

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

鎌本 優^{†,††}, 守谷 健弘^{†,††}
西本 卓也[†] 嵯峨山 茂樹[†]

本稿では、多チャンネル時系列信号を対象に、各チャンネルごとの時間領域の線形予測とその残差信号のチャンネル間相関に基づく適応的差分を柔軟に利用した可逆圧縮符号化方法を提案する。予測残差信号のチャンネル間相関関数を基準に、符号化対象のチャンネル信号から参照チャンネルに最適重み係数を乗じた信号を引いた差信号を求め、信号の振幅が小さくなる傾向を利用してエントロピー符号化を行う。参照チャンネルと符号化対象チャンネルは逐次芋づる式に求め、参照チャンネルの番号と重み係数を補助情報とする。また全体のチャンネルを複数の集合に分割することで、処理量の増大を防ぐ手法や、重み係数と参照チャンネルの探索を複数回繰り返す改善法を示した。通常の2から8チャンネルの音響信号、256チャンネルの脳磁計を対象に本手法の符号化実験を行い、2チャンネルを対にした従来のジョイント・ステレオ符号化より最大約3%圧縮率を改善できることを示した。これらの提案手法はISO/IECのMPEG (Moving Picture Experts Group) で標準化が進められているロスレス・オーディオ信号の符号化モジュールとして提案され、オーディオ信号だけでなく、広範囲の時系列信号の符号化に国際的に広く利用される見込みである。この研究は東京大学大学院情報理工学系研究科と日本電信電話株式会社との産学連携講座に基づく共同研究の成果である。

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

YUTAKA KAMAMOTO,^{†,††} TAKEHIRO MORIYA,^{†,††}
TAKUYA NISHIMOTO[†] and SHIGEKI SAGAYAMA[†]

Lossless compression method has been proposed for multi-channel time series, where time-domain linear prediction residual signals are adaptively subtracted according to the inter-channel correlation analysis. The residual signal of objective channel is subtracted by the signal of reference channel multiplied by an optimum weighting factor, and is efficiently coded by entropy coding making use of the amplitude reduction. The relationship between the reference channel and the coding channel is sequentially searched. Each channel includes the reference channel number and the weighting factor as side information. Additionally, complexity reduction method in reference channel search procedure and the iterative application of the proposed coding procedure are shown. As a result of coding experiments for 2-8 channels audio signals and 256-channels MEG signals, up to 3% of relative improvement of compression ratio have been observed compared to the conventional joint-stereo coding. All these methods have been proposed to the standardization activities of lossless audio coding at ISO/IEC MPEG (Moving Picture Experts Group), and may be widely used not only for audio signals but also other general time series. This result is based on the joint research work of a technical liaison program between the University of Tokyo and NTT.

1. はじめに

今までにオーディオ信号の圧縮符号化の研究がさかんに行われ、インターネット音楽配信 (MPEG-

1 LayerIII/TwinVQ などを使用)、ミニディスク (ATRAC/ATRAC3 を使用)、地上波デジタル放送 (MPEG-2 AAC を使用) などに広く使われてきている。これらの圧縮符号化法は時系列信号の相関や分布の偏りを利用し、さらに聴覚心理モデルに基づく歪制御を行っている。これにより、原信号の完全復元はできない不可逆圧縮方式であるが、聴感的な品質をほとんど損なわずに、原音の情報量を 1/10 程度以下まで圧縮できる¹⁾。

これに対し、特に音楽などのデータ蓄積保存や編集

[†] 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology,
The University of Tokyo

^{††} 日本電信電話株式会社コミュニケーション科学基礎研究所
NTT Communication Science Laboratories
現在、日本電信電話株式会社
Presently with NTT

加工では原音の完全再生が必須である。また利用者からも高品質化の要望があり、高いサンプリングレート、大きい振幅語長、多チャンネルの信号への要求が高まりつつある。音響空間をそのまま表現することを試みる Wave Field Synthesis (波面合成) 信号、生体信号、地震波信号、センサアレー信号なども多チャンネル化の傾向があり、チャンネル数は 100 以上となる場合も多い。これらの用途では情報量の増大が問題となりつつあり、また、対象別の圧縮法 (たとえばオーディオ信号の場合は聴覚特性を利用) は使えないため、多チャンネル信号の完全再生を保証する汎用の圧縮符号化が必要である。

これまでに、時系列信号、特にオーディオ信号に対するいくつかの可逆圧縮符号化の研究開発が進められていて、時間領域の予測または周波数領域への変換と、エントロピー符号化が有力である²⁾。このうち、DVD のオーディオ信号用のロスレス符号化 MLP や SACD (Super Audio CD) の 1 ビットオーディオ信号に対する可逆圧縮符号化が商品に組み込まれている³⁾。しかし、これまでは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などが主体で^{2),3)}、多チャンネルの相関を積極的に利用した研究は少ない。また、チャンネルごとに独立に符号化するほうが圧縮率が高くなる場合もあり、実用的利点が少なかった。

本稿ではチャンネルごとの時間領域の予測符号化を基礎にし、その性能を失うことなく、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する。

2 章では、時系列信号の可逆圧縮の典型例の 1 つである時間領域の線形予測符号化について述べ、これを発展させた提案手法を 3 章で述べる。4 章では、提案手法を評価するための圧縮実験結果を示し、演算量と効率を議論する。

実用的な符号化方式では、国際的な互換性が保証されること、計算プラットフォームに依存しないこと、少なくとも 100 年単位での保守継続性があることが必要であり、オープンソースでの開発が重要となる。そのため、ISO/IEC の MPEG (Moving Picture Experts Group) により、2002 年より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。本稿の提案手法は 2004 年 7 月に MPEG 標準化にも提案され、他機関との相互検証や評価作業が進行している。

また本研究は東京大学大学院情報理工学系研究科と NTT 研究所との戦略型 IT 連携講座に基づく共同研究の成果であり、双方の得意とする研究背景、動機、実績と密接な議論によって達成された。産から学へは客

