

論文の内容の要旨

論文題目 : **Efficient lossless coding of multichannel signal
based on time-space linear predictive model**
(時空間線形予測モデルに基づく多チャンネル信号の
高効率可逆圧縮符号化の研究)

氏 名 : 鎌 本 優

音声音響符号化は現代のデジタル通信において大きな貢献をしてきている。例えば携帯電話、IP 電話 (インターネット・プロトコル電話)、デジタル放送、携帯音楽プレーヤ、光ディスクなどは圧縮符号化技術なしには実現できなかったであろう。現在広く使われている歪のある音声音響符号化方法は、40 年以上のさまざまな分野の研究成果に基づき、元のビットレートの約 10 分の 1 の大きさに圧縮できている。音声音響信号処理のみならず聴覚や脳科学を知識も用いて、人間が気付きにくい歪を許すことにより効率よく信号を伝送している。例えば、携帯電話で使われている AMR (Adaptive Multi-Rate) 符号化方式は音声信号に特化し、低遅延・低ビットレートでなおかつ通信回線の誤り耐性に強い方式である。また、携帯オーディオプレーヤや地上デジタル放送などでも使われている AAC (Advanced Audio Coding) 符号化方式は、一方通行の通信ということもあり、遅延等を許す代わりに低ビットレートで高品質の音楽を伝送することができる方式である。一般的に圧縮符号化には強みと弱みのトレードオフがあり、何かを犠牲にすることにより性能を維持している。現在の電話や放送で用いられている符号化方式は復号しても元の信号には復元することができない歪のある方法となっている。

一方、身の回りで使われている歪のない圧縮符号化方式として ZIP などが挙げられる。この符号化方式は文書ファイルやプログラムファイルなど 1 ビットの誤りもあってはいけないようなデータを圧縮するのに適しており、実際、メールに添付してファイルを送るときなどに頻繁に用いられている。ただしこの方法にも弱点はあり、いくつかのファイル (少なくとも 1 つ) をひとまとめにして圧縮するのが想定されているため、実時間で圧縮しながら伝送するような、電話や放送の用途には向かない。またこれらの圧縮は入力を限定しないことが設計指針となっており、入力を音声や音響信号に限定した場合の圧縮率は非常に低く、音を伝送するという観点から言えば実用的な利点は小さい。

近年、高速にネットワークの品質は改善し記憶容量も大きくなってきている。そのような背景のもとで、音声や音楽向けの歪みの無い符号化 (可逆圧縮符号化/ロスレス符号化) が普及しつつある。音声音響信号の可逆圧縮符号化はビットストリームから元の音声音響

波形を完全に復元することができる。また、歪のある符号化と比べて大きな伝送容量を必要とするが、音質の劣化を気にせずにするため、安心して素材音源や文化的音源を実時間で伝送することができる。

圧縮符号化方式は、International Telecommunication Union (ITU) や International Organization for standardization (ISO) および International Electrotechnical Commission (IEC) などで標準化されることが多い。通信や放送で使うためには独自規格ではなくお互いに同じ規格を用いることにより相互接続性が確保でき、機器の普及にもつながる。また世界的な協力によって標準が策定されるため、一部の国や企業に利益が集中することも少なく、長期的なメンテナンスも期待できる。実際にアナログレコード（既に劣化が始まっているものや、そもそもプレーヤのないものもある。）の長期保存のための国際標準も策定されており、安心して長期保存が可能となる可逆圧縮符号化方式の国際標準は文化的にも重要な技術である。

本研究では歪を許す音声や音響信号の圧縮符号化で長く研究されてきた処理や見識を十分利用して、歪を許さない符号化を実現する。また入力信号を音声と音響信号に限定せず、生体信号や地震波などの多チャンネル時系列信号を皇室よく圧縮することを目指す。

