

特集:2016年「カンファレンス」

## 国際標準 MPEG-4 ALS による ハイレゾ音源ロスレス伝送

日本電信電話(株) NTTコミュニケーション科学基礎研究所  
鎌本 優, 守谷 健弘, 原田 登

はじめに

モノ(物質)からコト(事象・体験)へ消費が移りつつあると言われている。移動手段としての自動車(所有)だけではなく安全や環境を考慮した自動車(安心志向)の販売台数が伸びていたり、化学調味料(調整)よりも無添加天然出汁(健康志向)の人気が出てきたりしている。図1に示すように、音楽を楽しむ場合にもレコードやCD(Compact Disc)を所有する「モノ」からライブや握手会に参加する「コト」に変化しつつある。ハイレゾ対応のスマートフォンも増えてきており、良い音を体験できる機会が増えてきている。音を運ぶ媒体がレコードやCDのように「モノ」に依存して音質がある程度固定化されていた頃とは異なり、ネットワークを用いて伝送する場合は、歪みのあるロッキー符号化から歪みの無いロスレス符号化および全く圧縮しないPCM(Pulse Code Modulation)形式があり、ユーザは自分の聞きたい形式を選択し、音楽を体験する「コト」ができる。通信回線速度がそれほど十分でない場合でもロッキー符号化によりステレオの音響信号を64 kbps~256 kbps程度に圧縮することで、ネットワークを通して音楽を楽しむことができるようになってきている。例えば、地デジの放送でも使われているMPEG-2 AAC(Advanced Audio Coding)により、放送局で制作されたコンテンツを家庭でもほぼ原音品質で楽しめるようになった。これから始まる超高精細度テレビジョン放送(4K/8K放送)では音声のチャンネル数は22.2 chまで拡張されるだけでなく、ロスレス符号化であるMPEG-4 ALS(Audio Lossless Coding)を使うことができるようになり、放送局で制作された音源をそのまま楽しめることができるようになってきている[1, 2]。しかたなくロッキー符号化を使っていた時代から、音質に妥協しないロスレス符号化を使うことができる時代になってきている。特に、ハイレゾ音源を伝送する場合には、ロッキー符号化を使ってしまうと原音を持つせつかくの情報量の多さ(特に高域の情報)を失ってしまうため、もったいない。また、非圧縮PCMで伝送することも考えられるが、限られた電波資源や通信資源を無駄なく使うためには、デジタル信号の劣化が全くないロスレス符号化は有効である。ロスレス符号化によりおおよそ半分の大きさのビットレート・ファイルサイズに圧縮できれば、使用するパケット量も半分ですみ、ダウンロード時間も半分ですむようになる。

