

ベクトル類似度を用いた順序制約付きステレオ符号化

Sequence Joint Stereo Coding based on Vector Angle

窪 太巻† 栗原 良尚‡ 濱崎 貴紀‡ 本間 永愛‡ 六井 淳†
Taichi Kubo Yoshinao Kurihara Takanori Hamasaki Hisae Honma Jyun Rokui

1. はじめに

現在、インターネット音楽配信やデジタル放送等で最も普及している圧縮手法は、高圧縮率を達成した不可逆圧縮符号化である。不可逆圧縮符号化は圧縮率の面で優れているが、原信号の完全復元が不可能である。音響信号や生体信号などの分野では、原音・原信号の情報を保持する必要がある。多チャンネル信号の可逆圧縮符号化は、通信インフラの整備と利用環境の向上により、今後幅広い応用が期待される。

本研究で取り上げるステレオ符号化(Joint Stereo Coding: 以下JSC)は、現在MPEG-4規格として国際基準が策定されつつある符号化手法である。本来、2ch(ステレオ)データ用に考案された手法であるが、2の階乗個のデータに対しても有用であるため、多チャンネル信号にも用いられている。多チャンネル信号はステレオの対に対して符号化するため、対の組み合わせによって圧縮率が向上する可能性が高い。現在策定中の基準では、線形予測とエントロピー符号化を組み合わせた手法が用いられている[1]。また、チャンネル間相関を用いることで、圧縮性能を向上する手法も提案されている[2]。

本研究では、チャンネル間のデータ類似度による対の形成と、チャンネル内データのフレーム分割類似度を用いた対の形成について検討を行った。

2. 可逆圧縮符号化の基本原則

2.1 MPEG-4 規格内可逆圧縮符号化

現在規格策定中の MPEG-4 圧縮規格に音声を含む多チャンネル可逆圧縮符号化手法とし、Audio Lossless Coding (以下: ALS)[3],[4]が用いられている。ALS では、符号化の中心技術として線形予測及びエントロピー符号化を用いている。時間軸方向にフレーム分割を適用し、これらの処理を施すことにより、時間軸領域での圧縮符号化を可能としている。

2.2 線形予測

線形予測法は、過去の入力信号により、現在及び未来の値を線形予測して、出力する手法である。この手法により得られた予測値 \hat{x}_t と、現在の実測値 x_t との差異を予測誤差 e_t として求め保持する。ここで、予測係数は a_k ($k=1,2,\dots,p$) で表される。予測誤差・予測係数及び過去のデータ p サンプルにより、現在の原信号を復元可能とする。音声信号などの、自己回帰構造を持つ既知の信号に対して効果が期待できる。JSC においては、データ伝送時に、原信号が復元可能かつ伝送効率の向上が期待で

† 島根大学総合理工学研究科

‡ 島根大学総合理工学部

きる手法として用いられている。

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

$$x_t = \hat{x}_t + e_t \quad \dots(2)$$

2.3 エントロピー符号化

代表的な情報の符号化手法として、エントロピー符号化が挙げられる。情報の出現頻度により符号語を配することで、効率的な符号化を図る手法である。符号アルファベットの要素数を n 、任意のシンボル s の出現確率を ps とすると、 $-\log_n ps$ の長さの符号を割り当てた場合が最短の符号長となる。

線形予測により得られた予測誤差 e_t は、予測利得の特性から、予測精度の向上により、絶対値は0付近に収束する傾向がある。このように、分布に明らかな偏りが期待できる情報を符号化する際に、エントロピー符号化は有用である。

2.4 ステレオ符号化

Lch と Rch の信号 x_t^L, x_t^R の差異 $d_t = x_t^L - x_t^R$ をとり、原信号を含めそれぞれを符号化し、符号長の短い2つを対として用いる。どの組を採用した場合でも、原信号が復元可能(図 1)であり、情報損失を伴わないという特徴をもつ。また、処理手順が簡素化されているため、符号化処理に要する時間及び資源の低減も図ることが可能である。

多チャンネルに用いる際に、チャンネル対の形成法が問題となる。現在の規格基準では、チャンネル間相関とは無関係に対を作り、符号化が行われている。本研究では、この対形成法について、ベクトル類似度の利用を提案する。

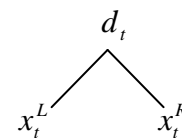


図 1 :Joint Stereo

3. 類似度を用いたチャンネル対選択

3.1 類似度対の形成

本研究では、原信号の符号化にあたり、類似度での対形成を行う。類似度対での減算により、同符号同士のと対の発生頻度が向上し、その振幅低減が期待できる。振幅低減により、符号化の効率が向上し、圧縮率向上に繋がると考えられる。類似度対生成時の、同符号対の発生イメージを図2に示す。図2は、幾何学的概念として各チャンネル内データは2値とし、チャンネル数4chの場合を表している。チャンネルを