員教授の派遣、学から産へは実習生を派遣し、連携をしながら共同で特許出願も行った。このような共同研究体制は以下のような事情で、特に国際標準化の貢献に適していると考えられる。国際標準規格は準備段階、提案改良段階を経て世の中に提供される。標準化に参画する場合には、規格制定後の維持管理まで含めた長期的サポートが要求される。企業としては最終的な商用サービスや製品化を目的にリソースを用意できると同時に、準備段階から維持管理まで長期的にサポートすることが可能である。しかしながら、提案改良段階では 1~2 年の短期間で集中的に専門性の高い新しい要素技術の提案やソフト開発をする必要があり、このリソースは企業では十分に確保できない。この段階では大学における専門性の高い研究リソースを生かした共同研究が有効である。大学にとっても優れた要素技術が国際標準に採用されることは非常に有意義であるが、大学単独では維持管理も含めた標準化の長期的サポートは容易ではない。

2. 従来法の可逆圧縮符号化の基本原則

2.1 MPEG-4 Audio Lossless Coding

時系列信号の可逆圧縮符号化の典型的な例の 1 つは時間領域の予測に基づくものである^{4),5)}。現在策定中である MPEG-4 Audio Lossless Coding 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^{6),7)}。これらの処理は、入力信号を時間方向の 2,048 サンプル程度のフレームに分割し、さらにチャンネルごとに分割されたブロックに対して行われている。本稿では、多チャンネル信号の新しい符号化アルゴリズムをこの枠組みを基礎として設計したので、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 x_i (i は整数) を整数値とし、その過去の p サンプルを用いて現在の値 x_t を、

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

の形式の自己回帰型の線形予測で予測する。ここで a_j ($j = 1, 2, \dots, p$) は線形予測係数であり、 $[\cdot]$ は整数化を表す。

予測係数の符号化には、安定判別、等価変換、量子化が容易な PARCOR (Partial Autocorrelation: 偏自己相関) 係数を用いる。PARCOR 係数は予測残差エネルギーの総和が最小となるように正規方程式を解くことで求め、逆サイン関数で非線形に変換して量子化する。符号器、復号器に共通な量子化後の PARCOR

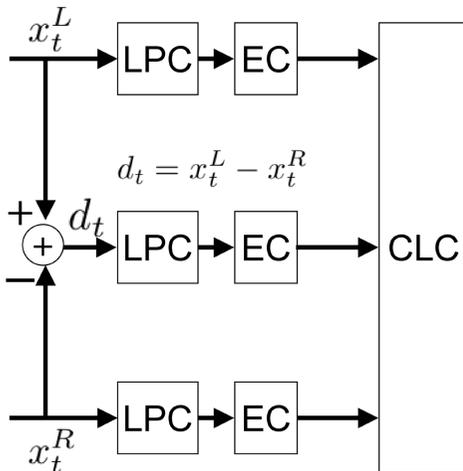


図1 ジョイント・ステレオ符号化方式のブロック図
(LPC: 線形予測残差を求める逆フィルタ, EC: エントロピー符号化, CLC: 符号長比較)

Fig. 1 Block diagram of joint stereo coding.
(LPC: The inverse filter for prediction residual, EC: Entropy Coding, CLC: Code Length Comparison).

係数から線形予測係数を求め、式(1)に従って波形の予測または合成を行う。符号器では予測残差 $e_t = x_t - \hat{x}_t$ が整数値として得られ、この予測残差を完全に伝送すれば、復号器では予測係数および過去の p サンプルから $x_t = \hat{x}_t + e_t$ により再帰的に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 の分散は原信号 x_t に比べ小さくなる(ただし、整数化するために量子化雑音加わり厳密には保証されない)ため、振幅値は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。たとえば、音響信号の予測残差信号の振幅値はラプラス分布で近似でき⁸⁾、ラプラス分布のエントロピーは分散の対数に比例するので、エネルギーが減少すればエントロピーも減少することが期待できる。実際、MPEG-4 Audio Lossless Coding では具体的なエントロピー符号化として Golomb-Rice 符号^{9)~11)}、BGMC (Block Gilbert Moore Code)^{12),13)} が併用されている。

以上のように、整数値信号に可逆圧縮の処理を施した結果、そのエネルギーが減少すれば、情報が圧縮されている可能性が高い。本稿では以後このような観点から論じる。

2.4 ジョイント・ステレオ符号化

ステレオ信号のチャンネル間相関を利用するジョイント・ステレオ符号化として、図1に表されるように、

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つを伝送する方法が用いられている。どの組合せでも原信号を復元できる。ここで、LPC (Linear Predictive Coding) は線形予測残差を求める逆フィルタ、EC (Entropy Coding) はエントロピー符号化、CLC (Code Length Comparison) は符号長比較である。

多チャンネルの場合にもこの手法を用いているが、チャンネル番号順に固定的に対を作って、チャンネル間の相関を分析せずに符号化処理を行っている。

3. チャンネル間相関の利用

3.1 時間方向と空間方向の情報圧縮

前章で述べたように、時間方向の相関を用いて線形予測することが効果的な圧縮技術として用いられている。そこで、同様にチャンネル間でも、相関を用いることで情報圧縮の効果が期待できる。すなわち、従来の時間方向に加えて新たにチャンネル間などの空間方向についても相関を考慮することで、より高効率の情報圧縮を行うことを考える。

近年、マイクロホンアレーや脳磁場計測のように、多チャンネル信号を扱う機会が増えてきており、チャンネル間相関を効果的に用いることで、従来法より圧縮率を向上させられることが期待できる。

3.2 残差信号の重みつき減算

各チャンネルの残差信号間の相関を利用してチャンネル間で適切な重みつきの減算をすれば、より信号振幅を低減でき、圧縮率を向上させられる可能性がある。ここで、提案手法の概要図を図2に示す。