本稿では、音響ロスレス符号化の国際標準であるMPEG-4 ALSについてハイレゾ音源をロスレスで伝送する仕組み、特に放送への応用について述べる。

ISO: International Organization for Standardization

IEC: International Electrotechnical Commission

MPEG: Moving Picture Experts Group

ARIB: Association of Radio Industries and Businesses

JEITA: Japan Electronics and Information Technology Industries Association

## 音楽（モノ）の所有

## 音楽（コト）の体験

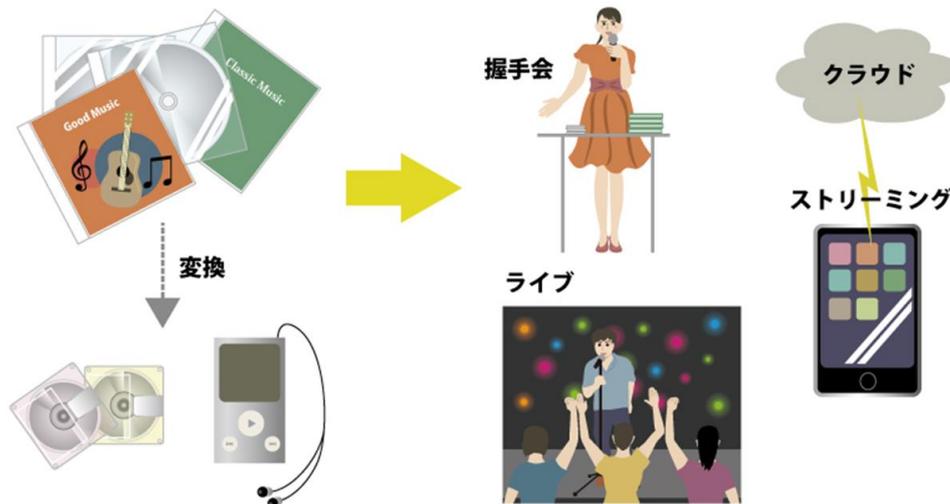


図1：モノ（レコードやCDの所有）からコト（ライブや握手会の体験）への変化

### 1. MPEG-4 ALS の概要

MPEG-4 Audio Lossless Coding (ALS) は ISO/IEC で規格化されたロスレス音響符号化の国際標準である。圧縮してもデジタル信号は必ず元に戻り、原音のデータが変わることは一切無い [3, 4, 5]。電子メールにファイルを添付するときに圧縮してから送ることが多いが、受信側でファイルが一切変わらないのと同様である。

原音に戻ること保証するため、ロスレス符号化のビットレートは指定することができず、入力信号に依存してしまう。例えば、白色雑音を入力した場合はほとんど圧縮できず PCM と同等のビットレートが必要となる。逆に、ほとんど無音に近い音を入力した場合にはロッキー符号化のビットレート (128 kbps や 256 kbps) を下回ることもある。図2に表すように、オーディオのビットレートはサンプリング周波数と量子化解像度とチャンネル数の積で表すことができる。ハイレゾの場合は CD に比べ、音のデータを入れる箱の体積が大きくなっていると解釈することもできる。データを入れる箱は大きくなっても、実際に詰め込む音には低域と高域で密度に差があるため、箱の形を少し変えることにより、音の情報はそのまま箱の体積を減らすことができる。これがロスレス符号化の比喩的概念である。白色雑音の場合には箱全体に均等に音情報が含まれているため箱の形を変えることができず、PCM と同じ大きさになってしまう。ハイレゾの箱にも関わらず低域しか音の成分が無かったり、量子化解像度が粗かったり、ほとんど音量の無いチャンネルが混ざっている場合には、ロスレス符号化により入力音の情報はそのままでも大幅に箱の体積を小さくして圧縮することができる。

図3は MPEG-4 ALS のブロック図である。線形予測分析とエントロピー符号化を組み合わせることにより圧縮を行っている。線形予測分析は音声符号化の分野で長く使われてきた技術で、携帯電話や VoLTE (Voice over Long-Term Evolution) で使われている。入力信号ベクトル  $\mathbf{x}$  を入力して線形予測分析を行い、残差信号ベクトル  $\mathbf{e}$  を求める。線形予測係数は PARCOR (PARTIAL autoCORrelation) 係数や LSP (Line Spectral Pairs) 係数に等価変換し量子化され

たものが使われる。携帯電話や VoLTE で使われている音声のロッシェ符号化の場合は残差信号ベクトル  $e$  をモデル化して低いビットレートで伝送するため元の音に戻すことはできないが、ロスレス符号化 MPEG-4 ALS の場合には残差信号ベクトル  $e$  を全部まじめに伝送するため、ビットレートは高くなってしまいが、元の音に戻すことができる。エンコーダとデコーダで同じ予測フィルタを用いるため、入力信号ベクトル  $x$  と同じ整数値信号を得ることができる。このように MPEG-4 ALS は長年使われてきている技術をもとにして単純な仕組みで歪みの全くない圧縮を実現している。

