

ロスレス符号化ツールの開発と評価用高精細音楽作品の制作

茂出木敏雄*

[抄録] 本稿では、主として高精細音楽制作における作業用サウンドデータを対象として高圧縮率を実現したロスレス符号化ツールのアルゴリズムの概要と、現状 Web を通じて配布されている各種ロスレス符号化ソフトウェアと筆者らが開発したロスレス符号化ツールとの性能比較実験を行った結果を述べる。ミックスダウンされた音楽作品のデータではいずれのツール間においても、圧縮率に顕著な差が生じないが、作業用の音楽ワークデータにおいては、ツール間で圧縮率に倍以上の差が生じることがあり、本提案手法の優位性が確認できた。併せて、本提案手法を含む各種符号化手法に対する品質評価を行うために制作した 96kHz/24bits/2ch 高精細音楽作品について紹介する。

Development of Lossless Encoder Tool and Production of a High Definition Music Work for Evaluation

Toshio Modegi*

Abstract: We have developed a lossless audio encoder software tool and produced a 96kHz/24bits high definition audio content for evaluation of miscellaneous audio compression tools including our developed tool. The feature of our proposed encoder tool is its very high compression ratio mainly applied to high definition working audio materials, because this tool can remove several types of redundant data included in audio data before a main linear prediction based lossless compression processor is executed. This paper describes our proposed lossless audio encoding algorithm and experimental results including compression performance comparison among several the other currently available lossless audio compression tools.

1. まえがき

高精細オーディオの制作においては、大容量のワークデータが発生し、品質を維持するために多少の冗長性をともなっても、基本的に非圧縮で保存や伝送が行われている。作業用デジタル環境やネットワーク環境が大容量化してきているが、それらが追従できない勢いで制作現場への負荷が増大しつつある。本研究は音楽制作で発生するワークデータに照準を当てた新規なロスレス符号化ツールを提案する。

オーディオのロスレス符号化は、文献[2]にあるように、線形予測と可変長符号化を基本とした手法が一般的であり、すでに DVD オーディオに活用され商用音楽制作ツールに組み込まれている事例もある[4]。また、Web 上に本機能をもつフリーウェア/シェアウェア/商用ツール等も多

*大日本印刷株式会社 研究開発センター 先端技術研究所

*Advanced Technology Lab., Research & Development Center, Dai Nippon Printing Co., Ltd.
(e-mail: Modegi-T@mail.dnp.co.jp)

数公開されている ([6]-[14])。しかし、これらの線形予測法を主体としたツールでは、平均圧縮率は60%前後であり、この程度では制作現場に導入する効果が弱く、DigiDesign 社 ProTools Steinberg 社 Nuendo といったメジャーな高精細音楽制作ツールに組み込まれていないこともあり、業務で活用されている事例がなく、現状普及していない。

そこで、筆者らは音楽スタジオにおいて各種デジタル音響編集において冗長度が増大する作業を分類整理し、それら各々に適合したロスレス圧縮アルゴリズムを開発し、ロスレス符号化の前処理として組み込むようにした[5]。一般的な線形予測法を主体としたロスレス符号化ツールとの相乗効果により、現状のロッキー符号化並の圧縮率をロスレスで実現できる場合があり、制作現場で導入するメリットが生じるレベルの圧縮効果が期待できることを確認できたため、以下その概況を報告する。

2. デジタル音響編集における冗長度拡大

本研究で対象とする PCM 形式でサンプリングされた音響信号に対してデジタル編集されたデータを分析すると、図1に示されるような5通りの操作で冗長度が増大することがわかる。(1)量子化ビット幅の拡大と(2)サンプリング周波数の拡大は、ミックスダウンを行う際に、ソースデータの品質を他の素材と合わせるために行われ、たとえばCD音源素材を高精細音楽制作で流用する際に発生する。(3)無音部は音声信号に主にみられ、他の素材とタイミングを合わせるため挿入されることがある。ポップスでは波形が飽和する箇所がみられることがあり、無音部と同様な信号平坦部の扱いになる。(4)特定パターンの繰り返しは、ループベース音楽制作ツールで作成されたリズム音楽パターンが代表的である。(5)ステレオチャンネルへの配置はモノラル素材をステレオ収録あるいは変換された際にみられる。(1)(2)の処理ではデータ容量が増大するのに対し、(3)から(5)においてはデータ容量の増大は必ずしも伴わないが、冗長性が相対的に増大する。