本研究の目標は、実用的な観点から十分な処理速度を持ちつつ、最小符号長規準のような理論的背景に基づいた高効率の可逆圧縮符号化技術を開発することである。可逆圧縮符号化は、モデルパラメータに与える符号長とモデルで表現できなかった残差に与える符号長のトレードオフの関係をもつ。また、実用の観点で言えば、演算量と圧縮率のトレードオフの関係をもつ。これら両方のトレードオフのなかから合理的な点を見つけることができれば、高効率な圧縮性能を得ることができる。ネットワークの伝送容量や蓄積媒体の記憶容量は確かに大きくなってきているが、わずかな演算量で品質をまったく変えずに伝送量を大きく削減できる（たとえば半減させる）のであれば、光ネットワークやルータなどのネットワーク機器の負荷を削減することができ、電力量削減にもつながる。実際、本論文に書かれている技術の一部は圧縮率と演算量のバランスに優れており、国際標準技術である、ISO/IEC MPEG-4 Audio lossless coding (ALS) や Recommendation ITU-T G.711.0 (Lossless compression of G.711 pulse code modulation) の必須の構成要素技術に採用されている。

論文の第1章では研究背景や動機について述べる。続く第2章では、2つの国際標準技術(MPEG-4 ALS と G.711.0) を用いて最先端の音声音響信号の可逆圧縮符号化の原理を説明する。

第3章では、時空間線形予測モデルを提案する。時間方向の線形予測残差信号に残るチャンネル間の冗長性を利用してチャンネル方向の線形予測を適用する。残差信号に対して重み付き減算処理を行うことで、振幅値を減少させエントロピー符号化の効率を向上させる。空間方向の予測によりチャンネル数の多い信号、例えばサラウンド音響信号のみならず脳磁計信号や心電図信号のような生体信号にも適用でき、実験により圧縮率が向上することを

示す。

第 4 章では、空間線形予測のためのチャンネル間相関の親子関係の組み合わせを高速に探索する効率的なアルゴリズムを提案する。時間方向の相関はサンプル列の順に並んでいるが、空間方向の冗長性は入力チャンネル順に並んでいるとは限らない。本稿での探索方法はグラフ理論に基づき、従来法と比べて演算量を大きく削減できており、最適解とほぼ変わらない圧縮率を得ることができる。

第 5 章では、空間方向の線形予測の効果をさらに高める時空間線形予測方法を提案する。従来法では時間方向の予測残差信号のみが最小となるように線形予測係数を求めていたが、提案法では直接的に時空間予測後の残差信号を最小化するように線形予測係数を求める。演算量削減の工夫を施すことにより、圧縮時間が同じであれば、従来法よりも圧縮率を向上させることができ、圧縮率を同程度にする設定では圧縮時間を大幅に削減できる。

第 6 章では、予測残差信号の ℓ_1 ノルムやそもそも Golomb-Rice 符号の符号長を最小化する規準で線形予測係数を求める方法を提案する。従来法での ℓ_2 ノルム（二乗誤差）最小化規準は厳密に言えば Golomb-Rice 符号の最小化規準ではない。補助関数法を用いることにより目的関数を効率よく減少させることができ、それを満たす線形予測係数を得ることができる。実時間程度の演算量を用いれば、圧縮率をさらに向上させることができる。

第 7 章では、高効率な PARCOR 係数の量子化方法と低演算量の予測次数推定方法を紹介する。提案する PARCOR 量子化法は予測残差のエントロピーが最小となる規準で設計されている。また、予測次数推定法は最小符号長規準を用いて PARCOR 係数から最適次数を求める。これら 2 つの提案法を組み合わせることにより、圧縮率と演算量のトレードオフの観点から妥当な性能に到達することができる。

最後に第 8 章で論文を締めくくる。本稿で提案されている方法の一部は国際標準規格の必須の構成要素技術として採用されており、今後世の中で使われることによって、品質が高く使いやすい通信システムの実現に微力ながら貢献できるものとする。