あるチャンネルのあるフレーム (N サンプル) の残差信号 $\{e_t\}$ ($t = 0, 1, \dots, N-1$) をベクトルと見なして $e = (e_0, e_1, \dots, e_{N-1})^T$ と表す。ただし e^T はベクトル e の転置を表す。チャンネル k の残差信号 e^k の振幅を低減するために、チャンネル l の残差信号 e^l を用いて、

$$\tilde{e}^k = e^k - \gamma^{k,l} \cdot e^l \quad (2)$$

のように重みつき差信号を構成する。チャンネル k の残差信号は、 $e^k = \tilde{e}^k + \gamma^{k,l} \cdot e^l$ により復元できる。ただし、重み係数 $\gamma^{k,l}$ は差信号の二乗ノルム $|\tilde{e}^k|^2$ すなわちエネルギーを最小化するように、

$$\gamma^{k,l} = \frac{e^{kT} \cdot e^l}{e^{lT} \cdot e^l} \quad (3)$$

により与える。以後、以上式(2)、式(3)の処理を減

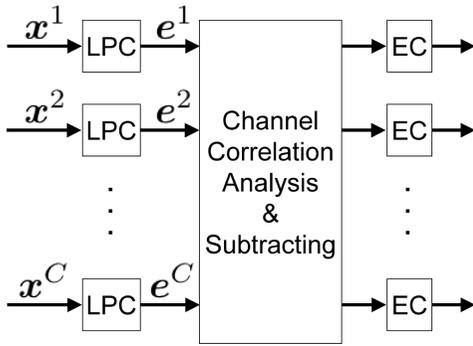


図2 提案手法のブロック図
(LPC: 線形予測残差を求める逆フィルタ, EC: エントロピー符号化, C: チャンネル数)

Fig. 2 Block diagram of proposed method.
(LPC: The inverse filter for prediction residual signal, EC: Entropy Coding, C: Number of channels).

算処理, 引かれるチャンネル k を子, 引くチャンネル l を親と呼ぶことにする. 重みつき差信号 \tilde{e}^k は, 子信号から親信号へ下ろした垂線ベクトルである.

以上では, 時間方向に線形予測を適用してから残差信号のチャンネル間相関を利用して差信号を求めたが, 逆に原信号のチャンネル間相関を利用して差信号を求め, その差信号に線形予測を適用させることも考えられる. そこで, 前者と後者を比較した予備実験を行ったところ, 線形予測の残差信号に提案手法を適用した方が, 線形予測分析の回数が少なく圧縮処理時間が短いにもかかわらず, 圧縮率が約 0.1% 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ることから, 残差信号のチャンネル間相関を利用する方針をとった.

減算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に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, 親チャンネル番号と重み係数 γ を表す情報より小さければ, 元の残差信号をそのまま符号化する. 提案手法を使うか否かを表すフラグが各ブロックごと必要であるが, わずか 1 ビットであり, チャンネル間にまったく相関がないという最悪の場合でも, 数千サンプルの振幅値を数ビットで表していることを考えると 1/1,000 未満の情報量なので, 十分無視できる. よって, 提案手法の性能はチャンネルごとに独立に線形予測符号化したものよりも低下しないと考えてよい.

3.3 親子関係の木構造表現

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 チャンネルの信号は N 次元の信号空間中の C 個の点で表せる. これら C 点す

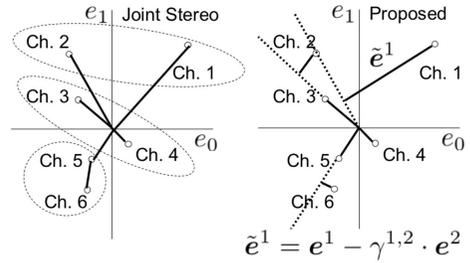


図3 ジョイント・ステレオ符号化方式(左)と提案手法(右)におけるチャンネル間の関係の幾何学的な解釈(2 サンプル値の場合を図示)

Fig. 3 Geometrical interpretation of relationship between joint stereo coding and proposed method. (illustrating the 2-channel case).

べてをつなぎ合わせる親子関係と重みつき差信号が与えられれば, これらは原点を根とする木構造をなすので, 個々の信号が復元できる. そこで, 振幅の低減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, 最適な木構造を求める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3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, 重み係数 γ が正の場合を表現している. 既存のジョイント・ステレオ符号化(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強い相関があるので, この場合の符号長は図中の実線部分と対応する. ここでは, すべてのチャンネル間で組み合わせると実線の長さ(ユークリッド距離)の総和を小さくするような配慮はされていない. たとえば, チャンネル2と3は相関があるにもかかわらず対を構成していないため, それぞれ独立に符号化される. チャンネル1, 2および3, 4は低相関と判定され独立に符号化されることが選択され, 結局チャンネル5と6だけが対を構成して符号化されている. 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持ち, 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. たとえば, チャンネル3, 4, 5は独立に符号化され, チャンネル2はチャンネル3を γ 倍した信号に垂線で表される信号を足すことで表現され, チャンネル1は芽づる式にチャンネル2を γ 倍した信号に垂線で表される信号を足すことで

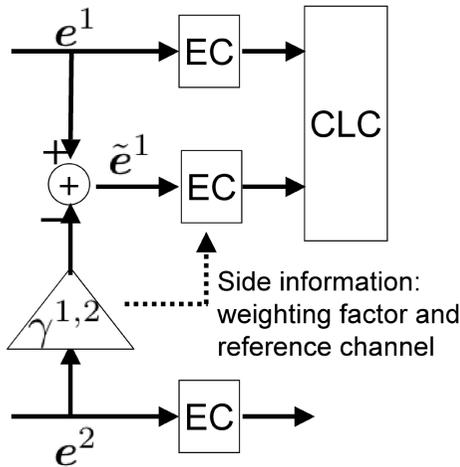


図4 提案手法におけるチャンネル間親子関係探索後のブロック図 (EC: エントロピー符号化, CLC: 符号長比較)

Fig. 4 Block diagram of proposed method after searching optimum relationship between channels.