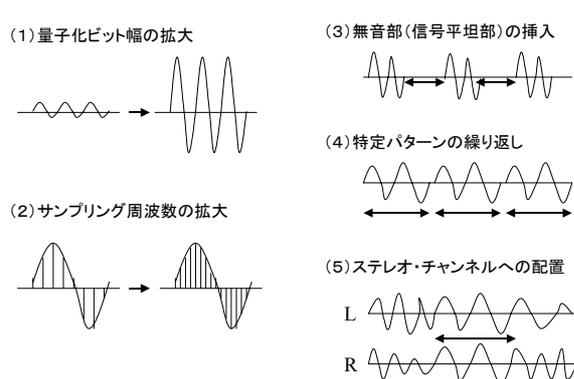


図1 各種デジタル編集に伴う冗長度の増大

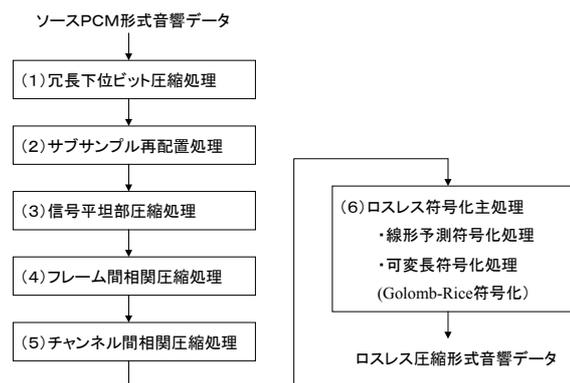


図2 提案するロスレス符号化ツールの構成

3. 提案するロスレス符号化ツールの構成

図2は図1の(1)~(5)に対応した冗長性を除去するための専用圧縮ツールをフロントエンドに配

置したロスレス符号化ツールの構成である。

(1)では、冗長な固定値（0など）で下駄をはかされた下位ビット成分を削除する。また、下位ビットの再現性を無視するニアロスレスモードの場合は、固定値でなくても削除する。

(2)では、リサンプル処理など計算補間されたサブサンプルをオリジナルサンプルと別扱いで符号化できるよう再配置する。サブサンプルは前後のサンプル値の平均値などで予測する方法をとり、主サンプルのみを後段の線形予測符号化の適用となるようにする。

(3)では、無音部や信号飽和部など同一サンプル値が連続する信号平坦部を検出除去する。

(4)では、時系列に配置された同一の波形パターンのセットを検出除去する。データを適当なフレーム単位に探索し、同一データが配置されているフレームを後方から前方に探索する。前方に同一フレームが検出された場合は後方に位置するフレームを削除する。

(5)では、チャンネル間に分散配置された類似パターンを検出除去する。左右チャンネル間に同一波形パターンが存在する場合は、Rチャンネル側のパターンを削除するようにし、類似した波形パターンの場合は、Rチャンネル側をR-Lの差分データに置換する。

(6)のロスレス符号化主処理では、まず Levinson-Durbin 法[1]に基づいて、線形予測係数の算出を適応的に行う方法をとるとともに、最大5次までの過去サンプルに基づく予測値から最も予測誤差が小さくなる次数を動的に選別するようにした。線形予測係数は1秒弱程度の間隔でアップデートすると、効果的であった。

続いて、予測誤差値の可変長符号化法として Golomb-Rice 法[2]に基づく手法を採用した。これは、局所的な平均符号長により下位ビットを二進コードで符号化し、あふれた上位ビットを単進コードで符号化する方法であるが、二進コードの幅を過去数100の予測誤差符号長の平均値で動的に決定するようにした。

勿論、(6)の処理だけでもロスレス圧縮機能は動作するが、前節で述べたような冗長性が存在すると、線形予測で予測誤差が逆に増大することがあり、本提案のような前処理を行わずに十分に圧縮をかけることが難しい。

4. 比較評価実験

表1および表2は、本提案手法に基づいて開発したロスレス圧縮ツール「ゲンコーダ」とWeb公開されている10種の他社競合ツール（[6]～[14]）と同一素材により圧縮性能比較を行った結果である。各マスの左上k単位数値は圧縮後の容量[kbytes]を右下の%数値は上段記載の原容量に対する圧縮率を示しており、ハッチングされている箇所は本提案より圧縮成績が優れたものを示す。

本提案より圧縮率の優れたツールは幾つか存在し、特に表2では“LA”のが殆どの素材において最良であったが、本提案との差は顕著ではない。表1では素材により圧縮率の差が倍以上に開くことがあったが、本提案では常に優秀なツール群に属し、表1の2つのワーク素材および表2の高精細ミックス素材において、本提案は2位以降と顕著な差をつけて最良になった。

