



## MPEG-4 ALS — 歪みを許さない「ロスレス・オーディオ符号化」の国際標準

NTTコミュニケーション科学基礎研究所

もりや たけひろ はらだ のぼる かまもと ゆたか せきがわ ひろし しらやなぎ きよし  
守谷 健弘 / 原田 登 / 鎌本 優 / 関川 浩 / 白柳 潔

NTTは、高品質サービスへの利用を目的に2002年より歪みを許さないロスレス・オーディオ符号化技術の開発に取り組み、その国際標準化の策定を積極的に主導してきました。昨年12月にMPEG-4 ALSとして承認され、本年3月に出版されました<sup>(1)</sup>。その技術内容や応用用途について紹介します。

### 開発の背景

ネットワークやデジタル機器のプロ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）の要望の高まりに伴い、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。NTTコミュニケーション科学基礎研究所では、オーディオ信号を歪みなく圧縮する技術の研究に取り組んできました。これに並行して相互接続性、長期保守性、特許権利者の明確化のために、技術を国際標準化することが急務であると判断し、2002年よりMPEGオーディオ委員会が標準規格の策定を主導してきました。具体的には、NTTが中心になって要求条件の確定や技術公募を企画し、ドイツのベルリン工科大学、アメリカのリアルネットワークス社、シンガポールのI2R等と協力して、性能改善の技術提案、相互検証、参照ソフトの作成を進めました<sup>(2)~(4)</sup>。その後2回の国別投票と改定を経て、昨年12月に、「14496-4 3rd ED AMD 2（通称ALS：Audio Lossless）」として正式に国際標準として承認されました。

なお、ALSと並行して聴覚符号化から連続的にロスレス符号化につながる

SLS（Scalable Lossless）符号化も標準化されていますが、ここでは、ALSにしぼって紹介します。

これまで図1に示すように、MPEGオーディオでは世界で広く普及している重要な規格が策定されてきています。この中に含まれるMP3やデジタル放送用のAAC（Advanced Audio Coding）や標準外で普及しているミニディスク用符号等は、すべて歪みを許す聴覚圧縮符号化です。これらの符号化では情報圧縮に伴う波形の歪みを聴感上目立たな

いように制御しています。

したがって、元の音楽波形とは異なるものを聞いていることになります。今回の標準符号化は全く歪みを許さない圧縮符号化で、聴覚符号化との違いは表1のように整理できます。

### 技術の内容

#### ■基本原理

本技術は時間領域での線形予測が基本であり、符号化と復号化の基本ブロック構成を図2に示します。音声信号

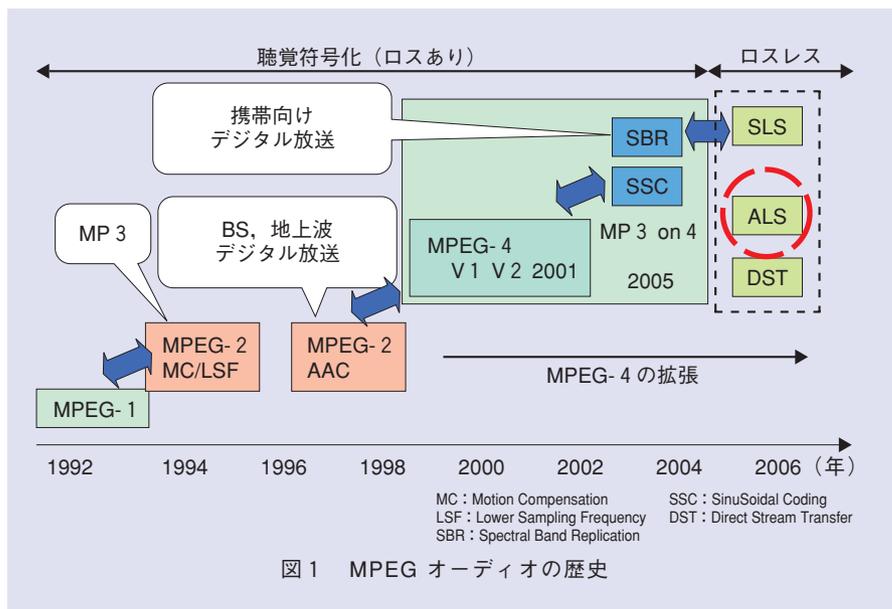
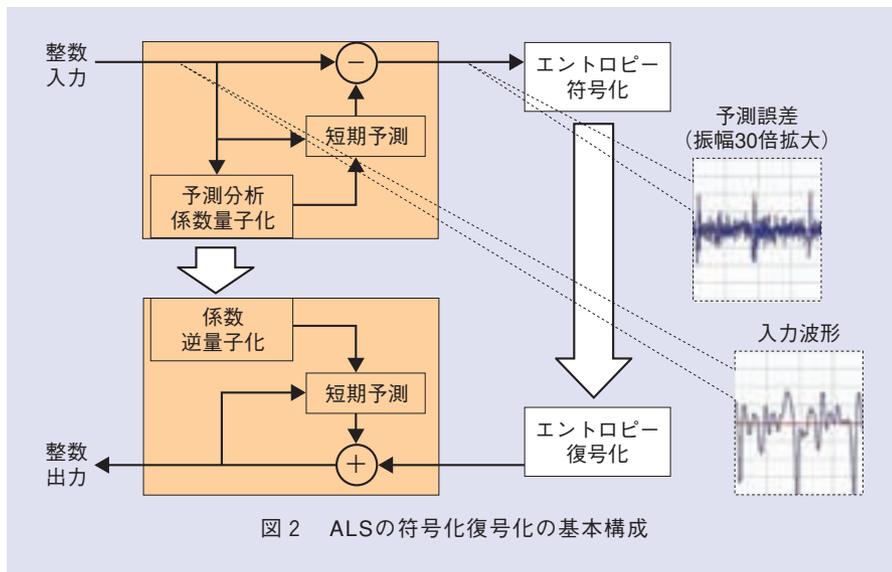


