

オーディオ符号化の最新標準化動向

守谷 健弘

NTT コミュニケーション科学基礎研究所

和文抄録

2002年以降ISO/IEC MPEGで標準化されたオーディオの符号化方式を紹介する。この中にはSBR (Spectral Band Replication), PS (Parametric Stereo), サラウンド符号化などのサブバンド領域での処理で低ビット化をねらう方式と、ロスレス(歪なし、可逆)符号化の3種の方式(ALS, SLS, DST)を含む。特に時間領域での簡易な圧縮方式 ALS に焦点をあてて、標準化の背景と経緯、技術、性能、用途、今後の展望について説明する。またロスレス符号化の重要な用途である長期保存、アーカイブ用途に関するフォーマット(Professional Archival Application Format) の動向や、今後の標準化予定についても紹介する。

Recent standardization activities of Audio Coding

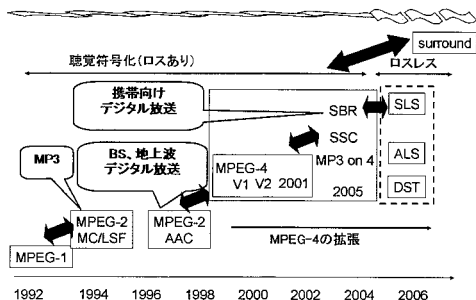
Takehiro Moriya

NTT Communication Science Laboratories

Abstract

Audio coding standards defined in the ISO/IEC MPEG since 2002 have been introduced. There are roughly two categories. One is low bit rate coding based on the processing in sub-band domain, such as SBR (Spectral Band Replication), PS (Parametric Stereo) and Surround Coding. The other is lossless coding which includes three schemes (ALS, SLS, DST). Among all, simple predictive coding, ALS (Audio Lossless coding) is focused. Various aspects including its background, technologies, performances, applications, and future activities are explained. In addition, a new standardization activity of Professional Archival Application Format and future MPEG standards are introduced.

MPEGオーディオの歴史



MPEG-4の拡張規格(1)

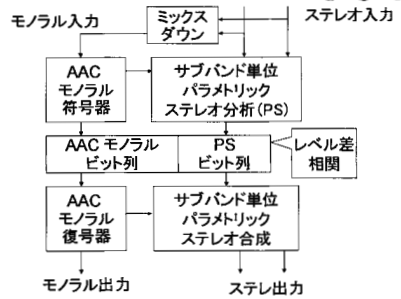
- 帯域拡張 SBR (Spectral Bandwidth Replication)
 - MP3プロの技術。少ないパラメータ(スペクトル包絡など)で高域側を表現。低域はAACなど。
 - 24 kbit/sで 32kbit/s AACより高品質
 - Coding Technology、松下、NECが貢献し、完了。
- パラメトリック楽音符号化
 - Philipsが主に貢献
 - 特にパラメトリックステレオ符号化が有効。モノラルに、わずかな追加パラメータでステレオ化が可能。

MPEG-4拡張規格(2)

- MP3 on MP4
 - (MP3+JPEG+XML)→MP4システム統合
- ロスレス符号化
 - ALS (時間領域の予測符号化)
 - SLS (AAC+周波数領域でのスケラブル符号化)
 - DST (1ビット形式DSDの圧縮、SACD圧縮と互換)
- MPEG surround
 - 5.1ch を分析再生、2ch再生
- 楽譜の書式
 - SMR (Symbolic Music Representation)

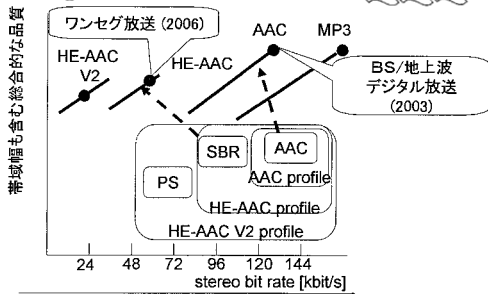
3

MPEG PS (HE-AAC v2)



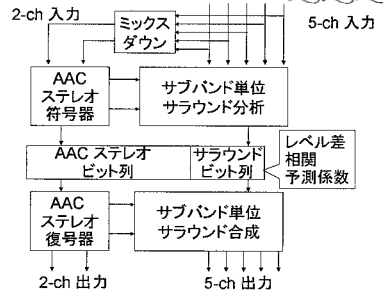
6

SBRとPSによる品質改善



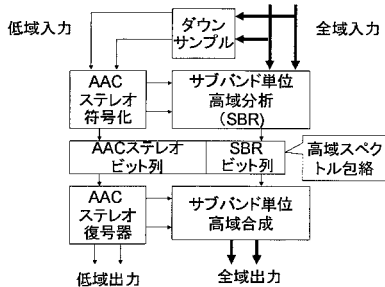
4

MPEG surround



7

MPEG SBR (HE-AAC)



