

低遅延 MPEG-2 ビデオトランスコーダ符号量制御方式の検討

笠井 裕之^{† †} 花村 剛^{† † †} 西村 敏[†] 亀山 渉^{† †} 富永 英義^{† †}

[†] 早稲田大学 理工学部 電子・情報通信学科
〒169-8555 東京都新宿区大久保 3-4-1

^{† †} 早稲田大学 国際情報通信研究センター
〒169-0051 東京都新宿区西早稲田 1-3-10

^{† † †} 株式会社 メディアグルー
〒169-0072 東京都新宿区大久保 2-4-12

kasai@tom.comm.waseda.ac.jp

本稿では、ネットワーク中継ノードにおけるビットレート削減器としてのトランスコーダに着目し、低遅延 MPEG-2 ビデオトランスコーダ符号量制御方式を提案する。まず、MPEG-2 トランスコーダ符号量制御方式への要件について整理する。そして、低遅延性、GOP 構造非依存性を実現する MPEG-2 トランスコーダ符号量制御方式を提案し、その詳細を説明する。さらに、トランスコーダにおける送受信バッファを解析することによりトランスコード処理に要求される遅延量について解析する。最後に、符号量変動、画質変動、遅延量、GOP 構造非依存性の観点から本提案方式を評価する。その結果、提案符号量制御方式は、画質に関して符号量制御方式として TM 方式を適用した従来方式とほぼ同等の特性があり、入力されるビットストリームの GOP 構造に依存せず、シーンチェンジによる GOP 構造変更のあるビットストリームに対しても低遅延で安定した符号量制御を実現できることを明らかにした。

キーワード トランスコーダ, 符号量制御, 低遅延性, MPEG-2

Rate Control Scheme for Low Delay MPEG-2 Video Transcoder

Hiroyuki KASAI^{† †} Tsuyoshi HANAMURA^{† † †} Satoshi NISHIMURA[†]
Wataru KAMEYAMA^{† †} Hideyoshi TOMINAGA^{† †}

[†] Dept. of Electronics, Information and Communication Engineering, WASEDA University
3-4-1 Ohkubo Shinjuku-ku, Tokyo, 169-8555 JAPAN

^{† †} Global Information and Telecommunication Institute, WASEDA University
1-3-10 Nishi-Waseda Shinjuku-ku, Tokyo 169-0051 JAPAN

^{† † †} Media Glue Corporation
2-4-12 Ohkubo Shinjuku-ku, Tokyo 169-0072 JAPAN

kasai@tom.comm.waseda.ac.jp

In this paper, we propose the rate control scheme for low delay MPEG-2 video transcoder which reduce the bitrate of MPEG-2 bitstream in network node. At first, we state the requirement of the rate control for low delay transcoder. Next the rate control scheme for low delay transcoder is proposed. And, we analyze the required delay time for the proposed scheme. At last, we evaluate the proposed scheme from the image quality, delay time and GOP-structure independency. As a result, we show the effectiveness of proposed scheme.

keyword Transcoder, Rate Control, Low Delay, MPEG-2

1. はじめに

異種帯域混在型ネットワークへのビデオ配送システムにおいて、ネットワークノードにおけるビットレート削減機能を実現する処理器として、ビデオトランスコーダ(以下、トランスコーダ¹と呼ぶ)が提案されている⁽¹⁾。本トランスコーダにおいては、低遅延性が要求されており、文献(1)では、H.261⁽²⁾を対象とした低遅延トランスコーダの検討が行われている。特に、H.261では、符号化アルゴリズム、符号化画像構造等が簡易であり、各フレーム間の符号量変動も小さくスキップ等も可能なため、低遅延トランスコーダのための符号量制御は比較的簡易に実現できる。しかしながら、MPEG-2⁽³⁾はH.261と比較した場合、各ピクチャにおいてピクチャタイプが存在し、それらのピクチャ周期等によりGOP構造が依存するだけでなく、それらGOP構造がシーンチェンジ等によりビットストリーム中で変化するというビットストリーム画像構造可変性を有する。そこで、GOP構造に非依存な符号量制御の実現を考えた場合、ビットストリームの復号符号量情報に基づき符号量制御が実現可能であるが、復号処理遅延が生じてしまうという問題がある。これに対して、文献(4)ではビットストリーム中にGOP構造をはじめとした符号化情報を記述しておくことにより低遅延レート制御を実現しているが、限定されたビットストリームのみしか対応できない点で問題が存在する。このように、ネットワークノードにおけるトランスコーダでは、入力GOP構造に非依存な低遅延符号量制御方式の開発が要求される。そこで本稿では、GOP構造に非依存な低遅延MPEG-2ビデオトランスコーダ符号量制御方式を提案し、符号量変動、画質変動、遅延量、構造非依存性の観点から本提案符号量制御方式を評価し、有効性を示す。

2. MPEG-2 トランスコーダ符号量制御方式への要件

Test Model 5⁽⁵⁾(以下、TMと呼ぶ)符号量制御方式では、予めN、M値等のGOP構造パラメータを決定し、GOP内の残符号量と残ピクチャ数及びピクチャタイプ情報とから、当該ピクチャの符号量を配分する。しかしトランスコーダに入力されるビットストリームは、シーンチェンジ等に起因してGOP構造が変化することが想定されるため、ビットストリームの復号処理以前に、GOP構造情報を仮定し、当該ピクチャへの符号量を行うことは難しくなる。よってトランスコーダ符号量制御において、N/M値等のGOP構造に依存しない符号量制御を採用する必要がある。