表1 聴覚符号化とロスレス符号化

	聴覚圧縮符号 (歪みあり)	ロスレス符号化
圧縮率	5-10%	15-70%
歪み	聴感的無歪	完全無歪
用途	放送, 携帯プレーヤ, 試聴	保存, 編集, 鑑賞, 音楽以外
方式	MP3, AAC, ミニディスク	フリーソフト各種, SACD, DVD-audio



に対する線形予測分析はNTTが世界のパイオニアであり、携帯電話などの音声圧縮符号化の必須の基本技術として世界に広く使われています。予測分析とは、予測回数に対応する個数の過去のサンプルから予測誤差が最小になる予測係数を計算することです。整数値として得られた予測誤差を復号器に渡し、復号器は予測誤差と予測係数によって符号器と逆の演算を行うことによって、完全に元の信号を再構成することができます。

一方、予測誤差として振幅の小さくなった各サンプルをエン트로ピー符号化\*1することによって符号量を削減することができます。表2のようなライス符号\*2として知られているエン트로ピー符号が使われており、符号量が振幅に比例

\*1 エン트로ピー符号化：元の符号の頻度に合わせて、一意に復号できる異なる長さ（頻度が高いものに短い符号）の符号を与えることで平均符号長を短縮する符号化法。

\*2 ライス符号：振幅が指数分布のときに最適となる可変長符号で、結果として表2のように振幅と符号長が比例します。

することが分かります。また予測係数はPARCOR（Partial Autocorrelation：偏自己相関）係数に変換して量子化し、さらにこれもライス符号を使って圧縮します。PARCOR係数は処理が簡単で、予測合成フィルタの安定性が簡単に判別できる利点があります。予測回数は適応的に0から1023まで選択でき、過去のフレームの波形が使えないランダムアクセスの開始フレームでは予測回数をサンプルごとに漸増させる予測を行っています。

#### ■長期予測とマルチチャンネル予測

予測符号化の基本原理は、隣接するサンプル間の相関を使ったものですが、音声や楽音のピッチ周期に起因する相関や、マルチチャンネル信号のチャンネル間相関を利用して、さらに符号量を削減することができます。このような拡張処理を図3に示します<sup>(5), (6)</sup>。長期予測、チャンネル間予測の順で短期予測の誤差に対して適用し、それぞれ5タップ、3または6タップの重み係数を使った予測により、エン트로ピー符号化への入力信号の

表2 ライス符号の構成例

振幅値	バイナリ符号
4	111111110
3	11111110
2	111110
1	110
0	0
-1	10
-2	110
-3	111110
-4	11111110

振幅を小さくします。長期予測はフレーム内の系列の過去の信号の最適な遅延量を探し、重み係数を求めます。またチャンネル間予測は参照の元とするチャンネルと重みつき差分を適用するチャンネルを指定し、チャンネル間の相対的な遅延量を探したうえで、重み係数を求めます。これらの重み係数はすべてライス符号化され、その他のパラメータとともに補助情報として符号化されます。