5

聴覚圧縮符号化から歪のない符号化

- 歪のない圧縮
 - 高品質・高サンプルレート・マルチチャンネル
 - 過去の大量のアナログデータの永久保存
- ブロードバンドサービスには必要
- 100年後にも解凍できなくてははいけない
- 媒体の大容量化速度 < コンテンツ増大速度

国際標準が必要

歪なし	歪あり
非圧縮	圧縮

8

ロスありとロスなし

- 信号にロス(歪)が生じる圧縮符号化
 - MPEGレイヤーⅢ (MP3)、AAC、ミニディスク等
 - 原音の情報を1/5から1/10に圧縮
 - 聴感上の劣化は殆どないが、再生波形は異なる
- 信号にロス(歪)が生じない圧縮符号化
 - 元のデータを復元可能(無歪)
 - マスターデータ・編集用素材の保存にも最適

圧縮優先 品質ベストエフォート

品質優先 圧縮ベストエフォート

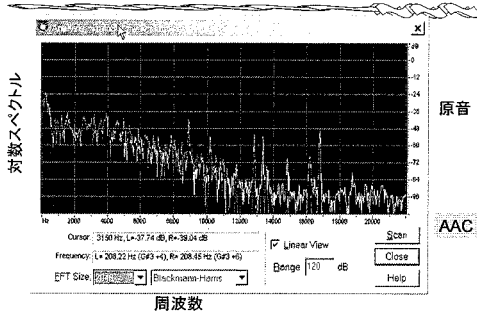
9

3種類のロスレス符号化

- ALS (Audio Lossless)
 - 時間領域
- SLS (Scalable Lossless)
 - 周波数領域
 - AACなどをコアとし、ロスレスまで細かく拡張
- DST (Direct Stream Transfer)
 - DSD (Direct Stream Digital: 1ビットオーバーサンプル形式)の圧縮
 - SA-CDの圧縮と互換

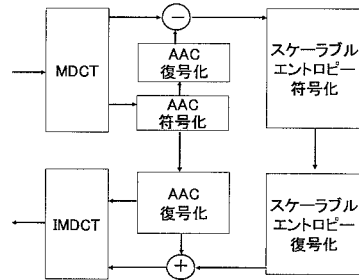
12

AACと原音のスペクトルの差



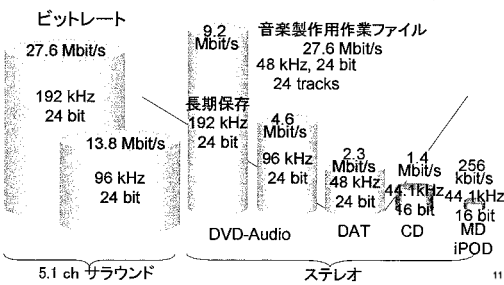
10

SLS (スケーラブル・ロスレス符号化)



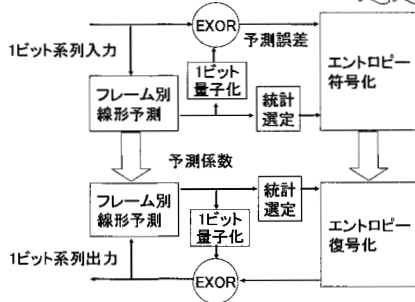
13

コンテンツの情報量の増加



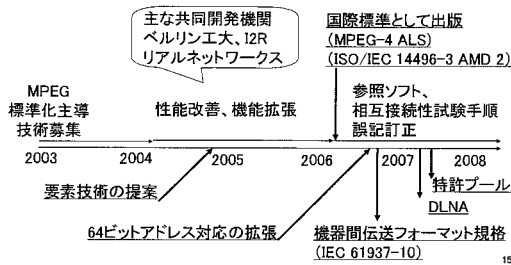
11

DST (1ビット量子化のロスレス符号化)



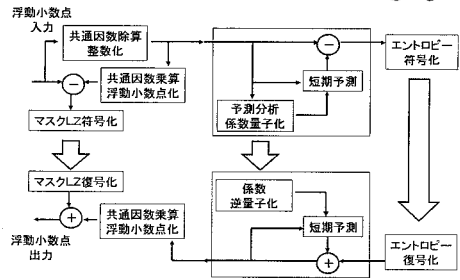
14

標準化の経緯 (MPEG-4 ALS)