(EC: Entropy Coding, CLC: Code Length Comparison).

表現され、チャンネル6はチャンネル5を γ 倍した信号に垂線で表される信号を足すことで表現される。このように、提案手法はジョイントステレオ方式と比べて、チャンネル間相関を積極的に利用し情報圧縮を行っている。木構造を求めた後は、図3の親子関係の1つを例として用いると、チャンネル1が子チャンネルとなりチャンネル2が親チャンネルとなるので、図4に示される方法でチャンネル1の重みつき差信号を求め、振幅縮小効果によるエントロピーの低下を利用することで振幅値を表現する符号長は短くなるが、これに補助情報を加えても各チャンネルを独立に符号化する場合に比べて符号長が短いならこれら \tilde{e}^1 と重み $\gamma^{1,2}$ および親チャンネルを表す情報を符号化し、そうでなければチャンネル1の信号 e^1 を独立に符号化する。チャンネル2は親チャンネルなので減算処理を行わずに、信号 e^2 を独立に符号化する。

3.4 親子関係の探索アルゴリズム

3.4.1 近似解を得るアルゴリズム

親子関係の木構造を探索するアルゴリズムは提案手法の性能に直接関わってくる。最適な芽づる式親子関係を求めるのは組合せ最適化問題となり、 $O((n!)^2)$ 通りの親子関係を探索しなければならないが、これは現実的ではなく、チャンネル数が多くなると、計算量が爆発する。そこで、好ましい親子関係を1個ずつ逐次的に決める現実的な近似探索アルゴリズムを提案する。この方法では距離の総和の真の最小化は保証されないが、 $O(n^2)$ で近似解を求めることができる。

また、本アルゴリズムでは、信号全体のエネルギーを小さくするために、重みつき差信号だけでなく減算処理をされていない信号のエネルギーも考慮して探索している。

3.4.2 符号器における処理

復号時には減算処理されていない元の信号が1つ以上必要となるので、初めに、根となる親チャンネル l ($l = 1, 2, \dots, C$)のエネルギーとこれを親とする子チャンネル k ($k = 1, 2, \dots, C; k \neq l$)のエネルギーの和が最も小さくなる親子対を探索する。ここで、親と子を合わせたエネルギーは、

$$\begin{aligned} E^{k,l} &= |\tilde{e}^k|^2 + |e^l|^2 \\ &= |e^k - \gamma^{k,l} \cdot e^l|^2 + |e^l|^2 \end{aligned} \quad (4)$$

と表すことができる。これら $C \times (C - 1)$ 通りの組合せの中から、エネルギーが最小となる親子関係をなすチャンネルを選択し符号化する。

次に、すでに符号化されたチャンネルは親になることができるので、これらから芽づる式につながる子を探索する。符号化されている親チャンネルを n 、それにつながるまだ符号化されていない子チャンネルを m ($m \neq n$)とすると、子だけのエネルギー $\hat{E}^{m,n}$ は、

$$\begin{aligned} \hat{E}^{m,n} &= |\tilde{e}^m|^2 \\ &= |e^m - \gamma^{m,n} \cdot e^n|^2 \end{aligned} \quad (5)$$

と表されるので、(すでに符号化されている親チャンネルの数) \times (まだ符号化されていない子チャンネルの数)通りの組合せの中から、このエネルギーが最小となる子チャンネルを探索して符号化する。新たに符号化された子チャンネルは以後、親チャンネルになることができる。ただし、 $\hat{E}^{m,n}$ よりも小さい $E^{k,l}$ が存在すればそれを採用する。以上を、すべてのチャンネルが符号化されるまで繰り返す。

すなわち、図5で示されるような以下のステップで親子関係を決定する。

- (1) 最小の $E^{k,l}$ を構成する親子関係を符号化。
- (2) $\hat{E}^{m,n}$ を小さい順に探索し、親チャンネル n がすでに符号化されていれば子チャンネル m を符号化。
- (3) 符号化されたものは親になれるので芽づる式に探索。
- (4) $\hat{E}^{m,n}$ が $E^{k,l}$ よりも大きい場合は(1)へ戻る。

この方法に従えば、親チャンネル番号と重み係数を表す補助情報を用いて原信号を復元することができる。

このような符号化手順をとる理由は以下のとおりである。最初の親の決定は性能に大きく影響すると考え

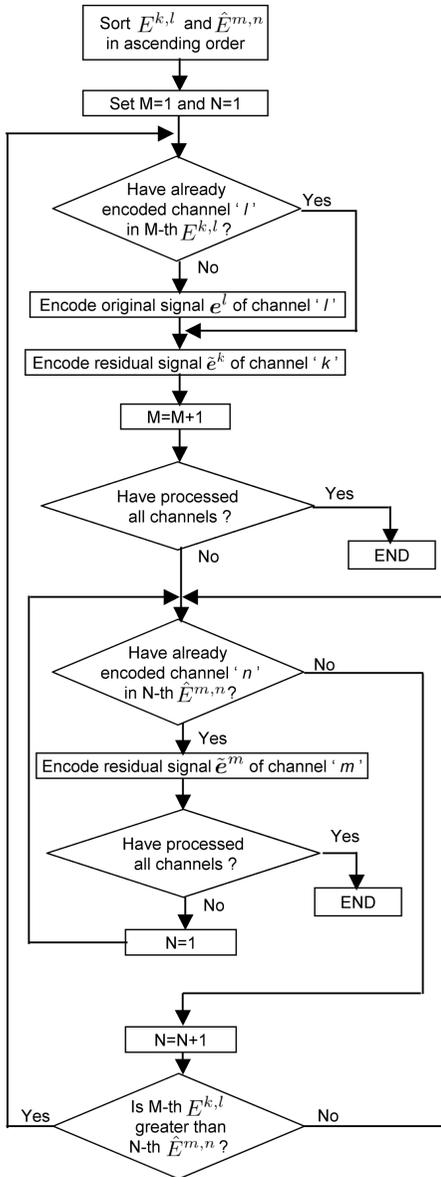


図 5 符号化の場合の、チャンネル間の親子関係の木構造を作成するアルゴリズムのフローチャート

Fig. 5 Flow chart of algorithm for making tree structure of the relationship between channels.