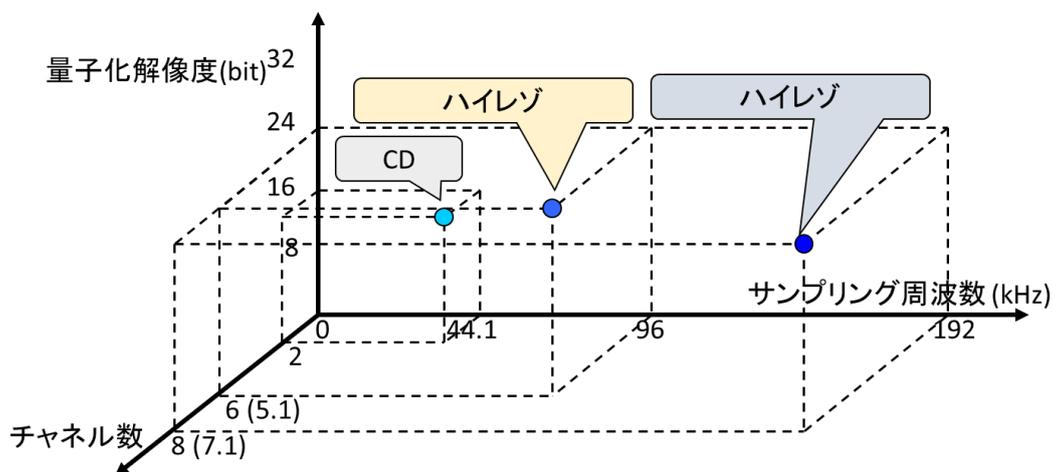


図2：音を入れる箱の体積の概念図

オーディオのビットレート＝サンプリング周波数×量子化解像度×チャンネル数

MPEG-4 ALS は仕様上では約 4 GHz サンプリング周波数、32 bit 量子化解像度、65536 チャンネルまで対応しているが、テレビやサウンドバー・デジタルアンプへの実装を考量して演算量やメモリ量を制限するために、Simple Profile が定義されている。放送用途が主な利用目的であったため、表 1 で示すようなパラメタで制限がされている。もちろん 384 kHz や 24 ch で伝送したい事業者がおり、普及が見込まれる場合には Level を拡張することも可能である。4K/8K 放送では Simple Profile の Level 2 と Level 4 を使うことができるため、放送局で制作された 48 kHz, 24 bit の原音をそのまま家庭まで届けることができる。

MPEG-4 ALS ではアルゴリズムの詳細が規定されているソフトウェア（リファレンス・ソースコード）だけではなく、国際標準準拠であるか確認するためのコンFORMANCE・ビットストリームも用意されている。MPEG で国際的な協力の下での長期的なメンテナンスがなされ、他の MPEG 標準技術との親和性も高い。4K/8K 放送では映像符号化 H.265/MPEG-H HEVC(High Efficiency Video Coding)とロスレス音響符号化 MPEG-4 ALS を組み合わせて MPEG-H MMT (MPEG Media Transport) によって多重化して MPEG ファミリーの技術を組み合わせることで高品質の映像と音声を伝送することができる。ハイレゾ音源配信向けには MPEG-4 ALS と MPEG-DASH (Dynamic Adaptive Streaming over HTTP) を組み合わせた

高臨場感音楽ライブ伝送の試みも行われている [6]。このように MPEG-4 ALS は国際標準であるため、フリーソフトとは異なり、安心して長期的なビジネスにも利用することができる。