上記要件から、入力ビットストリームの復号処理により得られる積算入力符号量により、当該ピクチャ目標符号量を算出することでN/M値に非依存な符号量制御が可能になるが、当該ピクチャビットストリーム素片の復号処理に伴う遅延が問題となる。よって、入力ビットストリームの長時間に及ぶ符号量測定を行わない方式である必要がある。

一方、入力ビットストリーム符号量の長時間の測定を行わない場合においても、トランスコーダの受信側バッファ及び送信側バッファのビットストリーム滞留時間が増加する。よって、入力ネットワークの伝送ビットレートを考慮した、バッファ滞留時間を抑制する符号量制御方式を実現することが要求される。

さらに、出力ビットストリーム内のvbv_delay値が、トランスコーダ対象ビットストリームのそれより大きい場合、復号器での処理待機遅延が生じるため、出力ビットストリームのvbv_delay値を考慮した符号量制御が必要となる。

以下に、低遅延トランスコーダ符号量制御への要件を整理する。

- N/M値のGOP構造情報に非依存である
- ビットストリーム符号量長時間測定を行わない
- バッファ滞留時間を最小限に抑制する
- vbv_delay値の増大を抑制する

3. 低遅延MPEG-2トランスコーダ符号量制御方式

3.1 提案方式概要

本論文では、ピクチャタイプ毎の符号量配分を行わず、各マクロブロック(以下、MBと呼ぶ)毎の量子化スケールの算出に基づく符号量制御方式を提案する。具体的には、一定時間幅のウィンドウ(スライディングウィンドウ)をずらしながら、入力ビットストリーム中のフレームレート F 、入力ビットレート値 B_{in} 、出力ビットストリームのターゲットビットレート値 B_{out} をもとに、入力ビットストリーム所定時間の残符号量比率 $ioRatio(n)$ を算出することにより、TM5 Step 1と等価な処理を実現する。そして各MB毎に、出力側バッファ状態値、入力側バッファ状態値に $ioRatio(n)$ を乗じた値を元に、各MBの量子化スケールを算出す

¹一般的にトランスコーダとは広義では符号化体系変換器を指すが、本稿ではビデオ情報ビットレート削減器を指す。

ることで、画面内の発生符号量に基づくフィードバック制御を行うことにより、TM5 Step 2と等価な処理を実現する。

3.2 符号量制御アルゴリズム

3.2.1 Step 1: 入出力ビットレート比率更新制御

入力ビットストリーム所定時間内残り符号量 $R_{in}(n)$ 、及び出力ビットストリーム所定時間内残り符号量 $R_{out}(n)$ から、 $ioRatio(n)$ を設定する。

第 $n(\geq 0)$ 番ピクチャのトランスコード前に、 $R_{in}(n)$ 、 $R_{out}(n)$ を式(1)により更新する。但し、スライディングウィンドウの時間幅を W (単位は[frame])とする。

$$\begin{cases} R_{in}(n) = R_{in}(n-1) - S_{in}(n-1) + \frac{B_{in}}{F} \\ R_{out}(n) = R_{out}(n-1) - S_{out}(n-1) + \frac{B_{out}}{F} \end{cases} \quad (1)$$

式(1)において、 $S_{in}(n)$ は入力ビットストリームにおける n 番ピクチャ復号符号量、 $S_{out}(n)$ は出力ビットストリームにおける n 番ピクチャ発生符号量を示す。式(1)から、第 n 番ピクチャトランスコード時の $ioRatio(n)$ は式(2)で算出する。

$$ioRatio(n) = \frac{R_{out}(n)}{R_{in}(n)} \quad (2)$$

但し、 $R_{in}(n), R_{out}(n)$ の初期値は式(3)の通り。

$$\begin{cases} R_{in}(0) = B_{in} \times W / F \\ R_{out}(0) = B_{out} \times W / F \end{cases} \quad (3)$$

3.2.2 Step 2: 入出力バッファ管理による量子化パラメータ算出

各MBにおける入力バッファ状態値に算出した $ioRatio(n)$ を乗じて算出されるバッファ状態値と、出力バッファ状態値との差をフィードバックすることにより量子化スケールを導出する。まず、 j 番目MBのトランスコードに先立ち、仮想バッファ占有量 $b(j)$ を式(4)により算出する。

$$b(j) = b(0) + \tau \times \sum_{k=1}^{j-1} \left\{ B_{MB_{out}}(k) - B_{MB_{in}}(k) \times ioRatio(n) \right\} \quad (4)$$

式(4)において、 $B_{MB_{out}}$ は出力ビットストリームにおけるピクチャ先頭から k 番目のMBの発生符号量、 $B_{MB_{in}}$ は入力ビットストリームにおけるピクチャ先頭から k 番目のMBの発生符号量を示す。また、 τ はリアクション係数を示し、フィードバックの強さを制御する。式(4)から算出した仮想バッファ占有量を用いて、量子化スケールは式(5)により算出する。

$$Q(j) = \frac{31 \times F}{2 \times B_{out}} \times b(j) \quad (5)$$

式(4)において $b(0)$ は仮想バッファの初期占有量を示し、例えばTM5で用いられる初期値($=10 \times 2 \times R_{out} / (F \times 31)$)を使用する。また $b(MB_{cnt})$ は、次のピクチャに対する仮想バッファ占有量の初期値 $b(0)$ として用いる。但し MB_{cnt} は1ピクチャ内のMB総数を示す。

