

ビットストリーム情報に基づく MPEG-2/H.264 トランスコーダ向け 可変ビットレート制御

A Bitstream Behavior Based Variable Bit Rate Control Algorithm for MPEG-2/H.264 Transcoder

瀧本 崇博[†]
Takahiro Takimoto

児玉 知也[†]
Tomoya Kodama

1. まえがき

地上デジタル放送などで用いられている MPEG-2 ストリームを H.264 ストリームへトランスコードする場合、可変ビットレート制御(Variable Bit Rate Control, 以下 VBR)を採用することで画質変動を抑えることができる。VBR において画質を安定させるためには、トランスコード対象となる画像信号がどのような挙動を示すかをトランスコード前に調べるのが重要である。そこで本稿では、トランスコード前に MPEG-2 ビットストリームを解析し、それを基に適応的にビットレート制御を行う手法を提案する。本手法により、原画像を 2 回符号化する既存の 2 パス VBR と比較して、同等の符号化性能をより少ない処理量で実現できることを確認した。

2. ビットストリーム情報に基づくビットレート制御

H.264 は、MPEG-2 と同様に動き補償と離散コサイン変換を用いたハイブリッド符号化を利用しており、その符号化効率は一般に MPEG-2 の約 2 倍であると言われている。よって、各フレームの MPEG-2 での符号化特性に基づいて符号量割り当てを行うことで、シーケンスを通じて適応的に符号量割り当てができると考えられる。しかしながら、MPEG-2 と H.264 では、画面内予測や動き補償の手法に違いがあるため、フレーム毎に符号化効率の向上度合いが変動する。したがって、MPEG-2 の符号化特性をそのまま用いてシーケンス内の各フレームに適応的に符号量割り当てを行うことは困難であると考えられる。

3. 提案手法

提案手法は、まず入力ストリームである MPEG-2 のヘッダ情報を解析することで、シーケンス内の全フレームに対して H.264 での符号化特性を推定する。推定によって生じる H.264 の符号化特性の実値との誤差を低減するために、各フレームの推定符号化特性を GOP 単位にまとめあげ、シーケンス全体に対し GOP 単位で符号量を割り当てる。また Picture 単位の符号量割り当ては、符号化時に動的に行う。提案手法の具体的な処理は Fig.1 のように 3 つの処理から構成されており、以下に各処理について説明する。

A. MPEG-2 解析

入力された MPEG-2 ストリームのヘッダ情報から、シーケンス内の各フレームのピクチャタイプと発生符号量 ($R_{mpeg2}(t)$) と平均量子化スケール ($Q_{mpeg2}(t)$) を算出する。なお、ピクチャタイプに関しては Picture header 内の

picture_coding_type から取得する。 $R_{mpeg2}(t)$ は、Picture header 間のビット量から算出し、 $Q_{mpeg2}(t)$ には、Slice header にある量子化ステップから推定する。

B. 符号量割り当て

B-1. 符号化特性(RO 関数)算出

A で取得した $R_{mpeg2}(t)$ と $Q_{mpeg2}(t)$ を用いて、(式1)に従い各フレームの MPEG-2 の符号化複雑度 ($X_{mpeg2}(t)$) を求める。

$$X_{mpeg2}(t) = R_{mpeg2}(t) \times Q_{mpeg2}^g(t) \quad (式1)$$

ただし、 $t=0,1,\dots,N$ 、 N はシーケンス内の総フレーム数であり、 g はピクチャタイプ毎に定まる固定値である。

次に、算出した $X_{mpeg2}(t)$ を用いてピクチャタイプ毎に用意した所定の近似式に従い各フレームの H.264 の符号化複雑度 ($X_{h264}(t)$) を推定する。

