

チャンネル間相関を用いた多チャンネル信号の可逆圧縮符号化

Lossless Compression of Multi-Channel Signals Using Inter-Channel Correlation

鎌本 優[†] 守谷 健弘[‡] 西本 卓也[†] 嵯峨山 茂樹[†]
Yutaka Kamamoto Takehiro Moriya Takuya Nishimoto Shigeki Sagayama

1. はじめに

本稿では、多チャンネル信号の可逆圧縮符号化において、チャンネル間相関を用いる手法を提案する。

これまで、聴覚心理モデルに基づく圧縮率の高いオーディオ信号の圧縮法が研究され、デジタル放送、ミニディスク、インターネット音楽配信に広く使われているが、原信号の完全復原はできない不可逆圧縮方式であった。一方、ブロードバンドの普及や記憶装置の容量増加により、圧縮率よりも原音の完全再生を条件とした圧縮が重視されつつある。特に音楽などのデータ蓄積や編集加工では原音の完全再生が必須である [1]。そのため、MPEG-4 Audio Lossless では、音響信号の可逆圧縮符号化の国際標準が策定されつつある。

多チャンネル信号の可逆圧縮符号化は音響信号だけでなく、生体信号や地震波信号、センサーアレイ信号にも応用することにより、医療、気象、環境、音場合成用途への応用が期待される。これらの用途の場合には、チャンネル数は 100 以上となる場合も多く、チャンネル間の信号の相関を柔軟に利用した分析、符号化法により圧縮率の改善が期待できる。どのような分野でも、可逆圧縮ならば安心して利用できる。しかし、これまで多チャンネルの信号はステレオの対に分解して符号化する手法が暫定規格となっており、改善の余地がある。

2. 従来法の可逆圧縮符号化の基本原則

2.1 MPEG-4 Audio Lossless Coding

現在策定中である MPEG-4 Audio Lossless の時間領域での圧縮符号化法は、線形予測とエントロピー符号化が技術の中心となっている [2]。これらの処理は、入力信号を時間方向の 2048 サンプル程度のフレームに分割し、さらにチャンネル毎に分割されたブロックに対して行われている。以下に、その技術要素の基本原則を述べる。

2.2 線形予測

入力信号が整数値を取るとすれば、その過去の p サンプルを用いて現在の値 x_t は、

$$\hat{x}_t = \left[-\sum_{k=1}^p a_k \cdot x_{t-k} \right] \quad (1)$$

によりの線形予測される。ここで a_k ($k = 1, 2, \dots, p$) は線形予測によって得られ、伝送のために量子化された予測係数で、 $[\cdot]$ は整数化を表す。予測残差 $e_t = x_t - \hat{x}_t$ も整数値として得られる。この予測残差と予測係数お

び初めの p サンプルを送り、 $x_t = \hat{x}_t + e_t$ により原信号を復元することができる。

2.3 エントロピー符号化

上記の予測残差信号 e_t の値は、振幅が原信号 x_t に比べ予測利得に応じて小さくなり、0 付近に集中する傾向がある。このような分布の偏りを利用すれば、エントロピー符号化により平均符号長を短くすることができる。このように、整数値信号の振幅を縮小できれば圧縮率を向上できる可能性が高い。本稿では以後このような観点から論じる。

2.4 Joint Stereo Coding

左右チャンネルの信号の差 $d_t = x_t^L - x_t^R$ をとり、 x_t^L , x_t^R , d_t の 3 つの信号をそれぞれ符号化し、符号長の短い 2 つを送る方法が現在用いられている。どの組合せでも原信号を復元できる。

この手法は多チャンネル用にも用いられているが、チャンネル間相関を考慮せずにチャンネル番号順に対を作り、符号化処理が行われている。

3. 残差信号とチャンネル間相関の利用

3.1 残差信号の重み付き減算

我々は、各チャンネルの残差信号間の相関を利用して重み付きの減算することにより振幅が低減され、圧縮率が向上する可能性があるという点に着目した。

まず、予備実験において、線形予測の残差信号の差分を処理する方が、原信号の差分の線形予測誤差を用いるよりも、振幅の縮小効果が高い傾向が見られた。また、残差信号はフレーム境界のサンプルの扱いに特別な注意を払わなくてもよい利点もあり、以下のようなチャンネル間の重み付きの差分をとる際に扱いやすい。以上から、残差信号のチャンネル間相関を利用する方針を取る。

1 チャンネルの 1 フレームの残差信号 $\{e_t\}$ をベクトルと見なして $e = (e_0, e_1, \dots, e_{n-1})^T$ と表す。チャンネル i の残差信号 e^i の振幅を低減するために、チャンネル j の残差信号 e^j を用いて、 $\tilde{e}^i = e^i - \gamma^{i,j} \cdot e^j$ のように重みつき差信号を構成する。チャンネル i の残差信号は、 $e^i = \tilde{e}^i + \gamma^{i,j} \cdot e^j$ により復元できる。ただし、 $\gamma^{i,j}$ は差信号の二乗ノルム $\|\tilde{e}^i\|^2$ を最小化するように、

$$\gamma^{i,j} = \frac{e^{i,T} \cdot e^j}{e^{j,T} \cdot e^j} \quad (2)$$

により与える。以後、引かれるチャンネル i を子と呼び、引くチャンネル j を親と呼ぶことにすれば、差信号 \tilde{e}^i は、子信号から親信号へ下ろした垂線ベクトルである。