られる（実験的にもその傾向が見られた）ので、最初の親と子の対に関してのみ全探索を行えば、性能を向上させつつ探索を $O(n^2)$ 通りに抑えられる。すなわち、初段とそれ以後の探索との計算量のバランスをとることで、性能向上と計算量抑制を同時に狙うものである。

3.4.3 復号器での処理

復号側ではすでに親と重み係数の情報が分かっている

るので、復号できるものから順次復号していけばよい。具体的には、

- (1) 自分自身が親であるものを復号、
- (2) 親が復号されていれば子を復号、
- (3) (1)へ戻る、

というステップを繰り返すことにより、すべてのチャンネルを復元することができる。復号側では従来と同程度の時間で処理をすることができる。

3.4.4 冗長性の削減

予備実験より、チャンネル間の親子関係は前後のフレームで同じであることが多い傾向が見い出された。そこで、このことを利用して情報量を削減するために、補助情報として、親チャンネルが前のフレームと同じであるかどうかを表すフラグを用意したところ、圧縮率が約 0.02% 向上するという結果が得られた。

また、自分が親、すなわち減算処理されずに送られた場合は、重み係数は必要ないので、さらに、補助情報を削減することができる。

3.5 チャンネルのサブブロック分割

チャンネル数が大きくなるにつれて、親子関係を探索するための計算量は膨大になる。そのため、符号化の計算量を現実的に抑えるために、全チャンネルをサブブロックにグループ分けして、探索空間を小さくする工夫を施した。

チャンネル数が C のとき、探索空間のサイズは $C \times C$ となるが、これを S 個のサブブロックに等分割すると、探索空間のサイズは $(C/S) \times (C/S) \times S$ になり、探索の計算量をおよそ $1/S$ に削減できる。

探索空間が狭くなるので圧縮の性能は低下する可能性があるが、あらかじめセンサ配置が既知であるような場合は、相関の高くなりそうなチャンネルどうしを同じサブブロックにまとめることにより、性能の低下を抑えることができる。

3.6 繰り返し処理による振幅の低減

以上に述べたアルゴリズムを繰り返すことにより、次々と振幅を低減させることができる。しかし、情報量は親のチャンネルと重み係数を表す補助情報の付加とのトレードオフになる。実際の処理は、まず、1度減算処理された重みつき差信号 \tilde{e}^k を用いて、子チャンネル k の 2 度減算処理された重みつき差信号は、新たに探索された最適な親チャンネル m により、

$$\tilde{\tilde{e}}^k = \tilde{e}^k - \gamma^{k,m} \cdot \tilde{e}^m \tag{6}$$

と求められ、振幅を低減させることができる。さらにこの $\tilde{\tilde{e}}^k$ を用いて、新たに探索された最適な親チャンネル n との減算処理により

$$\tilde{\tilde{e}}^k = \tilde{e}^k - \tilde{\gamma}^{k,n} \cdot \tilde{e}^n \quad (7)$$

と3度減算処理された重みつき差信号を求めることで、振幅を低減させることができる。このように何度でも減算処理は行えるが、その分補助情報が付加されるので、減算処理された信号を表す情報に補助情報を付加した符号長と、減算処理されない信号を表す情報の符号長を比較して、効果のある回数だけ繰返せば圧縮性能を向上させることができる。

一方、復号は逆の手順で行う。3度減算処理された重みつき差信号 $\tilde{\tilde{e}}^k$ と補助情報である重み係数と親チャンネル番号から、 $\tilde{e}^k = \tilde{\tilde{e}}^k + \tilde{\gamma}^{k,n} \cdot \tilde{e}^n$ と計算して2度減算された重みつき差信号を求め、同様の手順で \tilde{e}^k を求める。繰返し処理することで、最終的にもとの信号 e^k を復元できる。

4. 性能評価実験

4.1 評価方法

MPEG-4 Audio Lossless Coding の規範モデルに提案法を追加実装し、効果を確かめるために性能評価実験を行った。比較対象としては、一般的に広く使われているファイル圧縮方式である ZIP、元の MPEG-4 Audio Lossless の規範モデルに備わっている2つのモードである mp4als-I と mp4als-J を用いた。なお、mp4als-I はチャンネルごとに独立に線形予測符号化したものであり、mp4als-J は前述のジョイント・ステレオ符号化方式を表す。

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、MPEG-4 Audio Lossless Coding で用いられている評価用データセットのステレオの音楽データ (Audio 2ch)、5チャンネルのサラウンド音楽データ (Audio 5ch)、さらに独自に録音した8チャンネルのマイクロホンアレー受信信号 (Audio 8ch) を用いた。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。各信号のパラメータを表1にまとめた。複数のファイルがあるデータに関しては、平均値を結果とした。元の規範モデルと提案手法はともにサンプリング周波数に応じてフレーム長を変化させており、音響信号では2,048サンプル、MEG 信号では1,024サンプルとした。

4.2 実験結果と考察

4.2.1 圧縮率の比較

可逆圧縮符号化では原信号が完全に復元されるので、歪みの大きさの評価や主観評価実験などは不要で、

表1 評価データのパラメータ

Table 1 Parameters of evaluation materials

data	#channels	word length	sampling rate	#files
Audio 2ch	2	24 bits	48,000 Hz	15
Audio 5ch	5	16 bits	48,000 Hz	95
Audio 8ch	8	16 bits	44,100 Hz	1
MEG-raw	256	16 bits	2,000 Hz	2
MEG-ave	256	16 bits	2,000 Hz	2
MEG-anr	192	16 bits	500 Hz	1