表2では、ベートーベン第9の第4楽章素材において、PentiumIII 1GHzマシン(WindowsXP, 384Mbytes)を用いて符号化時間の比較を行った結果も示しているが、最も圧縮性能の良かった最下段の“LA”を除き、どのツールもリアルタイム処理可能なレベルであった。

表1 本提案と競合圧縮ツールとのワーク素材による性能比較実験

音素材 圧縮ツール	ヘンデルコーラス 44.1kHz/16bits 774kbytes		ステレオ化 2-ch 1548kbytes Rchコピー		時間軸2倍拡大 88.2kHz/16bits 1548kbytes 補間サンプル		振幅軸256倍拡大 44.1kHz/24bits 1189kbytes 下位0ビット		リズムパターン 48kHz/16bits 2048kbytes 繰り返し		音声アナウンス 44.1kHz/16bits 1714kbytes 無音部	
	ファイルサイズ	圧縮率	ファイルサイズ	圧縮率	ファイルサイズ	圧縮率	ファイルサイズ	圧縮率	ファイルサイズ	圧縮率	ファイルサイズ	圧縮率
ゲンコーダ (DNP、日本) (オプションOFF)	459k	59.3%	459k	29.6% (919k, 59.3%)	526k	33.9% (815k, 52.6%)	459k	38.6% (847k, 71.2%)	149k	7.2% (693k, 33.8%)	403k	23.5% (472k, 27.5%)
T. Sandworm (個人、日本)	491k	63.4%	979k	63.2%	864k	55.8%	1142k	96.0%	906k	44.2%	495k	28.8%
Shorten (SoftSound, US)	482k	62.3%	964k	62.2%	831k	53.6%	有償サポート		704k	34.3%	414k	25.1%
WaveZip (GadgetLabs, US)	507k	65.5%	1039k	67.1%	863k	55.7%	有償サポート		757k	36.9%	431k	24.1%
Wave Archiver (Dennis Lee, CA)	493k	63.6%	483k	31.2%	831k	53.6%	1189k	100.0%	704k	34.3%	414k	24.1%
FLAC 国家プロジェクト (Open Source, US)	476k	61.4%	476k	30.7%	エラー発生		475k	39.9%	725k	35.4%	425k	24.7%
WavPack (ConiferSoft, US)	463k	59.8%	511k	33.0%	782k	50.5%	705k	59.2%	753k	36.7%	468k	27.3%
LPAC (T. Liebchen, DE)	455k	58.7%	455k	29.3%	748k	48.3%	455k	38.2%	685k	33.4%	401k	23.3%
Monkey Audio (Ashland, US)	452k	58.3%	452k	29.1%	778k	50.2%	841k	70.7%	660k	32.2%	536k	31.2%
RKAU (MalcomTaylor, NZ)	449k	58.0%	450k	29.0%	エラー発生		450k	37.8%	682k	33.3%	390k	22.7%
LA, LosslessAudio (M. Bevin, US)	437k	56.4%	548k	35.4%	721k	46.5%	サポートせず		646k	31.5%	401k	23.3%