4. 送信バッファ解析に基づくトランスコーダ遅延量解析

トランスコーダ受信バッファ及び送信バッファの解析に基づき、トランスコーダにおける総遅延量を理論的に解析する。ここで、トランスコーダのレート変換処理に要する処理時間は考慮しないものとする。

4.1 受信バッファ解析に基づく遅延量解析

本提案符号量制御方式では、トランスコード対象となる当該MBの復号処理終了後にトランスコード処理が開始されるため、処理開始時刻に対する制約条件は、受信バッファ状態値が当該MB入力ビットストリーム符号量に相当する蓄積量に満ちることである。よって本制約条件は、処理開始時刻 $T_{in}(j)$ が式(6)を満足することと等価である。

$$\int_0^{T_{in}(j)} B_{in} dt \geq B_{H_{in}} + \sum_{k=0}^j B_{MB_{in}}(k) \quad (6)$$

但し、式(6)における $B_{H_{in}}$ は入力ビットストリームにおけるヘッダ情報符号量を示す。以上から、トランスコーダにおける j 番目MBに対する処理開始時刻 $T_{in}(j)$ は式(7)で表現される。

$$T_{in}(j) = \frac{B_{H_{in}}}{B_{in}} + \frac{1}{B_{in}} \sum_{k=0}^j B_{MB_{in}}(k) \quad (7)$$

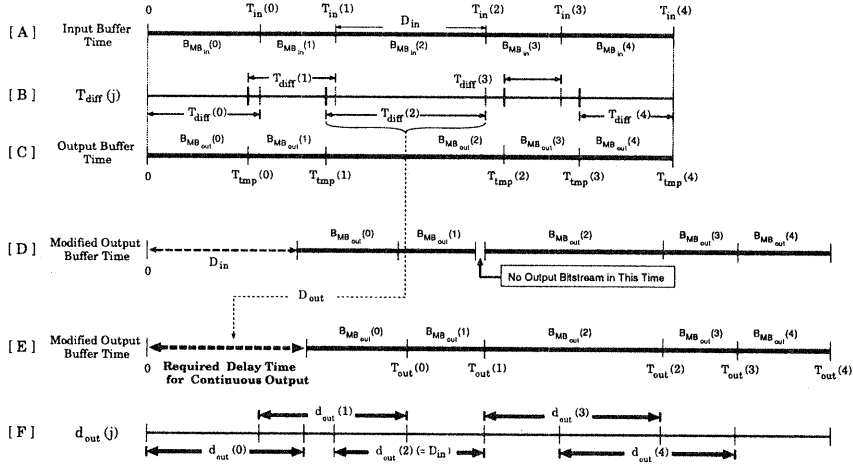


図 1: 送信バッファに基づく処理開始要求総遅延量及びトランスコーダ総遅延量

以上から、 j 番目 MB に対する処理開始要求遅延量 $d_{in}(j)$ は式 (8) で示される。

$$d_{in}(j) = T_{in}(j) - T_{in}(j-1) \quad (8)$$

よって式 (8) から、受信バッファ解析に基づく処理開始要求遅延量 D_{in} は式 (9) で表現される (図 1:[A])。但し、式 (9) において NMB は入力ビットストリーム内 MB 総数を示す。

$$D_{in} = \max_{0 \leq j \leq NMB} \{ T_{in}(j) - T_{in}(j-1) \} \\ = \frac{\max_{0 \leq j \leq NMB} B_{MB_{in}}(j)}{B_{in}} \quad (9)$$

4.2 送信バッファ解析に基づく遅延量解析

一方、 j 番目 MB 出力ビットストリームのトランスコーダ出力完了時刻 $T_{tmp}(j)$ は式 (10) で表現される (図 1:[C])。

$$T_{tmp}(j) = \frac{B_{H_{out}}}{B_{out}} + \frac{1}{B_{out}} \sum_{k=0}^j B_{MB_{out}}(k) \quad (10)$$

但し、式 (10) における $B_{H_{out}}$ は出力ビットストリームにおけるヘッダ情報符号量を示す。以上から、 j 番目 MB 入力ビットストリーム入力開始時刻、つまり j 番目 MB 処理開始時刻 $T_{in}(j)$ と、 $j-1$ 番目 MB 出力ビットストリーム出力完了時刻 $T_{tmp}(j-1)$ の差分値を $T_{diff}(j)$ と定義した場合、 $T_{diff}(j)$ は式 (11) で表現される (図 1:[B])。

$$T_{diff}(j) = T_{in}(j) - T_{tmp}(j-1) \quad (11)$$

ここで、 $T_{diff}(j)$ が負の場合、 j 番目 MB 入力ビットストリームの処理開始時刻の前に、 $j-1$ 番目 MB 出力ビットストリームの出力処理が完了していることを意味し、この場合、送信バッファから、ビットストリームが非連続的に出力されることを意味する (図 1:[D])。よって、入力ビットストリームのトランスコーダ処理開始時刻に対する必要条件は、 D_{in} の最小化に加えて、ビットストリーム出力の連続性確保であり、これはトランスコーダ送信バッファがアンダーフローしない条件 $T_{diff}(j) \leq 0$ と等価である。よって、本必要条件を満たすための送信バッファ要求処理開始遅延量 D_{out} は式 (12) で表現される (図 1:[E])。

$$D_{out} = \max_{0 \leq j \leq NMB} T_{diff}(j) \\ = \max_{0 \leq j \leq NMB} \{ T_{in}(j) - T_{tmp}(j-1) \} \quad (12)$$

