

MIDIデータのリアルタイム圧縮評価

6B-6

下川 義弘
NTT情報通信網研究所

1 はじめに

ISDNを始めとするデジタル網が普及してきたことから、加入者線の空き時間を利用してデータ通信を実現するデジタルオフトークサービスを検討している。当該サービスを広く一般家庭にまで浸透させるためにはBGM(Back Ground Music)、更に静止画、テロップを組み込んだBGVのような親しみやすいメニューを取り入れるのが効果あるが、例えばN-ISDNを例にとると通信速度はBチャンネル1本(64kbps)を使った場合、たかだかアナログ帯域に換算して7kHz、ラジオのモノラル音声AM放送1回線程度に過ぎず、既存のメディアと比較して魅力あるサービスとはいえない。音響信号を提供するには回線容量に問題はあるが、音符符号ならば不足を感じずに済む。音符符号にはキャプションシステムで提供されているような簡易なものもあるが、当該サービスを検討するにあたっては最近のMIDI(Musical Instrument Digital Interface)^{[1][2]}対応楽器、音源等の低価格化を考慮し、当該コードを対象に検討中である。本稿では前処理を施したMIDIコードを多チャンネル有線放送イメージで多重伝送する際に、リアルタイムデータ圧縮処理を加えることにより回線使用効率を改善した効果等について報告する。

2 デジタル網とMIDI

MIDIは、楽器間の演奏情報をデジタル通信するためのデファクトスタンダードである。15m以内のローカルな環境の下で調歩同期 31.25kbpsの通信ができる仕様になっているが、タイミング情報を含み、かつ、ピーク時に対応した値であることから、多少のバッファリングで平滑化すれば実際の伝送に必要な容量は更に下回る。従って、MIDIコードをN-ISDN等のデジタル網を通して送信しようとする、多重化、或はマルチメディア化できる容量の余裕がある。多重化を実現する手法として、各曲を予め単位時間毎にセグメント化したファイル構造に変換して(図1)、単位時間分だけ各曲をサイクリックに送信する手法が考えられる(図2)。現在市販されているMIDI楽曲データを例にとると、セグメント化後の伝送容量は1曲当り約250バイト/秒であり(図3)、ピークの継続時間が短く、他曲とピークが重なる確率は小さいことから、約

30曲の多重化が可能な見積りになる。

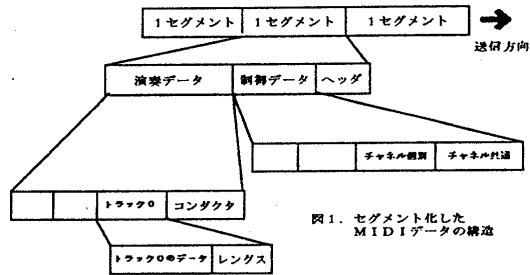


図1. セグメント化したMIDIデータの構造

多重化手法は実現できたが、セグメント化前のMIDIデータ1曲あたりのデータ量が数10%程度増大する問題点が残る。また、MIDIで定義されたエクスクルーシブを利用して、静止画やテロップなどの付加情報を加えたマルチメディア化まで実現しようとする、更に伝送量が増大することが考えられ、データ圧縮技法の導入による回線使用効率の改善が必要となる。

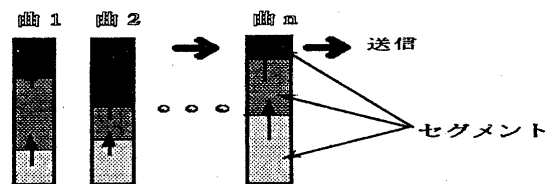


図2. サイクリックな送信

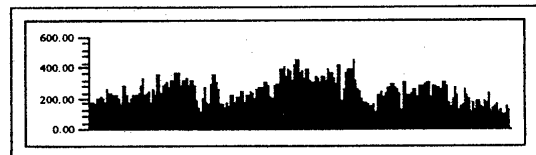


図3. データ伝送量の時間推移例(横軸-秒,縦軸-バイト)

3 LZW手法の適用による圧縮効果

オフトークの情報源は必ずしも蓄積データではなく、圧縮伝送にはリアルタイム処理できるものが適している。リアルタイム系の技法として代表的なLZW(Lempel-Ziv, Welch)^[3]の手法を検討対象とした。圧縮効率のみを考えればスライディング辞書法であるLZSS(Lempel-Ziv, Storer, Szymanski)^[4]の手法の方が、より効果が高いが、CCITT V.42bis^[5]でBTLZ(British Telecom-modified LZ)が採用されている

Study on compression of MIDI code using Lempel-Ziv, Welch method
Yoshihiro SHIMOKAWA (simokawa@nttcom.ntt.jp)
NTT Network Information Systems Laboratories, Musashino-Shi, Tokyo