表2 本提案と競合圧縮ツールとのミックス素材による性能比較実験

音素材 圧縮ツール	ベートーベン交響曲第9番 カラヤン指揮ベルリンフィル盤 44.1kHz/16bits/2-ch				DNPオリジナル「音響モザイク」	
	第1楽章 162,711kbytes	第2楽章 113,781kbytes	第3楽章 169,990kbytes	第4楽章 243,644kbytes	96kHz/24bits 2chミックス原曲 54,198kbytes	44.1kHz/16bits CD-DA版 16,598kbytes
ゲンコーダ (DNP、日本) (オプションOFF)	88,030k	56,214k	79,070k	127,537k	9,666k	8,620k
	54.1%	49.4%	46.5%	452sec, 52.3%	17.8% (36,107k, 66.6%)	51.9%
T. Sandworm (個人、日本)	119,841k	81,166k	117,254k	179,595k	エラー発生	9,094k
	73.6%	71.3%	68.9%	109sec, 73.7%		54.7%
Shorten (SoftSound, US)	91,708k	57,372k	80,503k	132,976k	有償サポート	8,876k
	56.3%	50.4%	47.3%	78sec, 54.5%		53.4%
WaveZip (GadgetLabs, US)	95,575k	60,382k	84,201k	139,494k	有償サポート	8,964k
	58.7%	53.0%	49.5%	54sec, 57.2%		54.0%
Wave Archiver (Dennis Lee, CA)	94,121k	59,244k	83,364k	136,388k	54,198k	8,859k
	57.8%	52.0%	49.0%	75sec, 55.9%	100.0%	53.3%
FLAC 国家プロジェクト (Open Source, US)	88,733k	56,327k	78,209k	128,693k	16,763k	8,439k
	54.3%	49.5%	46.0%	167sec, 52.8%	30.9%	50.8%
WavPack (ConiferSoft, US)	88,070k	55,750k	78,502k	127,630k	26,997k	8,272k
	54.1%	48.9%	46.1%	52sec, 52.3%	49.8%	49.8%
LPAC (T. Liebchen, DE)	85,561k	54,465k	76,640k	123,675k	16,115k	8,265k
	52.5%	47.8%	45.0%	197sec, 50.7%	29.7%	49.7%
Monkey Audio (Ashland, US)	84,976k	54,070k	75,956k	122,861k	34,474k	8,068k
	52.2%	47.5%	44.6%	94sec, 50.4%	63.6%	48.6%
RKAU (MalcomTaylor, NZ)	84,368k	53,817k	74,700k	122,434k	エラー発生	8,152k
	51.8%	47.2%	43.9%	387sec, 50.2%		49.1%
LA, LosslessAudio (M. Bevin, US)	81,451k	52,062k	71,819k	118,474k	サポートせず	7,822k
	50.0%	45.7%	42.2%	1332sec, 48.6%		47.1%

5. 評価用高精細音楽作品の制作

前節の表2の性能比較実験で使用したDNPオリジナル素材「音響モザイク」は、各種ロスレス符号化ツールおよび既存のロッキー型符号化ツールにおける圧縮性能および品質の評価を行なえるように新規に制作した高精細音楽作品である。

特徴として、素材およびミックス作品のオリジナルを96kHz/24bitsでサンプリングおよびマスタリングすることにより、DVDオーディオからCD-DAレベルまでスケラブルな品質を提供可能である。表3の右端の数値は本提案手法でロスレス圧縮を行ったときの圧縮率で、符号化適性が異なる主要素材14点をピックアップし、述べ40以上の素材をミックスさせ、圧縮性能を一括評価が行えるようにした。また、主要素材に対しては素材別評価が行え、評価用素材を目的に応じて自在に再構成しリミックスできるようにした。

表3 評価用高精細音楽作品のミキシング構成

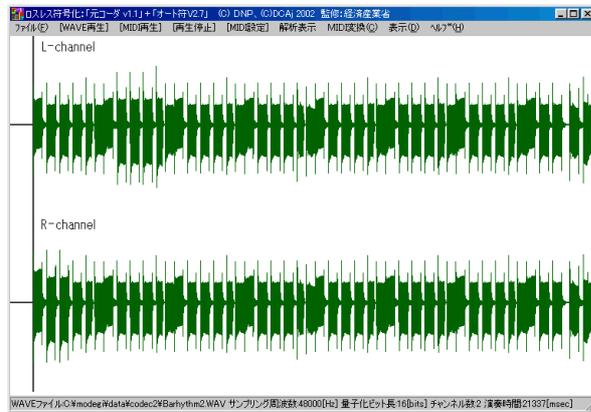
	10[sec]	20	30	40	50	60	70	80	90	圧縮率
ピアノ										38%
ハープ										57%
トランペット										53%
和太鼓										40%
デジェリドウ										38%
シタール										36%
ゴスペル歌声										61%
ソプラノ歌声										51%
アナウンスS										63%
アナウンスE										61%
鐘										35%
町の雑踏										55%
波										61%
蒸気機関車										54%
白色雑音										83%

6. あとがき

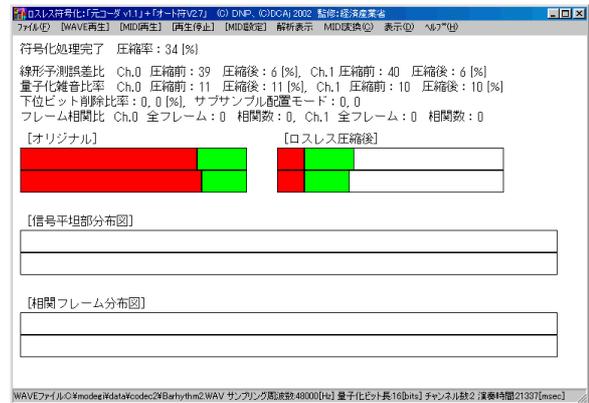
図3は本符号化ツールをループベースで制作したベース楽器パターンに対して適用した結果である。このような事例では、線形予測符号化のみの圧縮に比べ圧縮率で3倍以上の差が生じ、ロッキー符号化並の圧縮率を得ることができている。