式 (10)、(12) から j 番目 MB の実際の出力完了時間 $T_{out}(j)$ は式 (13) で表現される。

$$T_{out}(j) = D_{out} + \frac{B_{H_{out}}}{B_{out}} + \frac{1}{B_{out}} \sum_{k=0}^j B_{MB_{out}}(k) \quad (13)$$

よって、式 (7)、(13) から j 番目 MB における発生遅延量 $d_{out}(j)$ は式 (14) で示される (図 1:[F])。

$$d_{out}(j) = T_{out}(j) - T_{in}(j) \quad (14)$$

以上から、トランスコーダ処理開始における総遅延量 D_{total} は最終的に式 (15) で導出される。

$$D_{total} = \max_{0 \leq j \leq NMB} \{ T_{out}(j) - T_{in}(j) \} \quad (15)$$

ところで、 $0 \leq j \leq NMB$ の j において、 $B_{H_{in}}$ 、 B_{out} 、 $B_{MB_{in}}(j)$ 、 $B_{MB_{out}}(j)$ の関係が式 (16) を満たす場合、式 (7)、(13) から D_{total} は、式 (17) で示される。

$$\frac{B_{H_{in}} + \sum_{k=0}^j B_{MB_{in}}(k)}{B_{H_{out}} + \sum_{k=0}^j B_{MB_{out}}(k)} = \frac{B_{in}}{B_{out}} \quad (16)$$

$$D_{total} = \max_{0 \leq j \leq NMB} D_{out} \quad (17)$$

さらに、式 (16) が成り立つ時、 D_{out} が最小値を示すため、 D_{total} は式 (18) で示される。

$$D_{total} \geq \min_{0 \leq j \leq NMB} D_{out} \\ = \frac{\max_{0 \leq j \leq NMB} B_{MB_{in}}(j)}{B_{in}} = D_{in} \quad (18)$$

以上から、式 (16) が成り立つ時、総遅延量 D_{total} は D_{in} と等しいことがわかる。

一方、2. で述べた、入力ビットストリームの復号処理により得られる入力符号量積算に基づく符号量制御方式における総遅延量 D_{total} は、同様に式 (19) を満たすことが導出される。

$$D_{total} \geq \max_{0 \leq l \leq NPIC} \sum_{k=0}^{MBCnt} \frac{B_{MB_{in}}(l, k)}{B_{in}} \quad (19)$$

式 (19) において、 $B_{MB_{in}}(l, k)$ は l 番ピクチャ中 k 番 MB の復号符号量、 $NPIC$ はビットストリーム中のピクチャ総数を示す。

5. シミュレーション実験

提案符号量制御方式について、(i) 発生符号量変動及び画質変動、(ii) 遅延度、(ii) GOP 構造非依存性、の各観点についてシミュレーション実験から評価する。提案符号量制御方式を適用するト

□ 低遅延 MPEG-2 ビデオトランスコーダ符号量制御方式の検討

ランスコーダアーキテクチャとして、空間領域上のトランスコーダ(アーキテクチャ [A]), 及び DCT 係数領域上のトランスコーダ(アーキテクチャ [B]) の2つのアーキテクチャを採用する。アーキテクチャ [A] は逆量子化 / 再量子化処理の間に IDCT/MC/DCT のループ処理を有するのに対して、アーキテクチャ [B] は逆量子化 / 再量子化処理のみでビットレート削減処理を実現する^{(6) (7)}。ここでアーキテクチャ [A] への適用方式をループ処理 ON 方式、アーキテクチャ [B] への適用方式をループ処理 OFF 方式とする。

本実験では、評価対象方式として表 1 に示す各方式を採用し、比較・考察する。但し、TM 方式で使用する N, M 値は、入力ビットストリーム GOP 構造 (N, M)=(15, 3) と同一にすることで GOP 構造の差異による画質低下を回避した。また、実験においては、表 2 に示す実験用 MPEG-2 ビットストリームを使用する。さらに、表 3 に示す GOP 構造情報 N, M 値が異なるビットストリームを使用する。また本実験に共通して、トランスコード出力レートの目標値は $B_{out}=5$ [Mbps] とする。

表 1: 評価対象方式

方式	符号量制御方式	ループ処理
1	提案符号量制御方式 トランスコーダ	OFF
2	提案符号量制御方式 トランスコーダ	ON
3	TM 符号量制御方式 トランスコーダ	OFF
4	TM 符号量制御方式 トランスコーダ	ON
5	TM 符号量制御方式 符号化器	-

表 2: 実験用ビットストリーム符号化特性

符号化方式	MPEG-2 Video MP@ML (ES)
符号化レート	10, 6 [Mbps]
フレームレート	29.97 [frame/sec]
符号量制御方式	TM 方式 (Step 3 OFF)
ピクチャ構造	フレーム構造
動き補償方式	フレーム / フィールド MC
テストシーケンス	Bus, Mobile&Calendar Flower Garden, Table Tennis
画像フォーマット	ITU-R BT.601 4:2:0 Format 輝度信号 704[pe] × 480[line] 色差信号 352[pe] × 240[line]
符号化フレーム数	150 [Frame]
ビットストリームの種類	5 種類 (A, B, C, D, E) [表 3]