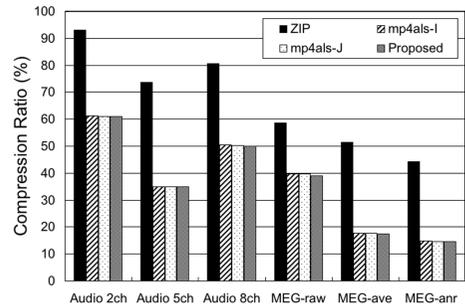


図6 ZIP、独立符号化 (mp4als-I)、ジョイント・ステレオ符号化 (mp4als-J)、提案手法を用いた場合の圧縮率

Fig.6 Compression ratio of ZIP, independent coding (mp4als-I), joint stereo coding (mp4als-J) and the proposed method.

性能の評価としては圧縮率が重要である。ここでは、チャンネル間相関の効果を評価するために、提案手法はチャンネルのサブブロック分割せず、繰返し処理は行わない条件とした。

まず、前述の4手法の圧縮率を図6に示す。ここで圧縮率は、

$$\begin{aligned} \text{圧縮率 (\%)} &= \frac{\text{圧縮後のファイルサイズ}}{\text{元のファイルサイズ}} \times 100 \quad (8) \end{aligned}$$

とした。

MPEG-4 Audio Lossless Coding の規範モデルは、オーディオ信号に対して最適化されており、チャンネル間独立符号化でも ZIP よりはるかに圧縮率が高いことが分かる。さらに提案手法の効果を明示するために、ジョイント・ステレオ符号化と提案手法を比較した圧縮率の相対改善率を図7に示す。相対改善率はチャンネルごとに独立に線形予測符号化した mp4als-I を基準とし、mp4als-I による圧縮後のファイルサイズを I 、比較対象手法による圧縮後のファイルサイズを J として、

$$\text{相対改善率 (\%)} = \frac{I - J}{I} \times 100 \quad (9)$$

により求めた。

いずれの信号についても、提案手法は既存の手法よ

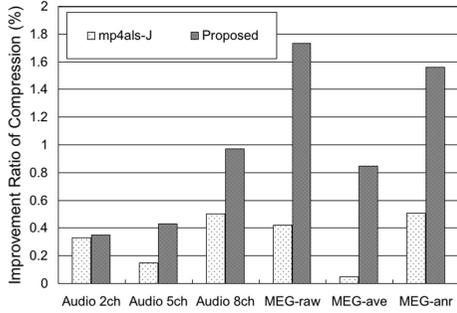


図 7 独立符号化を基準としたジョイント・ステレオ符号化と提案手法の相対改善率の比較

Fig. 7 Improvement ratio of compression comparing joint stereo coding with the proposed method relative to that of independent coding.

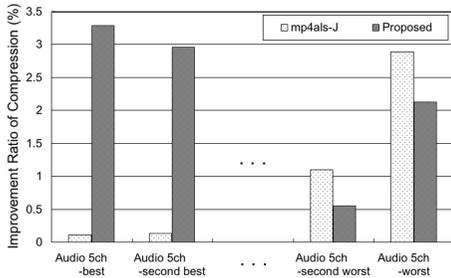


図 8 Audio 5ch のうち各手法の相対改善率の差が大きな場合の例：(提案手法の改善率 - mp4als-J の改善率) を降順に左から並べて両端から 2 つずつを表示

Fig. 8 Extreme examples of improvement ratio in Audio 5ch. (representing both ends of descending order of (Improvement ratio of the proposed method - Improvement ratio of mp4als-J)).

りも圧縮率を向上させることができた。2チャンネルの信号 Audio 2ch では従来のジョイント・ステレオ方式と大差はないが、5チャンネルの信号 Audio 5ch では従来手法と比べて平均で 0.2%の圧縮率の向上が見られた。

この 5チャンネルの信号の 95個のファイルの中には、図 8 に示されるように、圧縮率を最大約 3%向上させることができる場合もあったが、逆に従来手法を用いた改善率が提案手法を上回る場合もあった。図 9 のように、相関のある波形が必ずしもチャンネル番号順の対になっていない場合は、あらゆるチャンネル間の相関を利用して提案手法が効果的に圧縮率を向上させるが、図 10 のように、相関のある波形がチャンネル番号順の対になっていて、さらに対を組めない余ったチャンネルと他のチャンネルとの相関が小さい場合には、補助情報の少ない従来法を用いた方が圧縮性能が高くなる。

また、8チャンネルの信号 Audio 8ch では 0.5%の圧

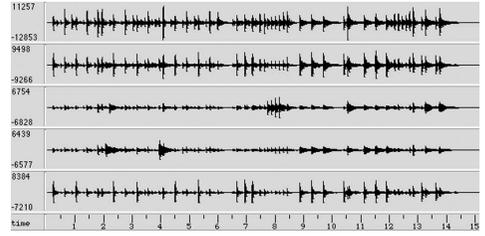


図 9 提案手法が効果的な信号 (Audio 5ch-best) の波形例 縦軸：各チャンネルの整数振幅値、横軸：時間 (秒)

Fig. 9 A set of signals which can be most efficiently encoded by proposed method (Audio 5ch-best) vertical axis: amplitude, horizontal axis: time (second)

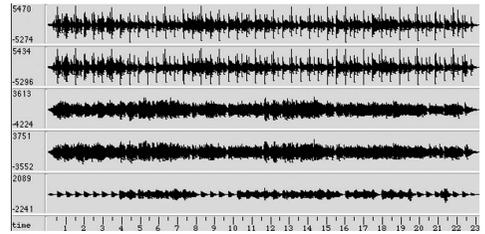


図 10 ジョイント・ステレオ方式が効果的な信号 (Audio 5ch-worst) の波形例 縦軸：各チャンネルの整数振幅値、横軸：時間 (秒)