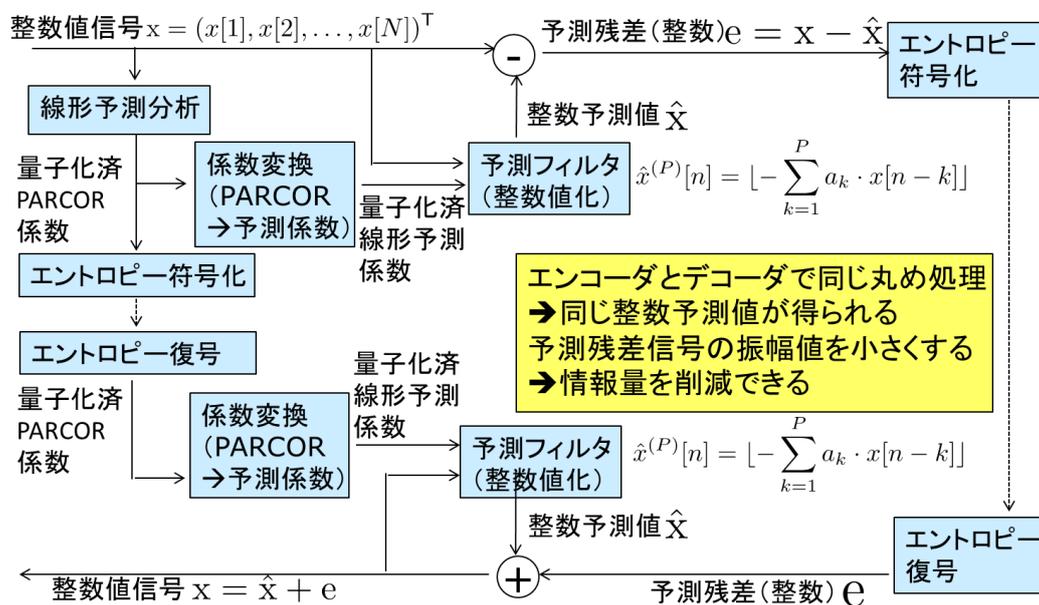


図 3：MPEG-4 ALS の単純化したブロック図

Level	Max. # of channels	Max. sampling rate (kHz)	Max. word length (bit)	Max. # of samples per frame	Max. prediction order	Max. BS stages	Max. MCC stage
1	2	48	16	4096	15	3	1
2	2	48	24	4096	15	3	1
3	6	48	16	4096	15	3	1
4	6	48	24	4096	15	3	1

表 1：MPEG-4 ALS Simple Profile

(BS: Block Switching tool, MCC: Multi-Channel Coding tool)

## 2. ALS に関連する標準化

4K/8K 放送では図 4 で示すように、ALS を使えばテレビ局のスタジオで制作されたコンテンツを原音のまま家庭で楽しむことができるような仕組みが制定されている。一方、AAC を用いてチャンネル方向の拡張を行い 22.2ch の空間的臨場感を再現する方法も利用できる。AAC と ALS の音響符号化切り替えをオーディオ内で閉じて行うために、MPEG-4 Audio で規定されて

いる Low Overhead Audio Transport Multiplex (LATM) / Low Overhead Audio Stream (LOAS) 形式のストリームを用いて伝送する。LATM/LOAS ヘッダを解釈すれば AAC の 22.2ch 音源なのか ALS の 2ch 音源なのかが分かり、ユーザは音声符号化方式を選択して視聴することができる。

4K/8K 放送の運用規定では、ALS は必須の符号化方式ではないため実装しないテレビも存在する可能性がある。この場合、良い音を楽しみたい視聴者は ALS に対応したサウンドバーやデジタルアンプを用意しなければならない。テレビからサウンドバーやデジタルアンプに ALS のビットストリームを伝送するために、IEC 61937-10 という国際標準規格がある [7]。CD の音を MD (Mini Disk) へ伝送するときいわゆる SPDIF で CD プレーヤと MD レコーダをオプティカルケーブルやコアキシャルケーブルで接続してきたが、ここで PCM 伝送を行うための規格が IEC 60958 である [8]。本来 PCM を伝送するための規格であるが、圧縮ビットストリームも伝送することができるように規格化したものが IEC 61937 シリーズである。4K/8K 放送向けに、IEC 61937-10 に LATM/LOAS 形式の ALS をそのまま透過する方式を追記作業中であり、近々第二版として IEC 61937-10 Edition 2 が標準化される見込みである。MPEG-4 AAC を SPDIF で伝送するための規格である IEC 61937-11 も 22.2ch 対応のために改訂中であり [9]、これらの規格が揃えば視聴者は自分のお気に入りのオーディオシステムで ALS による原音や AAC による 22.2ch 音源を楽しむことができるようになる。HDMI (High-Definition Multimedia Interface) 規格は IEC 規格を参照することが多いため、テレビとサウンドバーやデジタルアンプを HDMI で接続することにより ALS や AAC のビットストリームを伝送することができるようになる日が来るかもしれない。