表 3: 実験用ビットストリーム GOP 構造

種類	N	M	フレーム番号	シーケンス
A	15	3	0 - 149	Bus
B	15	3	0 - 149	Table Tennis
C	9	3	0 - 149	Bus
D	15	5	0 - 149	Bus
E	シーン チェンジ 対応	3	0 - 38	Bus
			39 - 71	Mobile& Calendar
			72 - 122	Flower Garden
			123 - 149	Table Tennis

5.1 発生符号量及び画質変動に関する評価実験

試験対象ビットストリームとして表 3 に示すビットストリーム A 及び B を使用し、提案方式の各パラメータは $W=15$, $r=1$ とし、発生符号量及び画質変動を評価する。 $B_{in}=6, 10$ [Mbps] から $B_{out}=5$ [Mbps] へのトランスコード処理における、ピクチャタイプ別平均符号量結果を表 4 に示す。表 4 から、方式 3, 4 では B_{in} , ループ処理の ON/OFF に関わらず、方式 5 に近い符号量割当を示すのに対し、方式 1, 2 では $B_{in} = 6$ の場合は方式 5 に近いものの、 $B_{in} = 10$ の場合は、その傾向は異なり、特に I ピクチャ (フレーム番号 92, 107) の割当符号量の減少と、B ピクチャ割当符号量の増加傾向がある。これは、方式 1, 2 は、入力符号量に対して、入出力ビットレート比率に比例した符号量配分を行うため、入出力ビットレート比率が大きくなるに従い、入力ビットストリーム内符号量占有量が大きい I ピクチャ程、削減符号量の比率が増加するためである。以上から、提案方式では、入出力ビットレート比率が大きくなるに従い、入力符号量に比例した符号量削減を行うため、TM 方式と傾向が異なり、特に I ピクチャの符号量削減の減少傾向が顕著になる。

次に、各方式の平均 SN 比を表 4 に示す。表 4 に関して、ループ処理 ON/OFF のアーキテクチャの相違についてみると、提案方式、TM 方式両方式に共通して、ループ処理 OFF の場合は ON の場合と比較して約 0.4 ~ 1.0[dB] 程度の SN 比低下の傾向がみら

表 4: ピクチャタイプ別平均符号量と平均 SN 比

SEQ	B_{in}	方式	平均符号量 [Mbit/Frame]				SN 比 [dB]
			I	P	B	Total	
Bus	6	1	0.482	0.226	0.111	5.001	32.76
		2	0.484	0.227	0.111	5.001	33.15
		3	0.509	0.242	0.102	5.002	32.54
		4	0.515	0.242	0.102	5.002	33.02
	10	1	0.440	0.229	0.114	5.001	32.10
		2	0.431	0.226	0.116	5.000	33.07
		3	0.538	0.250	0.096	5.002	32.21
		4	0.485	0.247	0.103	5.002	33.18
	-	5	0.523	0.248	0.098	5.000	33.72
	Table Tennis	6	1	0.493	0.239	0.105	5.005
2			0.495	0.239	0.105	5.004	31.72
3			0.515	0.242	0.102	5.002	31.15
4			0.546	0.257	0.093	5.003	31.75
10		1	0.430	0.235	0.113	5.004	30.74
		2	0.422	0.231	0.115	5.003	31.66
		3	0.485	0.247	0.103	5.002	31.03
		4	0.505	0.264	0.094	5.003	31.90
-		5	0.557	0.266	0.088	5.001	32.47

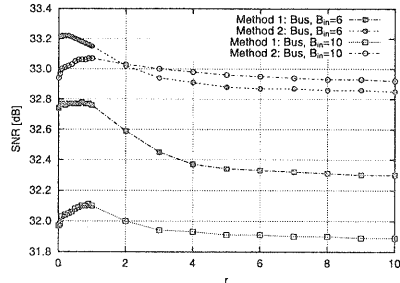


図 2: r と SN 比の関係

れる。これは、ループ処理を行わないことによる符号化器と復号器での参照画像の不一致に起因するドリフト誤差発生の影響である。次に、 B_{in} の相違、即ち入出力ビットレート比率の相違を合わせて考察すると、入出力ビットレート比率が小さい ($B_{in}=6$) 場合には、ループ処理 ON 方式では、提案方式、TM 方式とも同等の SN 比を示し、ループ処理 OFF 方式では提案方式の SN 比が上回っている。しかしながら、入出力ビットレート比率が大きい ($B_{in}=10$) 場合には、ループ処理 OFF 方式において提案方式の SN 比低下の傾向は大きく、これは、符号量変動に関する考察で前述したように、I ピクチャの符号量削減の傾向が大きいため、I ピクチャの画質低下が起こり、結果として、先に述べたドリフト誤差の影響がさらに顕著になるためである。しかしながら、ループ処理 ON 方式では B ピクチャの割当符号量が増加し、且つドリフト誤差の影響が小さいため、I ピクチャ以後の SN 比の低下が見られないことがわかる。