ベクトルと見なした場合、対形成を考慮しないJSCでは、差分は大きくなり、JSC本来の圧縮効果を得られない場合がある。類似度を考慮した場合、類似度の小さいチャンネルを選択することで圧縮効果の高い組み合わせを選択できる。

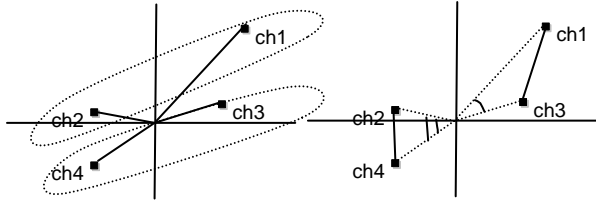


図2: JointStereo方式(左図)と類似度対方式(右図)の幾何学的概念(2値の場合を图示)

本研究では、チャンネルデータをベクトルとみなし、ベクトル類似度[5]を算出する。ベクトル x 及び y のなす角 θ は

$$\theta = \text{Arccos} \frac{\langle x, y \rangle}{\|x\| \|y\|} \quad \dots(3)$$

で与えられる。ここで、 $\langle x, y \rangle$ は x と y の内積を表し、 $\|x\|$ は x のノルムである。この θ が 0 に近い対を選ぶことにより、振幅縮小効果は大きくなる。このため、対の選定は信号毎に適時類似度を取り求める必要がある。また、類似度対により得られた差信号が原信号に対し、情報削減効果が得られない場合は、原信号をそのまま符号化する。

3.2 類似度対による木構造の形成

従来手法の JSC では、チャンネル番号順に対を形成し、ていた。対ごとに2つの原信号及び差信号を符号化し、その符号長の短いもの2つを保持する。この符号長が図2左の実線部と対応する。JSC では、対の形成にデータ相関が考慮されておらず、符号長はチャンネルの順番に依存し、符号長の最適化は行われていない。

類似度対の形成により、振幅の低減が図れるため、JSCは多チャンネルに対しても有効な手法となる。しかし、チャンネル対の形成に際し、チャンネル数を n とすると nC_2 通りの組み合わせが存在するため、従来法同様にその探索方法及び探索時間問題を抱えている。本研究では、類似度対の探索時間問題解決のため、8個のチャンネルごとにチャンネル間相関を考慮した対形成法を提案する。類似度対による符号長は図2右の実線部に相当する。8ch内で類似度による対の組み合わせ探索を行うことで、幾何学的に類似したチャンネルを対にすることができ、振幅の低減が図れる。また、8chを1グループとして処理することにより、高速な対形成処理が可能である。

以下に処理手順及び、類似度木を示す。

1. データ規則を考慮した小グループ分け
2. 小グループ内の8chについて類似度計算
3. 類似度が0に近いものから対とする
4. 各対について差信号を算出
5. 各対について原信号2・差信号をzip圧縮
6. 各対についてファイルサイズの小さい2つを保持

ここで、データ規則とは、チャンネル位置などの情報である。本研究の検証実験では脳波信号を用いているが、

頭部物理部位におけるチャンネル位置がこれに当たる。

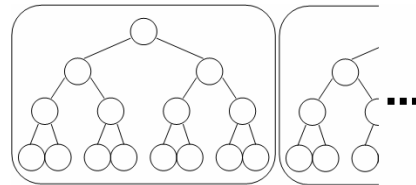


図3: 類似度木

4 実験

4.1 類似度圧縮効果

JSC に提案法を実装し、評価実験を行った。評価データとして、サンプリングレート 250Hz、インピーダンス 40K Ω 、データ採取時間約 15 分の 128ch の脳磁計(MEG)信号を用いた。

本研究における提案法との整合性を考慮し、JSC においても、 $8 \cdot 16 \cdot 32 \cdot 64 \cdot 128\text{ch}$ ごとに適用した。表1にその結果と図4にチャンネル毎の平均処理時間を示す。

評価データ	MEG-raw	JSC		提案手法		差分率
		サイズ(B)	圧縮率	サイズ(B)	圧縮率	
8	40,288,941	13,846,097	65.63%	13,659,196	66.10%	99.30%
16	80,843,752	27,825,920	65.58%	27,562,127	65.91%	99.50%
32	161,568,822	55,539,135	65.63%	55,050,218	65.93%	99.54%
64	319,624,106	110,001,064	65.58%	109,264,471	65.81%	99.65%
128	636,111,597	218,938,865	65.58%	218,653,721	65.63%	99.93%