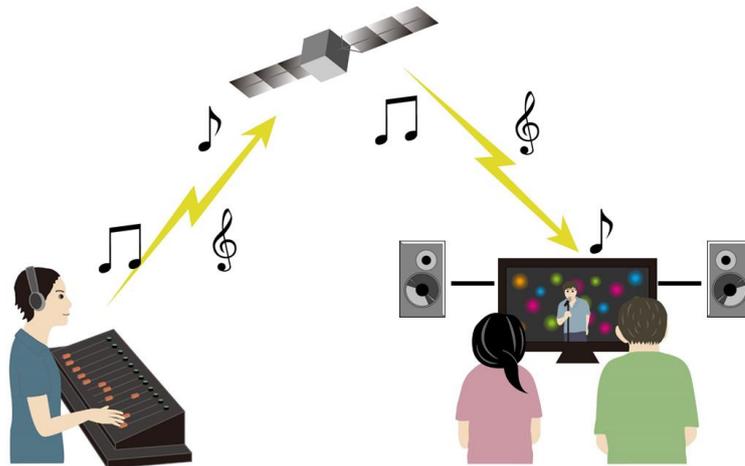


図 4：MPEG-4 ALS を用いた 4K/8K 放送の概念図

## 3. おわりに

音質を犠牲にしなければそもそも音が伝送できなかった時代とは異なり、近年は放送でも通信でも高品質音源を伝送できるようになってきている。伝送する仕組みは整ったので、コンテンツを製作する方々の音へのこだわりが、良い音を求めるユーザに届けることができるようになってきている。現在放送で使われている 48 kHz, 24 bit の音声信号は JEITA ではハイレゾと呼べるが日本オーディオ協会ではハイレゾとは呼べない [10, 11]。ハイレゾを普及させるためにも、映像と音声と一緒に場合には日本オーディオ協会の定義でも 48 kHz, 24 bit の音声信号もハイレゾと呼んでも良いのではないだろうか？そうすれば 4K/8K 放送で ALS を使った場合にはハイレゾ放送と呼ぶことができ、一般家庭へのハイレゾ普及が期待できる。そのような状況になれば、ハイレゾの良い音を聞くのが当たり前になると想定される。

## ■参考文献

- [1] ARIB STD-B32, “デジタル放送における映像符号化、音声符号化及び多重化方式,” 2014.
- [2] ARIB TR-B39, “高度広帯域衛星デジタル放送運用規定,” 2016.
- [3] ISO/IEC 14496-3:2009, Information technology—Coding of audio-visual objects—Part 3: Audio, 2009.
- [4] 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.
- [5] 遠藤, 原田, 鎌本, 守谷, “MPEG-4 オーディオ・ロスレスとストリーミング配信について,” JAS Journal, Vol.52, No.1, pp.22--29, January, 2012.
- [6] [https://labevent.ecl.ntt.co.jp/forum2017/elements/pdf\\_jpn/01/A-g\\_j.pdf](https://labevent.ecl.ntt.co.jp/forum2017/elements/pdf_jpn/01/A-g_j.pdf)
- [7] IEC 61937-10:2011, “Digital audio - Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 - Part 10: Non-linear PCM bitstreams according to the MPEG-4 audio lossless coding (ALS) format,” 2011.
- [8] IEC 60958, “Digital audio interface,” 1989.
- [9] IEC 61937-11:2010, “Digital audio - Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 - Part 11: MPEG-4 AAC and its extensions in LATM/LOAS,” 2010.
- [10] [http://home.jeita.or.jp/page\\_file/20140328095728\\_rhsiN0Pz8x.pdf](http://home.jeita.or.jp/page_file/20140328095728_rhsiN0Pz8x.pdf)
- [11] <https://www.jas-audio.or.jp/hi-res/definition>

■著者プロフィール



鎌本 優（かまもと ゆたか）

2005年 日本電信電話株式会社 入社。以来、音声音響符号化の研究開発と標準化（MPEG-4 ALS, ITU-T G.711.0, 3GPP EVS codec, IEC 61937-10 Ed. 2, ARIB STD-B32 など）に貢献。



守谷 健弘（もりや たけひろ）

1980年 日本電信電話公社 入社。以来、音声音響符号化の研究開発と標準化（PDC, MPEG, ITU-T, 3GPP, IEC, ARIB STD-B32 など）に貢献。



原田 登（はらだ のぼる）

1997年 日本電信電話株式会社 入社。以来、音声音響符号化の研究開発と標準化（MPEG-4 ALS, MPEG-A PA-AF, ITU-T G.711.0, 3GPP EVS codec, IEC 61937-10, ARIB STD-B32 など）に貢献。