次に、提案方式に関して、リアクシオン係数 r を変化させた場合の SN 比の変化を図 2 に示す。

図 2 から、 $0 \leq r \leq 1$ に最大値を有し、 $r=1$ から $r=10$ への増加において約 0.3 ~ 0.5[dB] の SN 比の低下に留まっていることが確認できる。

5.2 遅延度に関する評価実験

本節では、(i) 送受信バッファに基づく遅延量、(ii) vbv_delay 値の変動、各観点から提案方式の遅延度に関する評価実験を行う。シミュレーション実験では、表 3 の内、ビットストリーム A の 6[Mbps] を使用する。

5.2.1 送受信バッファに基づく遅延量の評価

本節では、各 MB における、 $d_{in}(j)$, D_{in} , 及び $d_{out}(j)$ を実測し、 D_{total} を算出することで、発生遅延量の評価を行う。ここで、アーキテクチャによる差が僅かだったため、方式 1 における $d_{in}(j)$, D_{in} の変動結果を図 3 に示す。但し、図 3 では、シーケンス最初の 5 ピクチャのみの結果を示し、各ピクチャの区切りを縦線点で示す。また、 $W=15$ とし、ピクチャ順序はビットストリー

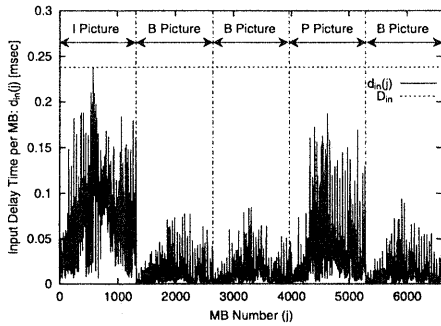


図 3: $d_{in}(j)$ の変動及び D_{in}

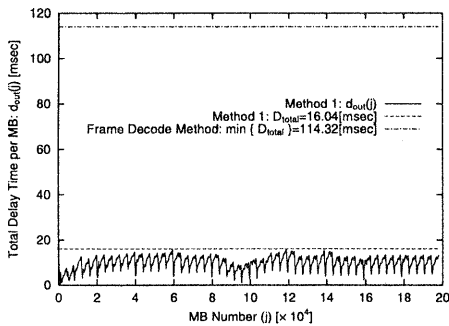


図 4: $d_{out}(j)$ の変動及び D_{total}

ム内順序であるものとする。図 3で、 $d_{in}(j)$ はピクチャタイプにより大きく異なり、I または P ピクチャの場合、B ピクチャと比較して、各 MB の復号符号量が大きいため $d_{in}(j)$ が大きくなる。よって、処理開始時刻に関する必要条件は $d_{in}(j)$ の最大値に依存することから、 D_{in} は最大符号量を有するピクチャ (図 3 では I ピクチャ) 中に存在することが分かる。

一方、 $d_{out}(j)$ 及び D_{total} の変動結果を図 4 に示す。但し、対照方式として、2 で述べた入力ビットストリームの復号処理により得られる入力符号量積算に基づく制御方式を採用し、式 (19) に示される $\min\{D_{total}\}$ を入力発生符号量から算出し、その結果を図 4 に同時に示す。図 4 から、 $d_{out}(j)$ は各 MB において約 5[msec] ~ 15[msec] であり、 D_{total} は約 16[msec] である。一方で、対照方式では、最低約 114[msec] の処理遅延量が必要であり、これは式 (19) に示される D_{total} の最小値であるため、送信バッファに基づく要求遅延量を考慮した場合、さらに D_{total} は増大する。以上から、提案符号量制御方式は、対照方式と比べて低遅延でトランスコード処理が可能である。

さらにシーケンス Bus に対して、リアクション係数 r を変化した場合の D_{total} の変化を図 5 に示す。図 5 から r の増加に対し D_{total} は単調減少し、 r を制御することで D_{total} を抑制できる。ここで図 5 と図 2 の結果を合わせると、 r を大きくするに従い D_{total} は小さくなるが、反面 SN 比が低下し、 D_{total} と SN 比とのコスト関数を定義することで最適な r の算出方法は今後の課題である。

5.2.2 vbv_delay 値の変動

本節では、トランスコーダへの入力ビットストリームの vbv_delay 値と、トランスコーダからの出力ビットストリームの vbv_delay 値との差を評価する。方式 1 のフレーム番号 30 ~ 60 の vbv_delay 値の変動を図 6 に示す。

図 6 から、入力ビットストリーム内 vbv_delay 値の変動と、出力ビットストリーム内 vbv_delay 値の変動はほぼ同じであり、特に I ピクチャでは、入力ビットストリーム内 vbv_delay 値より低下していることが分かる。よって、出力ビットストリーム受信側の復号器バッファの待機時間が、トランスコード処理を行わないビット