[†]東京大学, The University of Tokyo

[‡]日本電信電話株式会社, NTT

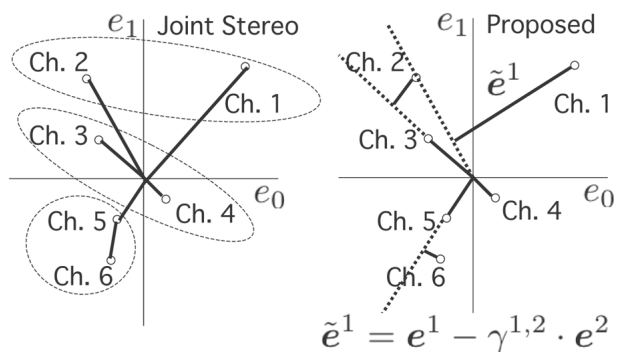


図 1: Joint Stereo 方式と本手法におけるチャンネル間の関係の幾何学的な解釈 (2 サンプル値の場合を図示)

差分処理による振幅縮小効果は、この親子関係の選び方に依存するので、信号に応じて適応的決めることが必要である。なお、この方法による情報削減効果が、親番号と重み γ を表す情報より少なければ、元の残差信号をそのまま符号化する。

3.2 親子関係の木構造表現と探索

信号のチャンネル数を C 、1 フレームを構成するサンプル数を N とすると、 C チャンルの信号は N 次元の信号空間中の C 個の点で表わされる。これらは、 C 点全てをつなぎ合わせる親子関係により、原点を根とする木構造で表され、個々の信号は復元できる。振幅の低減の効果が高くなるように親子関係の木構造を選べば、圧縮率の向上が期待できる。しかし、この問題は複雑な組合せ最適化問題になるので、実際的な方策が必要である。

図 1 は、以上の概念を幾何学的に表現したものである。図としての制約から、信号ベクトル次元 (フレーム内サンプル数) を 2、チャンネル数を 6 の場合を表現している。既存の Joint Stereo (同図左) では、チャンネル番号順に信号対を構成し、対ごとに 2 つの原信号と 1 つの差信号の符号長を比較して有利な 2 つを選ぶ。すでに述べたように符号長は信号の振幅とおおむね対応するので、この場合の符号長は図中の実線部分と対応する。すべてのチャンネル間で組み合わせて実線の長さの総和を小さくするような配慮はされていない。一方、提案手法 (同図右) では、木構造のチャンネル間の親子関係をもち、さらに係数 γ を用いて図中の破線で表される親のベクトル方向へ下ろした垂線で信号が表現できるため、距離の総和を小さくでき、符号長が短くなることが期待できる。

チャンネル間の親子関係の探索については、我々は、距離の総和最小化は保証されないが、それを近似する実際的なアルゴリズムを案出して用いている。

4. 評価実験

4.1 評価方法

MPEG-4 Audio Lossless の規範モデルに提案法を実装し、性能を確かめるために評価実験を行った。ZIP、MPEG-4 Audio Lossless Independent (mp4als-

表 1: データ圧縮率の比較 [%] (括弧内は mp4als-I を 100%とした相対値)

評価データ	圧縮手法			
	ZIP	mp4als-I	mp4als-J	提案手法
Audio 2ch	93.6	49.9	49.7 (99.8)	49.5 (99.4)
Audio 8ch	80.7	50.5	50.2 (99.5)	50.0 (99.0)
MEG-raw	58.0	39.9	39.7 (99.6)	39.1 (98.1)
MEG-ave	45.3	17.4	17.4 (99.9)	17.1 (98.5)
MEG-anr	44.4	14.8	14.7 (99.5)	14.5 (98.4)

I)、MPEG-4 Audio Lossless Joint Stereo (mp4als-J)、提案手法の 4 つの手法を比較し、原信号ファイルサイズに対する圧縮後のファイルサイズを % 値で表 1 に示す。ZIP は広く使われているファイル圧縮手法である。mp4als-I はチャンネル毎に独立に線形予測符号化したものであり、mp4als-J は前述の Joint Stereo 方式を表す。

実験評価データとして音響信号と脳磁計 (MEG) 信号を用いた。音響信号は、ステレオの音楽データ (Audio 2ch) と 8 チャンネルのマイクロホンアレー受信信号 (Audio 8ch) を用い、MEG 信号は、計測した信号そのもの 256 チャンネル (MEG-raw) と、同期加算平均をとったもの 256 チャンネル (MEG-ave) と、雑音除去をしたもの 192 チャンネル (MEG-anr) の 3 つを用いた。

4.2 結果と考察

いずれの信号についても、既存の手法よりも圧縮率を向上させることができた。チャンネル数が多いほど圧縮効果が大きい傾向があるので、チャンネル間相関を利用した効果が現れていると考えられる。また、チャンネル関係の木構造生成方法を改善すれば、さらに圧縮性能を向上できる可能性がある。

5. まとめ

チャンネル間相関を利用する多チャンネル信号の可逆圧縮符号化の手法を提案し、従来法に比べて圧縮性能を向上させられることを実験的に確認した。また、音響信号だけでなく、生体信号にも応用できることを確認した。

謝辞

MEG のデータを提供して下さった、東京大学 新領域創成科学研究科 武田常弘教授 大脇崇史助手に感謝いたします。

参考文献

- [1] Mat Hans, et al.: "Lossless Compression of Digital Audio," *IEEE Signal Processing Magazine*, pp. 21–32, July 2001.
- [2] Tilman Liebchen, Yuriy A. Reznik: "MPEG-4 ALS: an Emerging Standard for Lossless Audio Coding," *Proc. IEEE Data Compression Conference 2004*, pp. 439–448, 2004.