

## 音響ロスレス符号化 MPEG-4 ALS の ハイレゾ音源適応の検討と考察\*

☆天田将太 (筑波大), 鎌本優, 原田登, 杉浦亮介, 守谷健弘 (NTT), 山田武志, 牧野昭二 (筑波大)

### 1 はじめに

ハイレゾ音源とはサンプリング周波数または量子化ビット数のいずれかが CD スペックを超えており、もう一方も CD スペックと同等以上である高音質デジタル音源のことである<sup>[1]</sup>。ハイレゾ音源の情報量は同録音時間の CD 音源と比べて数倍となる。一般的な音楽用 CD 容量である 700 MB 分の記録をするに CD スペックでは約 66 分間を記録できる。しかしこれを 96 kHz, 24 bit ハイレゾで記録すると約 20 分間しか記録できない。そのため圧縮の必要性がより高まるが、音質を高く保つために可逆圧縮が必要である。ハイレゾ音源を効率よく伝送するために可逆圧縮方式の圧縮率を改善していくことが求められる。

動画や音声データを扱う国際標準化団体である MPEG (Moving Picture Experts Group) では音響信号の可逆圧縮方式として MPEG-4 Audio Lossless Coding (ALS)<sup>[2, 3]</sup> を規格化している。

MPEG-4 ALS は、サンプリング周波数 192kHz、量子化ビット数 32 bit にまで対応するなど CD スペックの音源のみならず、ハイレゾを含む幅広い入力信号を扱える。放送用途では最大チャンネル数やサンプリング周波数、量子化ビット深度を限定して、受信端末に実装しやすくした MPEG-4 ALS Simple Profile が定義されている (Table.1)。本稿では ALS Simple Profile に定義されていないサンプリング周波数 96 kHz のハイレゾ音源に MPEG-4 ALS を適用した場合の性能を調査し、ハイレゾ音源の圧縮に適した符号化パラメータを検討した結果について報告する。

Table 1 The allowed parameters of each level of MPEG-4 ALS Simple Profile.

レベル	1	2	3	4
最大チャンネル数	2	2	6	6
最大サンプリング周波数 [kHz]	48	48	48	48
最大量子化ビット数	16	24	16	24
最大フレーム長	4096	4096	4096	4096
最大予測次数	15	15	15	15

### 2 MPEG-4 ALS のハイレゾ音源適応

#### 2.1 MPEG-4 ALS

MPEG-4 ALS では入力信号について線形予測分析を行い、予測残差信号と、予測信号を得るための係数 (PARCOR 係数) の 2 つを用いて表現することで可逆圧縮を行っている。ここで予測次数を大きくすれば予測残差信号を伝送するために必要な情報量は小さくなるが、PARCOR 係数を伝送するために必要な情報量は大きくなってしまふ。つまり予測残差信号のための情報量と PARCOR 係数のための情報量にはトレードオフの関係がある。よって全体の情報量がより小さくなる適切な予測次数を検討する必要がある。MPEG-4 ALS では可変予測次数機能が採用されており、適切な予測次数を求めることにより圧縮性能を改善することができる。また入力信号はフレームに分割されて処理されるため、フレーム長は圧縮性能に影響を与える。そこで MPEG-4 ALS にハイレゾ音源を適用する際に適切な予測次数とフレーム長について調査を行った。

#### 2.2 フレーム長と最大予測次数

MPEG-4 ALS Simple Profile では 48 kHz 音源について最大フレーム長は 4096、最大予測次数は 15 次を用いている。96 kHz 音源でフレーム長を 8192、最大予測次数を 31 次として 48 kHz 音源の 2 倍にすれば、同じ時間区間で処理することに相当する。本稿ではこの場合の圧縮性能に着目して解析を行った。

### 3 調査実験

#### 3.1 実験条件

実験に用いた音源の情報を Table 2 に示す。ジャズやオーケストラ、オペラを含む 15 音源を用意し、サンプリング周波数 96 kHz の音源をダウンサンプリングすることでサンプリング周波数 48 kHz の音源を作成した。

#### 3.2 調査結果

48 kHz 音源および 96 kHz 音源の各フレーム長で可能な予測次数の範囲を全探索して求めた最適予測次数の分布を Fig.1 と Fig.2 に示す。それぞれ各フレー

\* Experimental evaluation of the parameters of MPEG-4 ALS for High-resolution audio. by Shota Amada (The University of Tsukuba), Yutaka Kamamoto, Noboru Harada, Ryosuke Sugiura, Moriya Takehiro (NTT), Takeshi Yamada and Shoji Makino (The University of Tsukuba).

Table 2 Information of input sound items.