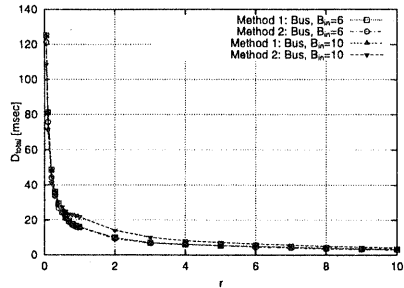


図 5: r と D_{total} の関係

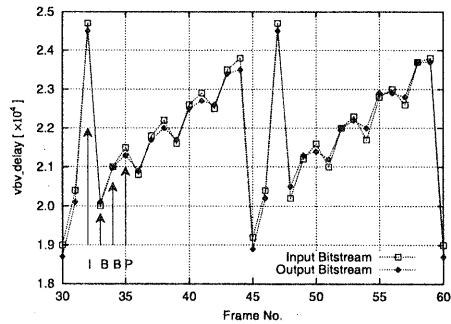


図 6: 各フレームにおける vbv_delay 値の変動

ストリームと同一となり、トランスコーダにより出力されるビットストリームによる復号器内での遅延量が十分無視できる量である。

5.3 GOP 構造非依存性に関する評価実験

表 3 に示すビットストリーム A, C, D, E を使用し、提案方式における W を変化した場合の D_{total} の変動を測定する。 D_{total} の変動結果を図 7 に示す。但し、図 7 には、方式 1 の場合の結果を示す。さらに、 $W = 15$ の場合におけるバッファ総遅延量 D_{total} 及び SN 比の結果を、GOP 構造情報を既知とする TM 符号量制御適用方式 (方式 3) の結果と併せて表 5 に示す。

表 5: 各ビットストリームのトランスコード処理における各方式の D_{total} [msec]、SN 比 [dB] の変動

B_{in}	ビットストリーム	方式 1		方式 3	
		D_{total}	SN 比	D_{total}	SN 比
6	A	16.04	32.76	12.88	32.54
	C	14.17	32.75	13.46	32.48
	D	18.86	33.04	18.52	32.78
	E	21.79	29.32	39.17	29.27
10	A	23.91	32.10	26.91	32.21
	C	24.93	32.31	25.29	32.34
	D	23.69	32.54	34.38	32.65
	E	32.06	28.93	89.68	29.28

まず、 D_{total} の変動についてみると、図 7 から、 $W < 5$ のように W 値が極端に小さい場合に総遅延量 D_{total} が大きくなる傾向があり、特にビットストリーム D において増加している。これは、スライディングウィンドウ幅の符号量である $B_{in} \times W/F$ 値が、I ピクチャ 1 枚の符号量に対して十分大きくないため、最も符号量が大きい I ピクチャトランスコーディング処理後に、入力残符号量が極端に減少し、式 (2) により算出される $ioRatio(n)$ が大きく変動することに起因すると考えられる。ビットストリーム D は N 値、M 値ともに A, C, D 中最大であるため、B ピクチャ数が最も多く、そのため I ピクチャの符号量が A, C, D 中最大であり、結果として総遅延量が増加したものと考えられる。一方、 $W \geq 5$ の W 値についてみると、 W の変化に対して D_{total} の変動が小さいこ

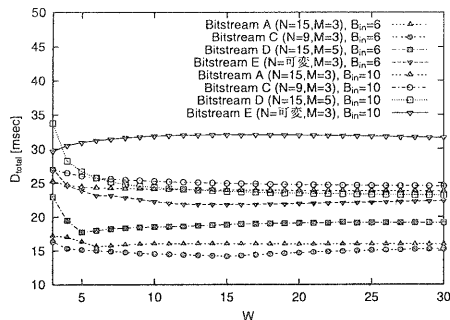


図 7: 各ビットストリームのトランスコード処理における W と D_{total} の関係

とから、 W の設定値に依存せず、安定した遅延量の下でトランスコード処理が実現できる。

また、各 N, M 値の相違についてみると、各ビットストリームの差は、 $B_{in} = 6$ [Mbps] で最大で約 5 [msec]、 $B_{in} = 10$ [Mbps] で約 2 [msec] であることから、 N, M 値の相違による遅延量の変動はほとんど無いものと判断して良い。さらに、TM 符号量制御適用方式と比較して、 $B_{in} = 6$ [Mbps] でビットストリーム A の場合においては約 3 [msec] 程度の増加が見られるが、それ以外はほぼ同等であり、 $B_{in} = 10$ [Mbps] でビットストリーム D の場合には約 10 [msec] 程度の遅延量減少が見られる。先にも述べたように、本実験の比較方式である TM 符号量制御適用方式は、GOP 構造情報を既知とする場合であるにも関わらず、その情報を既知としない提案方式は同等の遅延量である。よって、本提案方式が N, M 値及び W の設定値に大きく依存せず、TM 符号量制御適用方式と比較して、安定した低遅延量の下でトランスコード処理が実現できることが明らかとなった。

次にビットストリーム E の場合について考察すると、図 7 から W の変化に対する D_{total} の変動は小さく、ビットストリームの GOP 構造、 W の設定値に依存せず、安定した D_{total} の下でトランスコード処理が実現できる。また入力ビットレート B_{in} の相違についてみると、入力値が大きい場合には出力符号量の変動幅が大きくなるため、結果的に D_{total} が上昇する。そして、TM 符号量制御適用方式と比較した場合、図 7 及び表 5 から、いずれも提案方式の遅延量は小さく、特に方式 1、 $B_{in} = 10$ [Mbps] の場合においては、TM 符号量制御適用方式よりも約 57 [msec] 程度遅延量が低い。また、先にも述べたように TM 符号量制御適用方式はシーンチェンジ情報を既知とする場合であるため、プレスキューニングにより 1GOP 分の遅延量が発生する。以上から、提案方式では、シーンチェンジがあるシーケンスにおいても、 W の設定値に大きく依存せず、安定した低遅延量の下でトランスコード処理が実現できることが明らかとなった。