Fig. 10 A set of signals which can be most efficiently encoded by mp4als-J (Audio 5ch-worst) vertical axis: amplitude, horizontal axis: time (second)

縮率向上が見られた。そして、200チャンネル程度の信号 MEG-raw, MEG-ave, MEG-anr では、従来手法と比べて 1%程度の性能向上が確認できた。

このように、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積極的に利用した効果が現れていると考えられる。

4.2.2 チャンネルのサブブロック分割の効果

チャンネル数が多いほど親子関係の探索に時間がかかることから、サブブロックに分割した場合の代表的な結果として、MEG-raw についての圧縮率と圧縮符号化の処理時間のトレードオフの関係を図 11 に示す。これはチャンネル数の多い MEG 信号のみを対象とした。他の多チャンネル信号である MEG-ave や MEG-anr でも同様の結果が得られたので、ここでは省略する。図中の “32 Subblocks” はチャンネル全体を 32 のサブブロックに分けたときの相対改善率とその計算に要した相対圧縮処理時間を示している。相対圧縮処理時間もチャンネルごとに独立に線形予測符号化をした mp4als-I を基準とし、mp4als-I を用いたときの圧縮処理時間を τ_I 、比較対象手法を用いたときの圧縮処理時間を τ_J として、

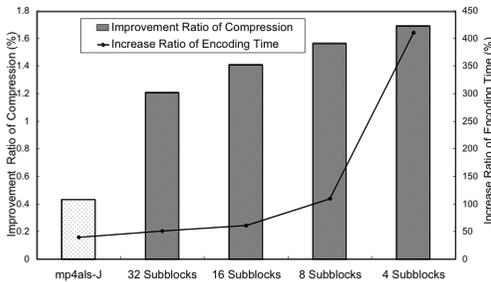


図 11 サブブロックに分割したときの圧縮率の相対改善率と相対圧縮処理時間の比較 (MEG-raw)

Fig. 11 Improvement ratio of compression and increase ratio of encoding time for MEG-raw in case of sub-block search.

$$\text{相対圧縮処理時間 (\%)} = \frac{T_J - T_I}{T_I} \times 100 \quad (10)$$

により求めた。なおここでは、ANSI C 準拠のライブラリ関数 clock を用いて計測した、ファイル入出力を含む圧縮処理に要する時間を圧縮処理時間とした。実験機の CPU は AMD の Athlon XP 1800+ (クロック約 1.5 GHz), RAM は 768 M バイト, OS は Windows 2000 であった。この条件下で、約 20 M バイトのファイルである MEG-raw における圧縮処理時間 T_I は 5.62 秒であった。

このように、圧縮処理時間と圧縮改善率とはトレードオフの関係にあるので、圧縮処理時間と圧縮率のバランスを考えて適当なものを選べば、圧縮処理時間の増大を抑えつつ既存の手法より圧縮性能を向上させることができる。たとえば図中の結果から、全チャンネルを 16 のサブブロックに分割することを合理的な選択とすることができる。また、センサの配置があらかじめ分かていなくても、初めの数フレームの親子関係の情報を利用してクラスタリングを行い、分割先のグループを決めておくことにより、圧縮性能を大きく変えずに圧縮処理時間を短縮することができる可能性がある。

4.2.3 繰返し処理による振幅の低減の効果

提案アルゴリズムを繰返すことによる圧縮率の向上を図 12 に示す。相対改善率は式 (7) による。ここでも代表的な MEG-raw についての結果を示したが、他の MEG 信号でも同様の結果が得られた。図中の iteration# は提案アルゴリズムを # 回繰返した結果を示している。繰返しの手法では、チャンネル数が多いと計算量が多いので、計算量と圧縮率のバランスが良い条件として、チャンネル全体を 16 のサブブロックに分割する方式を用いた。

サブブロックに分割すれば親子関係の探索の計算量

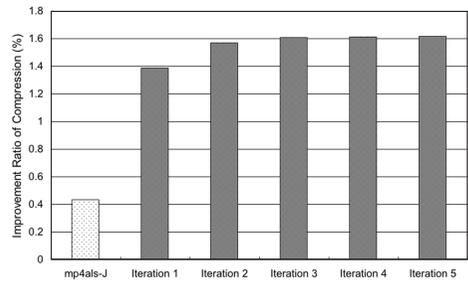


図 12 提案アルゴリズムを繰返し用いた場合の相対改善率
Fig. 12 Improvement ratio of compression by means of iterative application of the proposed method.

は抑えられるので、2 回繰返す処理は実際的であった。それ以上繰返すと振幅低減の効果よりも付加する補助情報の方が大きくなり、全体の圧縮に寄与できなかった。このような繰返しによる振幅低減手法も、圧縮性能と計算時間のバランスを考え、性能を選択することができる。

4.3 評価実験のまとめ

以上の結果から、提案手法は圧縮率の向上に効果があることを確かめることができた。

今回用いたアルゴリズムよりも、さらに最適解に近づくようにチャンネル関係の木構造生成方法を改善できれば、圧縮性能をさらに向上させられる可能性がある。また、提案手法は、圧縮性能と計算量を柔軟に選択できる特徴がある。サブブロックの分割と繰返しによる振幅の低減効果を組み合わせることや、クラスタリングや繰返し回数の自動決定などの工夫を施すことで、さらに性能向上が期待できる。

5. おわりに

多チャンネル時系列信号の線形予測分析後の残差信号に対して、チャンネル間相関分析に基づく重みつき差分を適応的に適用することで残差信号のエネルギーをさらに小さくし圧縮率を向上させる新しい圧縮符号化法を提案した。本手法はチャンネルごとの独立符号化より圧縮率が低下しないことを保証できるという特徴があり、用途によって符号器の演算量と圧縮性能のトレードオフを選択できる。