サンプリング周波数 [kHz]	48	96
チャンネル数	2	2
量子化ビット数	24	24
音源数	15	15
各音源の長さ [s]	30	30
ファイルサイズ ( $\times 10^6$ ) [byte]	8.64	17.28

ム長における各次数の出現頻度とピークをプロットした。最適予測次数のピークについて、48 kHz 音源では 15 次付近にあることが読み取れる。96 kHz 音源ではどのフレーム長においても 15 次を超えており、フレーム長 4096 では 31 次付近に割合が多い。サンプリング周波数の違いによって最適予測次数の分布が異なることが確認できる。

48 kHz 音源および 96 kHz 音源のそれぞれについて、フレーム長や最大予測次数を変えて MPEG-4 ALS を適用した場合の圧縮率の変化を Fig.3 と Fig.4 に示す。48 kHz 音源において圧縮率が最も高いフレーム長は 2048 であり、96 kHz 音源では 4096 である。さらに最大予測次数に注目すると 15 次から 31 次へ広げた場合に圧縮率が大きく向上する。

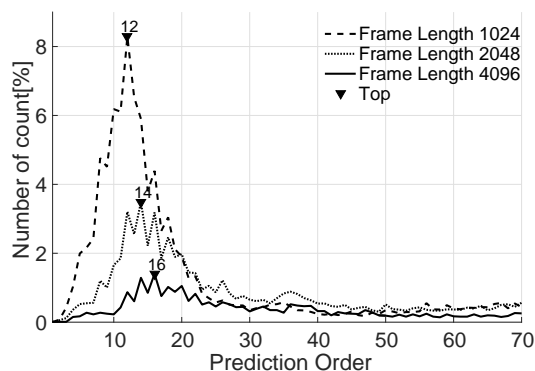


Fig. 1 Histogram of optimal prediction orders for each frame length of 48 kHz signals.

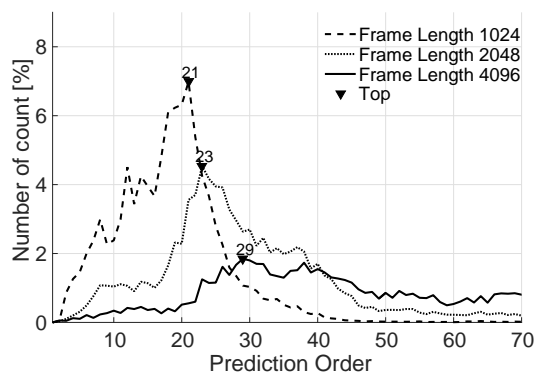


Fig. 2 Histogram of optimal prediction orders for each frame length of 96 kHz signals.

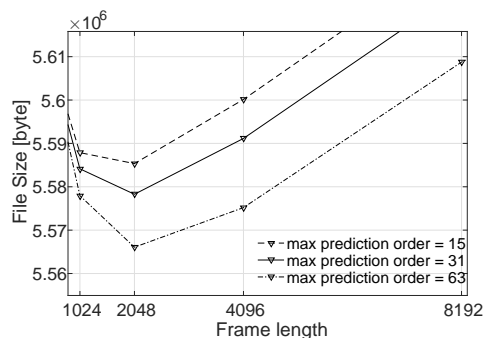


Fig. 3 The relationship between the encoded file size and frame length for 48 kHz signals.

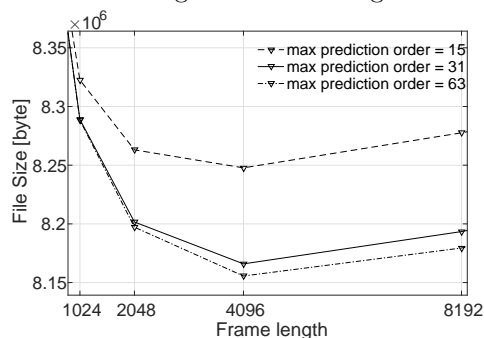


Fig. 4 The relationship between the encoded file size and frame length for 96 kHz signals.

#### 4 まとめ

最大予測次数や最大フレーム長を引き上げることで圧縮率の向上が見込めるが、一方で復号化や符号化に必要な演算量とバッファサイズは増加する。

サンプリング周波数 96 kHz のハイレゾ音源に MPEG-4 ALS を適用する場合、Simple Profile に定義されている 48 kHz 音源と同様に最大フレーム長 4096、最大予測次数 15 次としても十分な圧縮性能が見込める。ただし、フレーム長 8192 が最適な音源も半数ほど存在する。また最適予測次数は 15 次を超えたところに多く分布しているため最大予測次数を 31 次へ引き上げることで圧縮率のより大きな向上が見込める。

最大次数と最大フレーム長は用途に応じた選択をすることが重要である。圧縮率の向上を優先するならば最大予測次数 31 次、最大フレーム長 8192 とすべきであり、演算量やバッファサイズの増加を抑え、短時間計算および省消費電力を優先するならば最大予測次数 15 次、最大フレーム長 4096 とすべきであると考えられる。

#### 参考文献

- [1] 25JEITA-CP 第 42 号, 2014.
- [2] ISO/IEC 14496-3:2009/Amd 5:2015 ,2009
- [3] T.Liebechan, et. al., “The MPEG-4 Audio Lossless Coding (ALS) Standard - Technology and Applications ” ,2005.