大きくなります。また、プロの編集用途では、中間データのサイズが数GBになることも珍しくありません。

完全に元の信号が復元されることを保証する符号化方式は、ロスレス符号化、あるいは可逆符号化と呼ばれています。

MPEG-4 ALSのような可逆(ロスレス)符号化を用いることにより、高品質な音楽データのサイズを原音の品質を保ったまま圧縮し、取り扱いを容易にすることができます。

一方で、最近普及している携帯音楽プレーヤーでは、音楽情報のうち人間の耳に聞こえない成分を捨ててデータサ

イズを1/10~1/20程度に圧縮する技術を用いています。このような歪を許す符号化方式は、復号しても元の信号と完全に同じ波形にはならないことから、非可逆符号化、あるいは、ロッキーな符号化と呼ばれています。

可逆符号化が、元信号の完全な復元を保証するのに対して、非可逆符号化(ロッキーな符号化)では、省略しても人間の耳に差が知覚されにくい高周波成分の情報や、周波数的に他の音にマスクされる情報等を省略するた

め、聴感的にはほとんど音声品質の劣化を知覚されることなくデータサイズを小さくできますが、一方で、符号化・復号化を繰り返したり、符号化した信号を編集したりする場合に劣化が顕著に知覚される場合があります。

このため、楽曲のマスターデータの保存や、編集作業を前提としたデータ、高サンプリング周波数の信号の保存、原音を忠実に再現したいといった目的で用いるのであれば、ロスレス符号化を用いる必要があります。

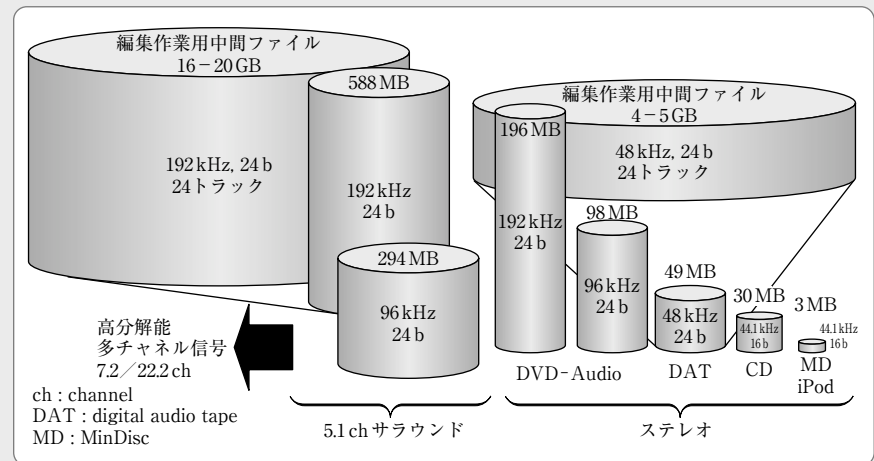


図3 3分の曲に必要な情報量

MPEG-4 ALSの圧縮性能と復号に必要な演算量

MPEG-4 ALSとその他のロスレス符号化ツールの圧縮性能を、圧縮後のデータサイズ（元のファイルサイズで正規化した圧縮後のサイズ：小さいほど望ましい）と、復号時間（短いほど使いやすい）の観点で比較したものを図4に示します。ISO/IECよりソースコードが公開されているMPEG-4 ALS RM18、および、互換性を保ったまま高速化したMPEG-4 ALS fastについて、2.39GHzのAMD Opteron プロセッサ250を用いて、30秒のファイルを復号するのにかかる平均時間と圧縮率をプロットしています（サンプリング周波数48~192kHz、量子化ビット数16~24ビット、ステレオファイル計51個の平均）。

また、サンプリング周波数や量子化ビット数、チャンネル数の代表的な組み合わせのいくつかについて、ALSビットストリームを、ARM 9Eでリアルタ

イムにデコードするために必要なCPUリソース（必要な動作クロック）の参考値を表2に示します。表より、例えば、CDと同様の44.1kHzサンプリング、16ビット、ステレオの信号であれば、ARM 9Eの動作クロック換算で、18MHz程度の演算処理速度があればリアルタイムでデコードが可能であることがわかります。

表2 MPEG-4 ALSビットストリームのリアルタイムデコードに必要なARM9E動作クロック

入力信号形式	ARM 9E動作クロック
44.1kHz, 16ビット, 2ch	18MHz
48kHz, 24ビット, 2ch	22MHz
48kHz, 24ビット, 5.1ch	65MHz
96kHz, 24ビット, 5.1ch	120MHz

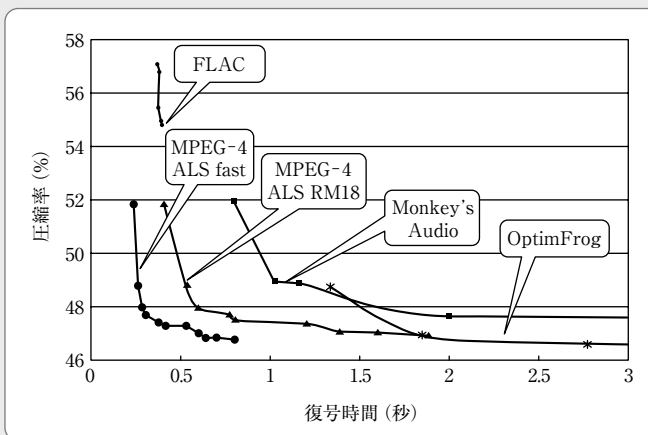


図4 復号処理性能比較

MPEG-4 ALSの応用分野

MPEG-4 ALSには、プロ用途から一般用途まで、次に示すようなさまざまなレベルの応用が考えられます。

- (1) オーディオファイルのインタネット配信（ストリーミング、オンラインミュージックストア、ダウンロード販売）
- (2) 高品質ディスクフォーマット
- (3) 映像符号化と組合せた利用（コンサート・オペラ等のライブ配信、IPTV等の放送用途）
- (4) 携帯ミュージックプレーヤ
- (5) アーカイブシステム（放送素材・

楽曲データの保存・蓄積）

(6) スタジオ編集（保存・蓄積、伝送、協調作業）

(7) 音響信号以外の時系列信号（MEG・EEG等の生体信号等）の保存

例えば、音楽ファイルのダウンロード販売等の用途では、ユーザはALSで可逆符号化された高品質な原音を購入し、自分の環境に合った非可逆符号化フォーマットに変換してから利用するという使い道が考えられます。これによって、プレーヤを買い替えたりした場合にも、それぞれのプレーヤに合った最適な品質で音楽を楽しむことができます。

また、アーカイブシステムやスタジオ編集用途では、同じくMPEG標準であるISO/IEC 63000-6 MPEG-A Professional Archival Application Formatなどのアーカイブフォーマットと組合せて用いることもできます。

さらに、MPEG-4 ALSは、元の信号を忠実に再現可能であることから、音楽信号以外にも、脳波や心電波形などの医療用生体信号の圧縮・保存にも用いることができます。

(2008年12月16日受付)

参考文献

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



原田 登 1997年、九州工業大学大学院情報工学研究科情報システム専攻修了。同年、NTT入社。現在、NTTコミュニケーション科学基礎研究所において、高効率音声音楽信号符号化の研究およびその標準化に従事。