#### ■浮動小数点对応

図4は基本構成を利用して、浮動小数点形式の入力を圧縮する構成を示します<sup>(7)-(9)</sup>。浮動小数点形式は振幅のオーバーフローやアンダーフローが生じないので、音楽編集時の作業ファイルとして使われます。しかし浮動小数点形式の表面的数値の連続性がないので、そのままではほとんど圧縮できません。そこで予測符号化のできる整数値に変換して、その誤差は別途符号化するという構成を考案し、圧縮を実現しました。またフォーマットは浮動小数点であっても、数値は整数または整数の定数倍である場合があります。これは入力機器のAD変換直後の波形はすべて整数値であり、編集時に定数倍される場合が頻繁にあるためです。このような場合に対して、その定数を推定して割り算することで整数の系列に変換する手法を考案し、圧縮効率を大幅に改善しました。さらに入力浮動小数点数値と整数の誤差は、誤差の性質に合わせて新たに考案したマスクLZ（Lempel-Ziv）符号化で効率よく

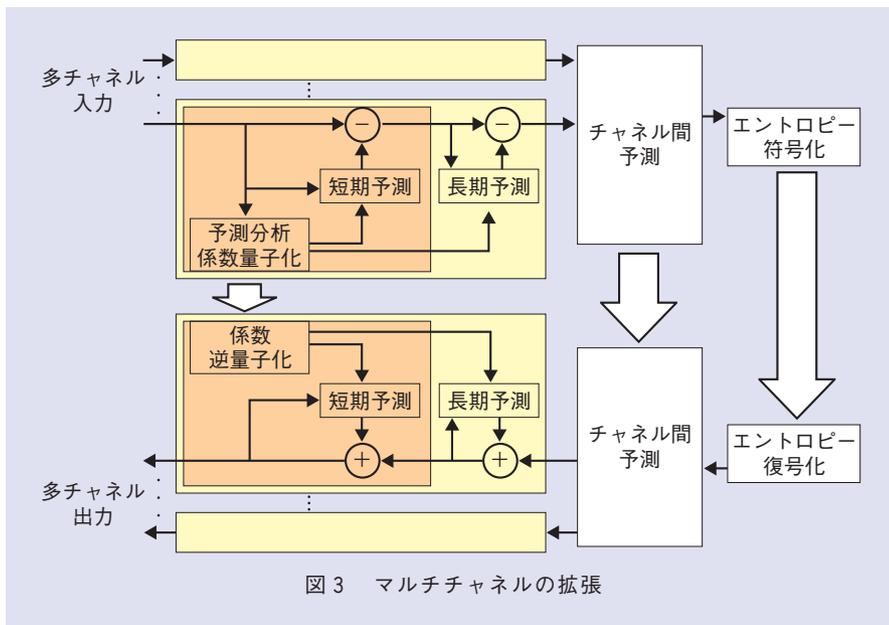


図3 マルチチャンネルの拡張

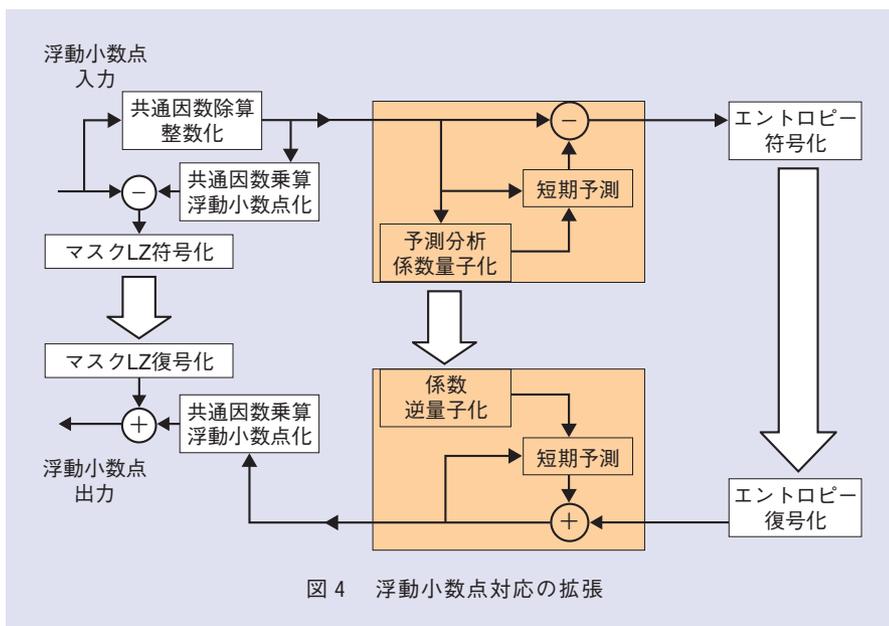


図4 浮動小数点对応の拡張

保ったまま高速化したNTT独自の処理ソフトの性能も合わせて示しました。この図から本標準化技術は、世界一の圧縮性能のレベルにあることが分かります。

本標準は、表3のようにCDの仕様よりはるかに高品質で柔軟なフォーマットまで対応可能で、ほとんどのアプリケーションに柔軟に対応できるという特徴を持っています。

また通常のPCのソフトで符号化する場合、オーディオ信号の再生時間の実時間以上の速度で実行でき、復号は実時間の10倍速以上で実行できます。このため、図6のような典型的な音楽データの配信システムにおいて、光ファイバ回線でのダウンロードであっても、ダウンロードと並行してそれより短時間で復号できるので、伝送・解凍のプロセスの時間は圧縮しない伝送に必要な時間より短縮できることとなります。すなわち、本技術による圧縮解凍の手続きによって増加する手間は非常に小さく、一方、圧縮による伝送速度の改善、蓄積コストの削減に大きな効果があります。