本符号化ツール開発および評価用高精細音楽制作は、財団法人デジタルコンテンツ協会の平成14年度「次世代デジタルコンテンツ制作支援事業」の一環として、経済産業省より助成を受け推進したものである。本稿で紹介した符号化ツールは、静止画・動画のロスレス圧縮に対応するよう展開を進めるとともに、機器組み込み用LSI開発も進めている。また、冗長度を解析する新たな音響解析ツールとして、生体音響信号やゲノム情報の圧縮など医療情報分野にも展開を検討

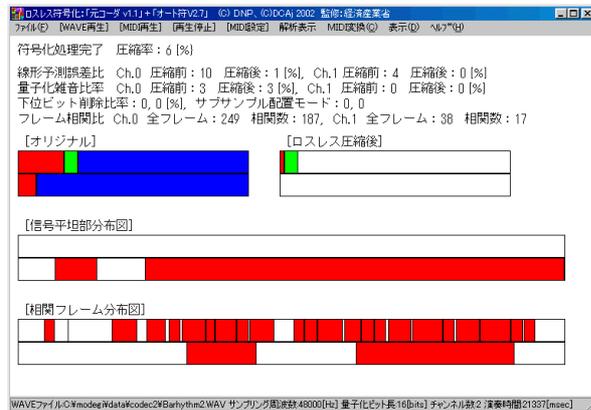
している[3]。本稿で紹介したツールは、既に幾つかの音楽プロダクションに無償配布し、試験運用を開始しており、同財団のHPを通じても無償配布を行う予定である (<http://www.dcaj.or.jp>)。



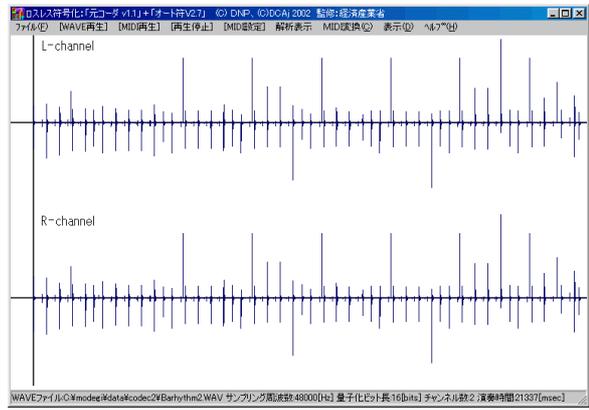
(A) ソース音響データ



(B) 線形予測のみによる圧縮結果



(C) フレーム間相関を用いた圧縮結果



(D) 符号化対象の予測誤差データ

図3 開発したロスレス符号化ツール「ゲンコーダ」の画面

参考文献

- [1] 藤原洋編：「マルチメディア情報圧縮」, インターネット時代の数学シリーズ5, 共立出版, (Mar.2000).
- [2] 奥村晴彦：「データ圧縮の基礎から応用まで」, Cマガジン, Vol.154, No.7, ソフトバンク pp.13-35 (Sep. 2002).
- [3] 茂出木敏雄：「生体音響信号のロスレス符号化手法」, 第22回医療情報学連合大会, ハイパーデモ H-09, pp.777-778 (Nov.,2002).
- [4] Khalid Sayood: "Lossless Compression of Audio Data", Lossless Compression Handbook, chapter 12, Academic Press, Elsevier Science USA, pp.255-267, 2003.
- [5] 茂出木敏雄：「高精度オーディオのロスレス符号化ツールの開発」, 情報処理学会 第65回全国大会論文集, Vol.3, pp.567-568 (Mar. 2003).
- [6] Tequila Sandworm: <http://www2u.biglobe.ne.jp/~tequila/audio000.html>
- [7] Shorten: <http://www.softsound.com/Shorten.html>
- [8] WaveZip: <http://www.5star-shareware.com/Music/WAVPlayersandUtilities/wavezip.html>
- [9] Wave Archiver: <http://www.firstpr.com.au/audiocomp/lossless/wavarc/0readme.html>
- [10] FLAC: <http://flac.sourceforge.net/>
- [11] LPAC: <http://www.nue.tu-berlin.de/wer/liebchen/lpac.html>
- [12] Monkey Audio: <http://www.monkeysaudio.com/>
- [13] RKAU: <http://rksoft.virtualave.net/rkau.html>
- [14] LA Lossless Audio: <http://www.lossless-audio.com/>