



2005年12月27日

NTTが主導で開発した技術がMPEGの国際標準規格として承認 ～歪みを許さないオーディオ符号化(MPEG-4 ALS)～

日本電信電話株式会社(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とりくみ、その国際標準化の策定を積極的に主導してきました。このたび最終国際投票の結果が公表され、NTTが他研究機関と協力して策定した技術がオーディオ・映像信号の圧縮方式の1つであるMPEG^{*1}の国際標準規格として承認されました。

この規格の中にはNTT研究所で30年以上前に発明されたPARCOR係数^{*2}をはじめ、NTTが提案した多くの要素技術が盛り込まれ、またNTTと東京大学大学院情報理工学研究科の嵯峨山茂樹教授の研究室(東京都文京区)との連携講座での共同研究成果も含まれております。

1. 開発の背景

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の普及、高品質化(高サンプリングレート、高振幅分解能、多チャネル化)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符号化する技術を開発するとともに、相互接続性、長期保守性、特許権利者の明確化のために、その仕様を国際標準化することが急務であると判断し、2002年よりISO/IEC^{*3}のMPEGオーディオ委員会で標準規格策定を主導してきました。

2. 国際標準規格承認までの取り組み等

今回の標準化活動において、NTTは要求条件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などと協力して、性能改善のための技術提案、相互検証、参照ソフトの作成を積極的に進めました。

その後、23カ国による2回の国別投票と改定を経て、このたびISO/IECでの最終の国別投票が締め切られ、「14496-3 3rd ED AMD 2 (通称

A L S : Audio Lossless) [*4](#)」として正式に国際標準として承認されました。

今回の規格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100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されることが期待されます。このため、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が整備されます。なお、本国際標準の実施に必須の特許は、今後権利者による任意団体による特許プールの形で実施許諾される見込みです。

3. 技術の内容

【ポイント】

- ◇ 符号化によりファイルの圧縮後も元のままに復号化（解凍）することが可能
- ◇ 圧縮性能は世界 1 のレベル
- ◇ 解凍に要する手間は小さく、伝送や蓄積コストの大幅な削減が可能

これまでのオーディオ符号化の主流はM P 3やデジタル放送用のA A C [*5](#)、ミニディスク用符号などで、すべて歪みを許す符号化でした。これらの符号化では、人間の聴覚の特性を利用して、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原信号の1/10程度まで圧縮しています。従って、波形情報としてはもとの音楽とは相当異なるものを聞いていることになります。

一方、今回のロスレス符号化では符号化により情報を圧縮しても復号化(解凍)プロセスにより、もとのデジタル波形を1ビットの違いもなく再構成することが保障できます。このため、波形の編集や長期保存も含む高品質オーディオ信号を扱う用途には不可欠です。その反面、圧縮率には限界があって、入力波形の性質に依存して異なりますが、もとのファイル容量の15－70%程度になります。ただしこの圧縮性能はテキスト用のZ I P [*6](#)より大幅に高く、オーディオ信号専用の入手可能なフリーソフトと比較しても世界一のレベルにあります。

[図1](#)は、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)の観点で比較して示したものです。今回の標準化技術は目的によって圧縮率と処理量を選択でき、通常モードでは高速に処理ができ、高圧縮モードでは高い圧縮率を実現できます。

また互換性を保ったままN T T独自の高速処理ソフトも実現されています。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わかります。

また入力の対象となるオーディオ信号は

- ・ サンプル周波数192 kHzまで（コンパクトディスク(CD)は44.1

kHz)

- ・ 32ビットまでの各種整数PCMオーディオフォーマット対応（CDは16ビット）
- ・ IEEE754 32ビット浮動小数点对応（CDは整数のみ）
- ・ チャンネル数65536まで（CDは2チャンネル）

であり、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、復号では実時間の10倍以上の速度で実行できます。このため、光ファイバ回線でのダウンロードであっても、ダウンロードと平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。

本標準化技術はNTTが30年以上前から取り組んできた音声分析や予測技術を土台に、予測誤差を小さくし、符号量を最小化する数多くの技術が盛り込まれています。

特にNTTから提案された主要要素技術として

- ・ PARCOR係数を使う時間領域の線形予測
- ・ マルチチャンネル符号化(NTTと東大の連携講座の成果)
- ・ 長期予測符号化(NTTと東大の連携講座の成果)
- ・ 浮動小数点用共通因数分析とマスク付適応圧縮
- ・ ランダムアクセス対応漸増次数線形予測

などが組み込まれています。またNTT研究所は、標準化と平行して、標準との互換性を厳密に維持したまま、ソフトの独自の高速化や、符号器の改良による性能向上を図ってきました。

4. 今後の展開

NTTコミュニケーション科学基礎研究所では引き続き、互換性確認試験手順策定などの標準化のサポートや、標準準拠の符号器の性能向上を図ります。

また併せてNTTコミュニケーションズは、標準準拠の実用的ソフトを、業務用の音楽の蓄積、配信の効率化を図るツールとして組み込んだ商用システムの販売を予定しています。

さらにNTTグループとして、業務用オーディオ編集ソフト、一般個人用携帯機器、編集ソフト、音楽以外の医療データや環境データの蓄積用途

への展開を見込み、N T Tのグループ会社内外との連携や、ライセンス提供も予定しています。

<用語解説>

*1 M P E G

Moving Picture Expert Group : ISO/IEC JTC1/SC29/WG11の標準化グループで1978年より、主にビデオ信号とオーディオ信号の符号化の標準を策定してきています。

*2 P A R C O R係数

Partial Auto Correlation (偏自己相関) : 1972年電電公社通信研究所で開発された音声信号を予測するパラメータの表現形式で、量子化特性や予測システムの安定性の保証の点で優れているため、世界的に音声合成などに広く使われました。

*3 I S O / I E C

ISOの正式名称は国際標準化機構 (International Organization for Standardization) 、IECの正式名称は国際電気標準会議 (International Electrotechnical Commission) で国際標準策定機関です。

*4 14496-3 3rd ED AMD 2 (ALS)

MPEG-4 オーディオ規格第3版追補が正式規格名であるが、M P E G A L Sと呼ばれています。

*5 A A C

Advanced Audio Coder (先進音響符号化) : 1997年に策定された高品質の音楽符号化でMP3より優れた品質をもち、日本の衛星および地上波デジタル放送や、携帯音楽プレーヤに使われています。

*6 Z I P

テキストやプログラムなどのデータをデータの特徴を学習しながら適応的に歪なく圧縮する技術のソフトで、P CのO Sにも標準的に組み込まれています。

・[図1. 競合技術との圧縮性能比較](#)

<本件に関する問い合わせ先>

N T T先端技術総合研究所

企画部 情報戦略担当

為近、甕(もたい)

Tel: 046-240-5152
<http://www.ntt.co.jp/sclab/contact>

NTT ニュースリリース 

Copyright(c) 2005 日本電信電話株式会社