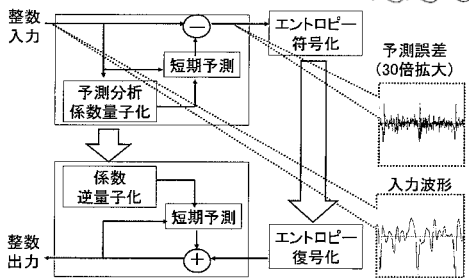
15

浮動小数点对応の拡張



18

ALSの符号化復号化の基本構成



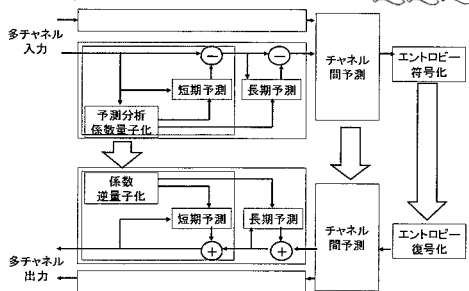
16

MPEG-4 ALSの緒元

サンプリング 周波数	44.1, 48, 96, 192 kHzに対応 (384 kHzにも対応可能)
振幅量子化 ビット数	整数PCM (32 bitに対応), IEEE754 32ビット浮動小数点
チャンネル数	最大65536チャンネル
ファイル フォーマット	Raw, WAVE, AIFF, BWF, Sony Wave64, BWF RF64
圧縮性能	入力信号を15%~70%に圧縮(入力依存)
演算量	48kHz 16bitステレオ、(PentiumM1.0 GHz) 符号化:実時間の1/2~1/100程度の時間 復号化:実時間の1/30~1/100程度の時間

19

マルチチャンネルの拡張



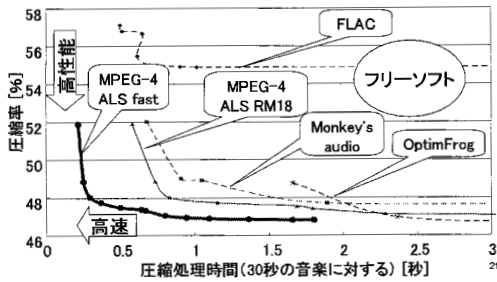
17

ALSの構成

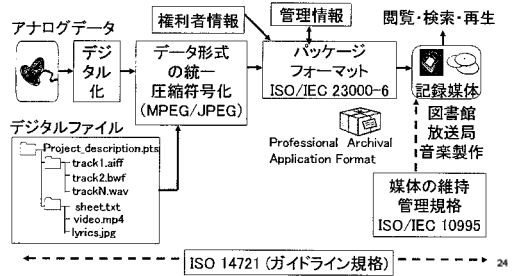
- 時間軸予測符号化
 - 前方適応予測と後方適応予測
 - 予測残差用BGMC符号
 - PARCOR 係数のGolomb-Rice符号
 - 漸増次数予測
 - 長期予測
 - 階層ブロック長選択
- 拡張
 - 浮動小数点对応
 - マルチチャンネル予測

20

ALSエンコーダの性能比較



長期保存のための技術



ALSの優位性

音響信号・時系列データを完全に復元可能な形で15~70%のサイズに圧縮する

- 機能の優位性
 - 符号化によりファイルの圧縮後も元のままの波形に完全復号(解凍)することが可能
 - 歪のない圧縮としての性能は世界1のレベル
 - 高速復号が可能で、大幅な伝送・蓄積コスト削減が可能
- 標準としての優位性
 - 他サービスと相互接続しやすい
 - 特許権利関係がクリアで安心
 - 100年後にも必ず使える

22

最新情勢

- AAC-ELD (Enhanced low delay) IS
 - 通信用途にも使える低遅延AAC (30ms)
- SAOC (Spatial Audio Object Coding) CD
 - 音源位置の情報、制御
 - surround coding tool を利用
- Speech and Audio Coding 公募中
 - 16-20 kbit/sで音声音楽に対する万能符号化

24

ALSの展開

- アーカイブ、専門家用圧縮解凍ソフト
 - 蓄積配信、ネットワークサービス
 - 音響データの蓄積編集(スタジオ・レコード会社)
- 一般用PC向け圧縮解凍ソフト
 - オンラインミュージックストア・携帯プレーヤ
 - OSのデータ圧縮機能(WinZIPなどへの組み込み)
- 家電製品、ホームネットワーク
- 音響データ以外への応用
 - 生体、医療、地震、センサーアレイ、臨場感音場