各種条件での音響信号、生体信号で従来の 2 チャンネルのペアごとの圧縮法およびチャンネルごとの独立の圧縮法と比較して圧縮率が改善されることを確認した。2 チャンネルや 8 チャンネルのオーディオ信号での改善率は小さく、5 チャンネルの場合の平均改善率は 0.2% であったが、信号によっては約 3% の圧縮率の改善が得られた。256 チャンネルの MEG の場合には圧縮率の改

善率は1%から2%であった。

本手法は ISO/IEC MPEG オーディオ標準化で策定中のロスレス・オーディオ符号化のコア実験に提案され、2004年12月現在、第三者による相互検証を経て暫定参照モデルに組み入れられている。本手法が MPEG 国際標準に採用されれば、音響信号だけでなく幅広い時系列信号の可逆圧縮ツールとして貢献することが期待される。

また、本研究は東京大学大学院情報理工学系研究科と NTT 研究所との戦略型 IT 連携講座に基づく共同研究の成果であり、探索アルゴリズムの考案、生体信号の提供と考察は主に大学側、符号化の枠組みの設計、標準化への提案は NTT 側からの寄与が大きく、産学連携研究の成功例である。大学のみでは難しい国際標準化への提案を行い、研究成果を実用的な形で公開できた。

謝辞 MEG のデータを提供していただくとともにセンサの位置や信号の分類に関してご指導していただいた、東京大学大学院新領域創成科学研究科武田常広教授、大脇崇史助手に感謝いたします。

参考文献

- 1) 藤原洋ほか：画像&音声圧縮技術のすべて，CQ出版株式会社 (2000).
- 2) Hans, M. and Schafer, R.W.: Lossless Compression of Digital Audio, *IEEE Signal Processing Magazine*, Vol.18, No.4, pp.21–32 (2001).
- 3) Konstantinides, K., et al.: An Introduction to Super Audio CD and DVD-Audio, *IEEE Signal Processing Magazine*, Vol.20, No.4, pp.71–82 (2003).
- 4) 守谷健弘：音声符号化，電子情報通信学会出版 (1998).
- 5) Itakura, F. and Saito, S.: Digital filtering techniques for speech analysis and synthesis, *Proc. Int. Acoust. Cong. 25C-1* (1971).
- 6) Liebchen, T. and Reznik, Y.A.: MPEG-4 ALS: An Emerging Standard for Lossless Audio Coding, *Proc. IEEE Data Compression Conference 2004*, pp.439–448 (2004).
- 7) Liebchen, T.: MPEG-4 lossless coding for high-definition audio, *AES 115th Convention paper 5872* (2003).
- 8) Robinson, A.J.: SHORTEN: Simple Lossless and Near-Lossless Waveform Compression, Technical report CUED/F-INFENG/TR.156, Cambridge University Engineering Department (1994).
- 9) Golomb, S.W.: Run-Length Encodings, *IEEE Trans. Information Theory*, Vol.12, No.7, pp.399–401 (1966).
- 10) Rice, R.F.: Some practical universal noiseless coding techniques — Parts I-III, Jet Propulsion Laboratory Technical Report JPL-79-22, JPL-83-17 and JPL-91-3 (1979, 1983, 1991).
- 11) Salomon, D.: *Data Compression*, Springer (1998).
- 12) Gilbert, E.N. and Moore, E.F.: Variable-Length Binary Encodings, *Bell System Technical Journal*, Vol.38, pp.933–968 (1959).
- 13) Szpankowski, W.: Asymptotic Average Redundancy of Huffman (and Other) Block Codes, *IEEE Trans. Information Theory*, Vol.46, No.7, pp.2434–2443 (2000).

(平成 16 年 9 月 1 日受付)

(平成 17 年 2 月 1 日採録)



鎌本 優 (正会員)

1978年生。2003年慶應義塾大学理工学部物理情報工学科卒業。同年東京大学大学院情報理工学系研究科システム情報学専攻入学。2004年からNTTコミュニケーション科学基礎研究所実習生。2005年4月NTT入社。多チャンネル信号処理の研究に従事。日本音響学会、情報理論とその応用学会各会員。



守谷 健弘 (正会員)

1955年生。1980年東京大学大学院工学系研究科計数工学専攻修士課程修了。同年電電公社入社。音響音響信号の符号化の研究開発、標準化に従事。1989年工学博士，同年AT&Tベル研究所客員研究員。現在，NTTコミュニケーション科学基礎研究所人間情報研究部長，東京大学大学院情報理工学系研究科戦略型IT連携講座客員教授，NTT R&D フェロー。電子情報通信学会，日本音響学会，IEEE 各会員。IEEE フェロー。



西本 卓也 (正会員)

1970年生。1993年早稲田大学理工学部卒業。1995年同大学大学院理工学研究科修士課程修了。1996年京都工芸繊維大学工学部助手。2002年東京大学大学院情報理工学系研究科助手。音声対話システム、福祉情報工学、擬人化音声対話エージェントの研究に従事。人工知能学会 1998年度研究会奨励賞。日本音響学会第20回粟谷潔学術奨励賞。日本音響学会、電子情報通信学会、人工知能学会、ヒューマンインタフェース学会各会員。



嵯峨山茂樹 (正会員)

1948年生。1974年東京大学大学院工学系研究科計数工学専攻修士課程修了。同年電電公社入社。音声情報処理の研究に従事。1990～1993年ATR自動翻訳電話研究所音声情報処理研究室長。1998年北陸先端科学技術大学院大学教授。2001年東京大学大学院教授。現在、同大学院情報理工学系研究科教授。音声・音楽・文字・エージェントの信号処理・情報処理の教育研究に従事。博士(工学)。発明協会発明賞、科学技術庁長官賞等を受賞。日本音響学会、電子情報通信学会、SICE、IEEE、ISCA、AVIRG各会員。