前述のように、本技術は処理が高速で、歪みなく圧縮できることを生かして多様な用途に応用できます。業務用としては蓄積配信システム、オーディオ編集ソフトとして利用価値が高いと考えます。また一般個人用にも携帯プレーヤ、編集ソフトに組み込むことが可能です。いずれ、テキスト圧縮ツールのZIPのようにオーディオ圧縮ツールとして幅広くPCや家電製品に組み込まれるかも知れません。さらに音楽以外の医療データや環境データの蓄積用途への応用も有望です。

## 今後の展開

MPEG-4 ALSは、世界共通の互換性のあるツールとして各種用途に普及し、たとえ10年後でも正確にかつ簡単に入力信号が再現できるように維持管理が必要です。このため、NTT研究所として引き続き標準化委員会で、参照ソフトのソースコード、相互接続性試験手順などを整備する予定です。なお、本国際

圧縮します。

### ■その他の特徴

- ① 予測パラメータを符号化せず、符号器と復号器で同一の適応予測フィルタを動作させる後方予測モードも選択可能で、演算量の増加と引き換えに高い圧縮率が達成できます。
- ② 階層的フレーム分割や圧縮率の高いエン트로ピー符号化選択等、符号器での演算量と圧縮性能の柔軟な選択が可能です。
- ③ 符号器におけるパラメータ設定に

は自由度が大きく、将来にわたって改良が可能です。

- ④ 波形全体の誤り検出符号を併用し、入力波形が正しく再構成されたことを保証することができます。

## 実用的性能評価

現在、入手可能な競合技術と標準化技術の圧縮性能を圧縮後の容量（元の信号の容量で正規化した圧縮後の容量：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5に示します。また互換性を

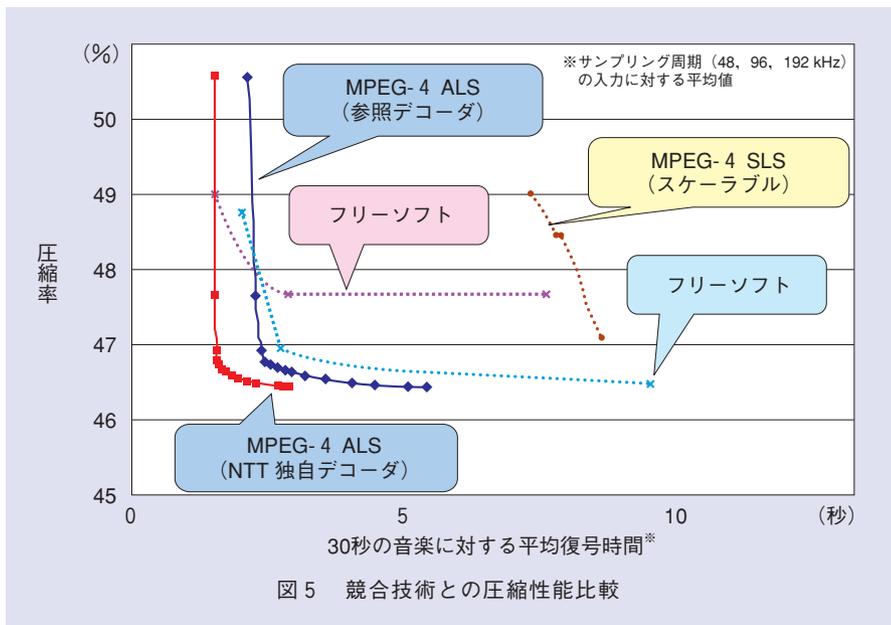
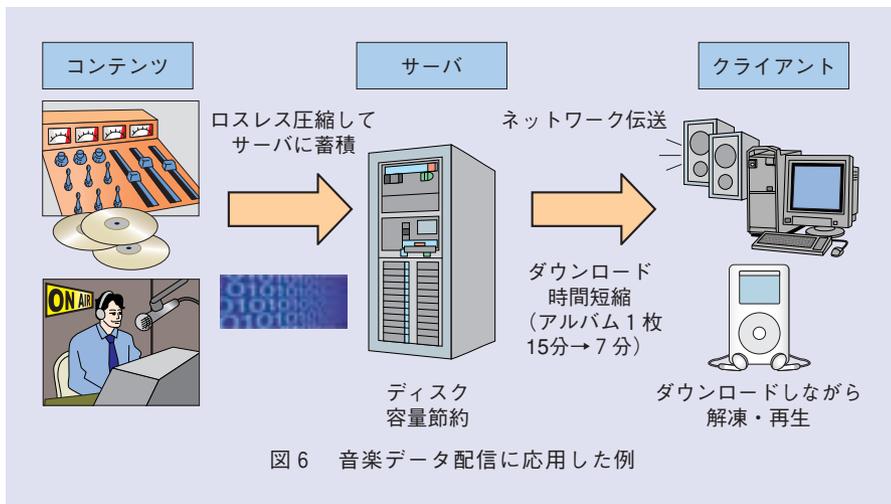