23

まとめ

- サブバンド領域での処理
 - SBR、PS、サラウンド、SAOC
- ロスレス符号化
 - ALS、SLS、DST
- 通信用符号化との境界領域
 - AAC-ELD
 - Speech and Audio
- ITU-Tでは広帯域化、スケーラブル化が進展

25

参考文献

1. ISO/IEC 14496-3:2008 Information technology - Coding of audio-visual objects - Part 3: Audio.
2. M. Hans, R. W. Schafer, "Lossless Compression of Digital Audio," IEEE Signal processing Magazine, pp 21 - 33, 2001.
3. 守谷, 原田, 鎌本, 関川, 白柳 "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 技術ジャーナル, vol. 18, No. 6, pp.42 -p.44, 2006.
4. T. Moriya, N. Harada, Y. Kamamoto and H. Sekigawa, "MPEG-4 ALS - International Standard for Lossless Audio Coding," NTT Review, Letters, Vol. 4 No. 8 pp. 40 - 45, 2006.
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", 情報処理学会論文誌, vol. 46, No. 5, pp. 1118 - 1128, 2005.
6. 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", 電子情報通信学会論文誌 VOL.J89-B No.2, pp. 204 - 213, 2006.
7. 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", 電子情報通信学会論文誌 VOL.J89-B No.2, pp. 214 - 222, 2006.
8. D. Yang and T. Moriya, "Lossless Compression for Audio Sources with IEEE Floating Point Format" Preprint 115th AES Convention, #5987, 2003.
9. T. Moriya, D. Yang and T. Liebchen, "A design of lossless compression for high quality audio signals," Proceedings of ICA (International Congress on Acoustics), 2004.
10. D. Yang, T. Moriya and T. Liebchen, "A Lossless Audio Compression Scheme with Random Access Property," Proceedings of ICASSP 2004, pp. III-1016 - 1019, 2004.
11. T. Moriya, D. Yang and T. Liebchen, "Extended Linear Prediction Tools for Lossless Audio Coding," Proceedings of ICASSP 2004, pp. III-1008 - 1011, 2004.
12. T. Liebchen, Y. Reznik, T. Moriya and D. Yang, "MPEG-4 Audio Lossless Coding," Preprint paper #6047, 116th AES Convention, Berlin, 2004.
13. N. Harada, T. Moriya, H. Sekigawa and K. Shirayanagi, "Lossless Compression of IEEE Floating-point Audio using the Approximate-Common-Factor Coding", Preprint paper #6352, 118th AES Convention, 2005.
14. T. Liebchen, T. Moriya, N. Harada, Y. Kamamoto and Y. Reznik, "The MPEG-4 Audio Lossless Coding (ALS) Standard Technology and Applications," Preprint paper #6589, 119th AES Convention, New York, 2005.
15. T. Moriya, N. Harada and Y. Kamamoto, "Performance-complexity tradeoffs of the MPEG-4 ALS lossless coding standard," WA7a-4, IEEE 40th Asilomar conference, 2006.
16. N. Harada, T. Moriya, and Y. Kamamoto, "An audio archiving format based on the MPEG-4 Audio Lossless Coding (ALS)," Preprint paper #6895, 121st AES Convention, 2006.
17. N. Harada, T. Moriya, and Y. Kamamoto, "MPEG-A Professional Archival Multimedia Application Format (MAF) Under Development," 9-2, 31st AES International Conference, 2007
18. Y. Kamamoto, N. Harada, and T. Moriya, "A multichannel linear prediction tool for MPEG-4 ALS compliant encoder," 023, AES 13th Regional Convention, Tokyo, 2007.
19. Y. Kamamoto, N. Harada and T. Moriya, "A Multichannel Linear Prediction Method for the MPEG-4 ALS Compliant Encoder," Proc. 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WP1-14, 2007.
20. 守谷 "オーディオ符号化技術とMPEG標準", 電気学会誌, vol. 127, No. 7, pp. 407 - 410, 2007.
21. 守谷 "音声音響符号化における標準", 日本音響学会学会誌, vol. 64, No. 2, pp. 114 - 118, 2008
22. 鎌本、原田、守谷、Csaba "ロスレス・オーディオ符号化 MPEG-4 ALS の高性能化", NTT 技術ジャーナル, Vol. 20, No. 2, pp. 11 - 18, 2008
23. 原田、守谷、鎌本, "MPEG-4 ALS の性能・応用と関連する標準化活動", NTT 技術ジャーナル, Vol. 20, No. 2, pp. 19 - 25, 2008