表1: データ圧縮率の比較

JSC 及び提案手法のサイズ欄は圧縮後のファイルサイズを、差分量の欄は圧縮後の JSC と提案法による差異を、差分率はその比率を示したものである。

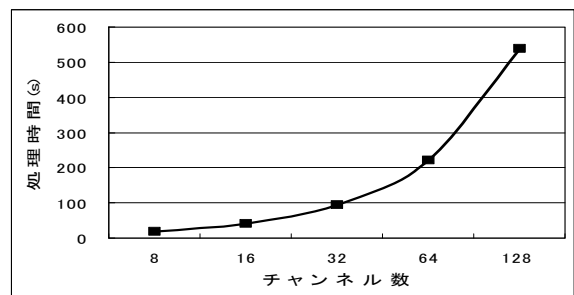


図4: チャンネル数-処理時間相関

また、表2に、8chについて対を総当りで計算し算出された最適対と提案手法による圧縮率の比較結果を示す。

表1より、8chでの木構造グループ分け時に最も圧縮効率が高いといえる。ベクトル類似度が0に近いものから選択し類似度木を生成するため、チャンネル数が少ない方が、より全体を最適に保つことが出来たためと考えられる。

	サイズ(B)	圧縮率
圧縮率	38,167,504	—
ZIP	13,058,973	65.79%
総当り	12,486,728	67.28%
提案法	12,502,682	67.24%

表2：総当りと提案法の比較

また、図4より、8chと128chでは処理時間に約29.3倍の開きがあり、1ch当たりの処理時間としても、約1.8倍の開きがある。これは、対の組み合わせ数が8chでは28通り、128chでは8128通り存在し、チャンネル数増加に伴う組み合わせ数増加により、探索に指数関数的に時間を要することを示している。

表2より、8chでの総当りによる最適対探索と類似度対の圧縮率は近似しており、ベクトル類似度による対形成法が十分に機能しているといえる。

以上より、8chでの木構造グループ分け時に、圧縮率・処理時間ともに最も効率が向上するといえる。

4.2 フレーム分割効果

新たな適用方法として、1chデータをフレーム分割し、フレーム単位に対選択を行う。JSC及び提案法を適用し、その結果を表3に示す。

フレーム幅	有効ch数	非圧縮	JSC		提案法	
		サイズ(B)	サイズ(B)	圧縮率	サイズ(B)	圧縮率
16,384	9	43,498,450	15,252,365	64.95%	15,169,397	65.13%
32,768	2	9,358,380	3,288,288	64.86%	3,248,343	65.29%
65,536	1	4,236,650	1,477,140	65.13%	1,422,422	66.43%

表3：JSC・提案法のフレーム単位適用

有効チャンネルは、提案法による類似度対がJSCと比較し効果を上げたチャンネル数である。

表3より、どのフレーム長についても提案方がJSCと比べ良い結果を示している。フレーム長が長くなるにつれ、類似度対の効果がJSCと同等程度になっている。フレーム長が短い場合、同位相対の発生率が増加し、信号振幅低減による符号長短縮効果が向上したためと考えられる。より短いフレーム単位で分割を行うことで、類似度対による圧縮率向上に繋がるものと考えられる。

5 まとめ

本研究ではベクトル類似度をチャンネル対選択の基準とした新たなステレオ符号化(以下：JSC)の枠組みを提案した。ベクトル類似度による対生成と、8chごとの木構造グループ分けによる類似度木の生成には十分な効果があるといえる。また、本研究では従来法より問題であった多チャンネル適用の際の対探索に伴う計算量問題に対し8chごとの木構造グループ毎に処理を行うことで、JSCの処理が簡便で高速であるという利点を損なわず、実装することができた。

また、1ch内のデータをフレーム分割し、各フレームに対しJSCを用いた場合、フレーム長が長い場合により圧縮効果があることが確かめられた。しかし、類似度の効果に

については、フレーム長を短く取った場合に有効であることが確認された。これにより、データ長に応じたフレーム分割を行ったチャンネルデータに対し、フレーム単位で8chの木構造グループを生成することで、さらなる圧縮率向上が望めると考えられる。

今後、グループ分けの明確な指標の策定とグループ毎に生成された類似度木同士の相関にも着目し研究を進める予定である。

6 参考文献

- [1] Tilman Liebchen, Yuriy A. Reznik, "MPEG-2 ALS: an Emerging Standard for Lossless Audio Coding", Proc. IEEE Data Compression 2004, pp.439-448, 2004.
- [2] 鎌本 優, 守谷 健弘, 西本 卓也, 嵯峨山 茂樹, "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", 情報処理学会論文誌, vol.46, No.5, pp.1118-1128, May, 2005.
- [3] Mat Hans, et al.: "Lossless Compression of Digital Audio," IEE signal Processing Magazine, pp.21-32, July 2001.
- [4] "Audio Lossless Coding (ALS), new audio profiles and BSAC extensions". ISO/IEC, 14496-3:2005/Amd.2:2006.
- [5] 友近 晋 「ベクトル解析」 共立出版株式会社 February, 1950
- [6] Itakura F and Saito S: "Analysis synthesis telephony based on the maximum likelihood method", Reports of the 6th Int. Cong. Acoust. C-5-5.1968.