ように、LZWの実用性は高い。LZW等のリアルタイム圧縮技法をデジタルオフトークを用いた放送型サービスに適用する場合、各受信端末がどの時点からデータを取り込み始めるかわからないことから注意が必要となる。いい替えると、曲の途中からの再生にどう対処するか、LZW処理のための送信用辞書の初期化をどのような契機で行うか、といった課題がある。図4にセグメント化したMIDIデータ1曲に対するLZW、LZSSとその他の圧縮技法による圧縮率を示す。

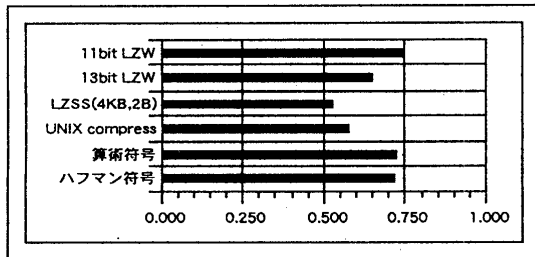


図4. 曲単位のデータ圧縮効果例

LZW手法と異なり可変長符号を排出するLZSS手法は扱いにくい面もあるが、圧縮率は一般にLZW手法よりも高く、辞書の初期化周期によってはLZW処理よりも適した領域があることに注意する。

4 辞書の初期化周期と圧縮率の試算

LZW手法で生成される符号化辞書は統計的に類似した集団別に複数用いると全体の圧縮効果が改善される可能性が高い。ここでは前提として各曲対応に辞書を割り当てることとした。その場合、辞書の初期化契機は圧縮効率から見て曲と曲の間が有効だが、途中再生の端末では最悪1曲分の再生待ち時間が発生する。待ち時間を短縮するためには辞書を途中で送信する方法と初期化周期をより短くする方法があり、LZW手法に適した後者の試算を行った(図5)。

図5において凡例(横軸)の11bitとは最大符号化長が11bitまでを許容されることを意味し、縦軸に辞書の初期化周期をセグメント数(秒数)を、垂直の軸にセグメント化データに対する圧縮率を示す。

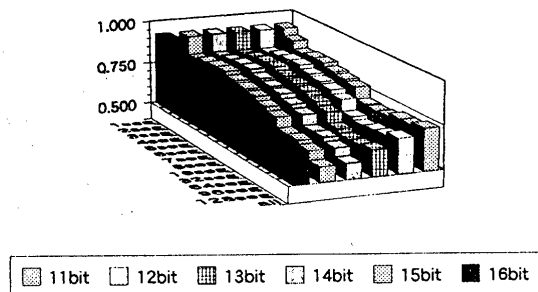


図5. 初期化周期と圧縮効果の例

5 考察

1セグメントのデータ量が250バイト程度であることから数10セグメントの周期も数kバイトのデータ量に過ぎず、数セグメントの周期と比べて改善効率は低い。むしろ、端末での実用的な待ち時間を考えれば5セグメントも小さくはなく、2ないし3セグメント程度を初期化周期とするのが適当な領域と見ることが出来る。最大符号化長を12bitまで許容すれば、周期を1セグメントから2セグメントに延ばすことで10%近い改善効果が発生する点が重要である。まとめると、LZW処理を行うことで20%程度の改善効果が期待でき、MIDIデータのセグメント化による多重化手法のもとで40ないし50曲多重(多チャンネル)が実現できる。

ところで辞書の初期化周期が数セグメント(数秒)程度に小さい場合にはスライディング辞書による検索はさほど問題にならない。むしろ、

- (1) MIDIコードは主に3バイト単位のデータ系列の集合であり、データバイトまで含めて同パターンのメッセージが非常に多い。特に同一MIDIチャンネルでは顕著。
- (2) チャンネルメッセージであるノートオン、オフとコントロールチェンジが圧倒的に多い。

ことから、LZW手法よりも系列一致の効率が高くなる可能性が考えられる。現在、LZSS手法(バッファ512バイト,最大一致長系列2,3バイト前後)の処理能力についてLZW手法と比較検討中である。

また、MIDIに静止画、テロップ等を追加することで、MIDIの統計特性と異なるデータの比率が高まった場合、追加されたデータの特性に応じた別の圧縮技法を個別に適用する必要がある。

6 むすび

デジタルオフトークサービスでMIDIを利用した多チャンネルBGMを実現するためのデータのリアルタイム圧縮効果を試算し、N-ISDNを例にとればおおよそ40~50曲程度の多重化が可能であることを確認した。今後は非リアルタイム圧縮技法も含めたデータ処理方法の検討とともに、MIDIのエクスクルーシブを利用したマルチメディア化への対応、一方向伝送の環境における誤り訂正制御との関係整理などが課題である。

参考文献

- [1] MIDI-1.0規格書, MIDI規格協議会, 1984
- [2] MIDI-1.0詳解(改定版), MIDI規格協議会, 1987
- [3] T.A. Welch, "A Technique for high performance data compression", Computer, vol. 17, No. 6, pp. 8-19, June 1984
- [4] James A. Storer, "Data Compression: Methods and Theory.", Computer Science Press, Rockville, MD., 1988
- [5] CCITT Study Group XVII Temporary Document 97-E rev., Geneva, 26-29, Sep. 1989