

音響信号のロスレス符号化

音の情報を歪なく圧縮する可逆符号化技術

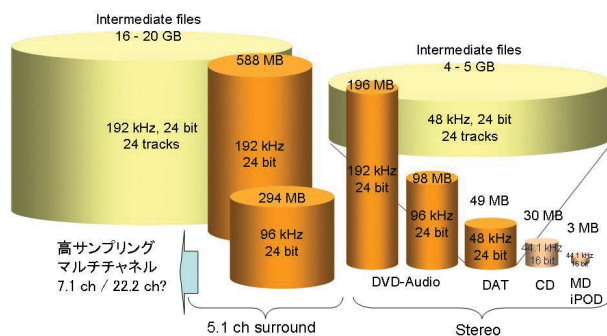
研究の背景

近年、ネットワークやデジタル機器のブロードバンド化が進展する一方、オーディオ信号の高品質化(高サンプリングレート、高振幅分解能、多チャンネル化)の要望が高まり、蓄積や配信のための情報量も飛躍的に増大しつつあります。本研究では、音の情報を歪なく圧縮するために、整数信号、多チャンネル信号、浮動小数点信号等に適用可能な可逆符号化アルゴリズムを検討しています。また、これと並行して相互接続性、長期保守性、特許権利者の明確化のために、MPEGオーディオ委員会が国際標準規格の策定を主導してきました。この結果、NTTが提案した技術は他機関の技術と統合され、国際標準MPEG-4 ALS (Audio Lossless Coding)として2006年3月に出版されました。

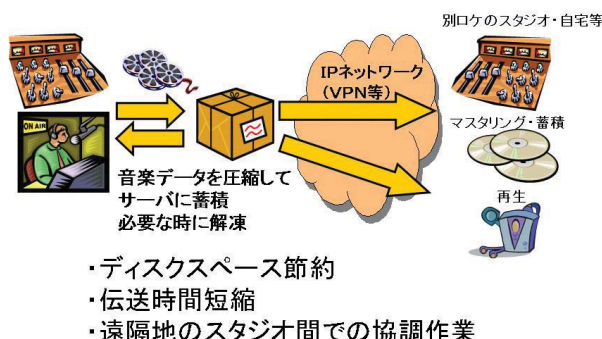
将来どのように使われるのか(応用例)

ロスレス符号化技術によって、さまざまな形式の高品質のオーディオ信号を歪なく15~70%のサイズに圧縮することができます。これによって、遠隔地で音楽データのやりとり、マスターデータの長期的な保存、高品質な音楽のネットワークからの購入、家庭用のPCや家電製品での音楽の蓄積などを効率的に行うことが出来るようになります。また音楽以外の時系列信号にも広く適用可能であるため、医療用波形信号やセンサーアレイの信号等を歪なく圧縮して、便利に取り扱うことが出来るようになります。国際標準規格による互換性を維持しつつ、多様な用途で、世界中で、世代を越えて便利に使ってもらうことを夢見ています。

3分の曲に必要なデータサイズ



ロスレス符号化の応用例



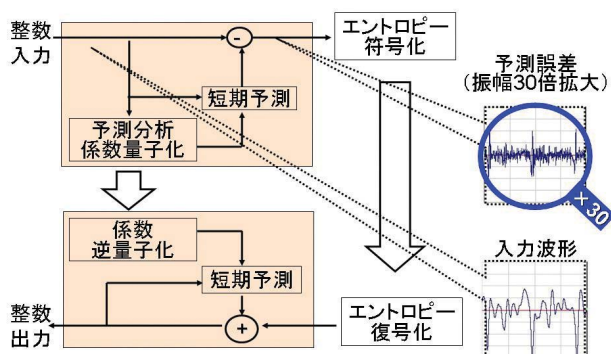
音響信号のロスレス符号化

私たちは、オーディオ信号を歪なく圧縮する技術の研究に取り組んでいます

国際標準MPEG-4 Audio Lossless Coding(ALS)

ALSは特にオーディオ信号の波形データを完全に再構成することを保証しつつ、データ量を圧縮できる国際標準規格の符号化技術です。この技術は過去の複数のサンプルの値に予測係数を乗じて足し合わせることで現在のサンプルを予測する線形予測技術に基づいています。図の上段の予測誤差を忠実に再現すれば、下段の逆の操作で元の波形を再構成することができます。予測係数は予測誤差の振幅が最小になるように計算し、得られた予測誤差の各サンプルをエントロピー符号化することによって符号量を削減します。さらに長期予測、多チャンネル信号符号化、浮動小数点信号の符号化技術等、さまざまな新しい技術が用いられています。

ALSの符号化・復号化の基本構成



多チャンネル信号と浮動小数点信号の圧縮

現在多くの皆さんが耳にする音楽はステレオ(2チャンネル)信号ですが、普及し始めているサラウンド(5.1チャンネル)をはじめとして、音響信号のチャンネル数を増加させることで臨場感を高めることができます。またロスレス圧縮が適用可能な生体信号(心電図や脳磁計信号)やセンサーアレイ信号は500チャンネルを超えることもあります。このような信号のチャンネル間での信号の類似性に着目し、チャンネル間の適応的予測手法やチャンネル間の適応的対応付け手法を考案しました。これにより効率的な処理で圧縮性能の向上を実現しました。

浮動小数点信号は振幅の精度を保てるので、プロ用の音楽録音、編集、保存、配信等に便利ですが、これまで効率よく圧縮することはできませんでした。浮動小数点信号を整数信号にマッピングして効率よく圧縮する手法を考案し、さらにその際に、事前に近似共通因数符号化を適用する方式、誤差信号をマスクLZ符号化を用いて圧縮する方式を考案しました。これにより特に整数信号に浮動小数点表現された共通乗数が掛け合わされて生成された信号の圧縮率が劇的に改善されることを示しました。

便利に使っていただくために

高い圧縮率によってデータ量を削減できても、圧縮や解凍の処理に時間がかかったのでは便利に使えません。そこでMPEG標準に準拠したままアルゴリズムやソフトウェアを改善して、圧縮率と処理速度の両立を図っています。図では左下ほど高性能であることを示していますが、NTTで開発したソフトは世界のトップレベルの性能を発揮していることがわかります。現状のPCで、演奏時間の1/100くらいの時間で圧縮解凍処理が可能です。さらにPC上で複数のファイルやフォルダ内のデータを一括して圧縮するためのフォーマットや、家電製品で共通に使っていただくための伝送フォーマットについて、その設計や標準化への提案も進めています。

圧縮率と処理時間の比較