表3 ALS標準が対応する波形の形式

入力波形の形式	ALS標準	CD (参考)
最大サンプリング周期	192 kHz	44.1 kHz
最大振幅語長	32 bit	16 bit
最大チャンネル数	65 536	2
数値形式	整数・浮動小数点	整数



標準の実施に必須の特許は、今後権利者による任意団体から特許プールのかたちで実施許諾される見込みです。

またNTTコミュニケーションズでは、標準準拠ソフトの商用システムへの組み込みや販売を予定しています。さらにNTTグループとして、多様な用途での社外との連携やライセンス提供も予定しています。

■参考文献

- (1) ISO/IEC 14496-3:2005/Amd.2:2006, Information technology—Coding of audio-visual objects—Part 3: Audio, AMENDMENT 2: Audio Lossless Coding (ALS), new audio profiles and BSAC extensions, edition 2006-03-15.
- (2) T. Moriya, D. Yang, and T. Liebchen: "Extended Linear Prediction Tools for Lossless Audio Coding," Proceedings of ICASSP 2004, pp. III-1008-1011, 2004.
- (3) T. Liebchen, Y. Reznik, T. Moriya, and D. Yang: "MPEG-4 Audio Lossless Coding,"

- 116th Audio Engineering Society Convention, Preprint paper #6047, 2004.
- (4) T. Liebchen, T. Moriya, N. Harada, Y. Kamamoto, and Y. Reznik: "The MPEG-4 Audio Lossless Coding (ALS) Standard Technology and Applications," 119th Audio Engineering Society Convention, Preprint paper #6589, 2005.
- (5) 鎌本・守谷・西本・嵯峨山: "チャンネル相関を用いた多チャンネル信号の可逆圧縮符号化," 情処学論, Vol.46, No.5, pp.1118-1128, 2005.
- (6) 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.
- (7) D. Yang and T. Moriya: "Lossless Compression for Audio Sources with IEEE Floating Point Format," 115th Audio Engineering Society Convention, Preprint paper #5987, 2003.
- (8) N. Harada, T. Moriya, H. Sekigawa, and K. Shirayanagi: "Lossless Compression of IEEE Floating-point Audio using the Approximate-Common-Factor Coding," 118th Audio Engineering Society Convention, Preprint paper #6352, 2005.
- (9) 原田・守谷・関川・白柳・鎌本: "ISO/IEC MPEG-4 Audio Lossless Coding (ALS) におけるIEEE754浮動小数点信号の可逆符号化," 信学論, Vol. J89-B, No.2, pp.204-213, 2006.



(後列左から) 原田 登 / 白柳 潔 / 関川 浩  
(前列左から) 守谷 健弘 / 鎌本 優

NTTの事業はもちろん、それだけでなく世界中の蓄積サーバ、個人のPC、携帯電話などにツールとしてさりげなく組み込まれ、便利にしかも次の世紀の人達にも使ってもらえることが夢です。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所  
人間情報研究部  
TEL 046-240-3676  
FAX 046-240-4709  
E-mail n-harada@theory.brl.ntt.co.jp