一方、各フレームの SN 比の変動を図 8 に示す。但し図 8 には、 $B_{in}=6$ [Mbps]、方式 1, 3, 5 方式のみの結果を示し、方式 3 は、シーンチェンジ情報を既知とする TM 符号量制御適用方式の場合を示した。図 8 から、シーンチェンジ情報を既知とする TM 符号量制御適用方式は、シーンチェンジ直前 GOP の画像の複雑さを示すパラメータ X_1, X_p, X_b に基づいて符号量制御を実現するため、シーンチェンジ直後の制御が正常に動作していない。例えば、フレーム番号 72 でのシーンチェンジ直後では、前 GOP の特性パラメータを使用しているため、シーンチェンジ直後の I ピクチャにおいて割り当て符号量が大きくなり、結果として GOP 終盤での符号量不足による SN 比の低下が現れており、それ以後の制御にも影響が及んでいる。一方、提案方式では、シーンチェンジ前後での急激な SN 比の向上等がなく、安定した性能を示した。

5.4 考察及び評価

5.1 ~ 5.3 から提案方式を総合評価する。提案符号量制御方式は、ループ処理 ON/OFF によるアーキテクチャの相違に関する画質性能は TM 符号量制御適用方式をトランスコーダに適用した場合と同様の傾向である。入出力ビットレート値比率の相違に関しては、入出力ビットレート比率が大きい場合、I ピクチャの SN 比低下が著しく、ループ処理 ON/OFF に関わらず TM 符号量制御適用方式と比較して 0.1 ~ 0.3 [dB] SN 比が低下するが、入出力ビットレート比率が小さい場合、TM 符号量制御適用方式と同等以上

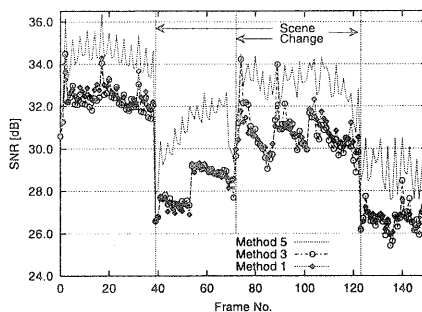


図 8: 各フレームにおける SN 比の変動

の SN 比である。以上から、提案符号量制御方式により提供される画質が、従来の TM 符号量制御適用方式をトランスコーダアーキテクチャ [A][B] に適用した場合とほぼ同等であることが明らかになった。

また、トランスコード遅延量については、TM 符号量制御適用方式にみられるような 1GOP 分の遅延や、従来方式に見られる長時間のビットストリーム統計に伴う遅延の発生はなく、送受信パuffers の滞留による遅延量を抑制でき、vbv_delay 値の増加を回避することで復号器での遅延発生を抑制可能である。以上から、低遅延トランスコード処理を実現できることが明らかになった。そして、リアクシオン係数 r の制御により総遅延量を抑制でき、それに伴う画質低下は 0.3 ~ 0.5 [dB] に抑えられることが示された。

最後に、GOP 構造の相違、シーンチェンジによる途中からの GOP 構造の変化、及び予め設定を行う残符号量時間幅に対しても、安定した符号量制御が実現可能であることが示され、提案符号量制御方式の遅延量変動、画質変動共に、入力ビットストリームの GOP 構造非依存であることが明らかになった。

以上から、提案符号量制御方式は、従来方式をトランスコーダに適用した場合と同等の画質を提供可能であり、入力ビットストリームに非依存で、且つ低遅延な符号量制御方式であることが明らかになった。

6. むすび

本稿では、低遅延 MPEG-2 ビデオトランスコーダ符号量制御方式を提案し、その有効性を明らかにした。本提案符号量制御を適用したアーキテクチャ [B] のトランスコーダは、安価に入手可能な PC (Pentium II 400 [MHz] 相当) において、リアルタイム処理 (1 秒あたりのトランスコード処理フレーム数が 29.97 [frame]) 以上を実現可能であることから、ネットワークノードへの実装による実用化への期待とその効果は大きいものと考えられる。現在、ATM ネットワークにより接続されるビデオサーバとクライアント PC との間に、本トランスコーダを実装したゲートウェイ PC を介在させ、リアルタイムトランスコード処理を行うシステムを実現している。

参考文献

- (1) D. Morrison: "Video Transcoders with Low Delay", IEICE Trans. Comms. (1997).
- (2) ITU-T, Recommendation H.261: "Video codec for audio-visual services at p x 64kbit/s" (1990).
- (3) ISO-IEC 13818-2, IS: "Generic Coding of Moving Picture and Associated Audio Recommendation H.262" (1996).
- (4) P. N. Tudor and O. H. Werner: "Real-Time Transcoding of MPEG-2 Video Bits Streams", International Broadcasting Convention, pp. 286-301 (1997).
- (5) ISO-IEC JTC1, MPEG 93/N225b: "Test Model 5 /draft" (1993).
- (6) 西村他: "リアルタイム MPEG-2 ビデオトランスコーダソフトウェアの開発", 情報研報, AVM25-5, pp. 25-30 (1999).
- (7) H. Sun, W. Kwok and J. W. Zdepski: "Architecture for MPEG Compressed Bitstream Scaling", IEEE Transaction on Circuits and Systems for Video Technology, Vol.6 No.2, pp. 191-199 (1996).