B-2. GOP 単位の符号量割り当て

B-1 の符号化特性算出で求めた各フレームの H.264 での推定符号化複雑度では 2 章で述べたようにフレーム単位では推定値と実値の誤差が大きいと考えられる。そこで、まず(式 2)を用いて各 GOP において、当該 GOP にある全フレームの H.264 推定符号化複雑度の総和を用いてシーケンス全体の平均ビットレートが所定の値となるように各 GOP に符号量を割り当てることとする。このようにすることで、中心極限定理により推定値と実値との誤差のばらつきを低減させることができる。

$$X_{gop(i)} = \sum_{k \in gop(i)} X_{h264}(k) \quad (式2)$$

ただし、 $i=0,1,\dots,M$ 、 M はシーケンス内の GOP 数を表している。

C. 符号化

ステップ B で求めた GOP 単位に割り当てた符号量を基に符号化を行う。符号化に際し、フレーム毎の目標符号量算出処理は次の 2 つのステップで構成される。

C-1. GOP 境界での処理

直前の GOP の符号化終了時での目標ビットレートとのずれと、B-2 で求めた当該 GOP の割り当て符号量を用いて、当該 GOP の割り当て符号量を算出する。

C-2. フレームの符号化時の処理

直前に符号化した各ピクチャタイプの符号化特性に合わせて、C-1 で算出した当該 GOP の目標符号量を GOP 内の未符号化フレームそれぞれに対して各フレームに比例配分し、これに基づき符号化を行う。

[†] 株式会社 東芝 Toshiba Corp.



Fig.1 提案手法のブロック図

4. 性能評価

提案手法の性能評価として、2パス VBR[1]および Test Model 5[2]に準拠した1パス CBRとの比較を行った。表1に入力画像として用いた MPEG-2 ストリームの符号化条件を、表2に H.264 の符号化条件を示す。Fig.2, Fig.3 に性能評価結果を示す。Fig.2 は GOP 毎の PSNR 変動を、Fig.3 には R-D 曲線を示し、各図において実線が 2パス VBR を、波線が 1パス CBR を、太実線が提案手法を表している。Fig.2 を見てわかるように提案手法は、2パス VBR に似た傾向を示していることがわかる。また 1パス CBR よりもビットレートが適応的に与えられることで、極端な PSNR の下落がなくなっていることがわかる。また、Fig.4 からわかるように、提案手法は 2パス VBR とほぼ同等の画質性能であることがわかる。また、表3に 1パス CBR を基準(=1)とした際の、処理時間増加率を示す。表3を見てわかるように提案手法は 2パス VBR と比較して約半分の処理量であることがわかる。

表1 入力 MPEG-2ストリームの符号化条件

フレーム数[frame]	約 8700
フレームレート[fps]	29.97
平均ビットレート[Mbps]	16.0
GOP 内のフレーム数[frame]	15
予測構造	IBBPBBPBBP.....
解像度	1440x1080

表2 符号化条件

Profile	High Profile
Level	4.1
平均ビットレート[Mbps]	8.0
最大ビットレート[Mbps]	20.0
最小ビットレート[Mbps]	7.2
GOP 構造	入力に従う
予測構造	
出力形式	H.264
解像度	1440x1080
エントロピー符号化	CABAC

表3 各手法における処理増加率

1パス CBR	2パス VBR	提案手法
1.00	2.10	1.05

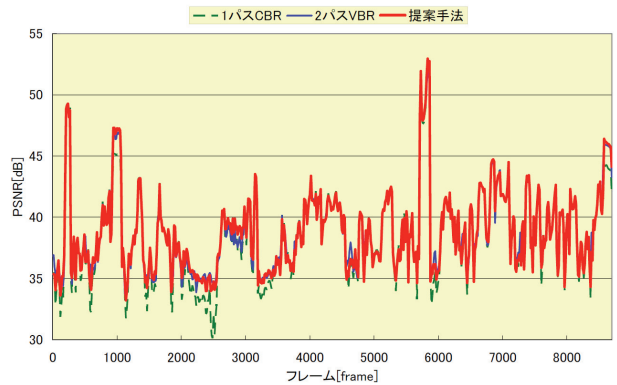


Fig.2 GOP 毎の平均 PSNR 変動

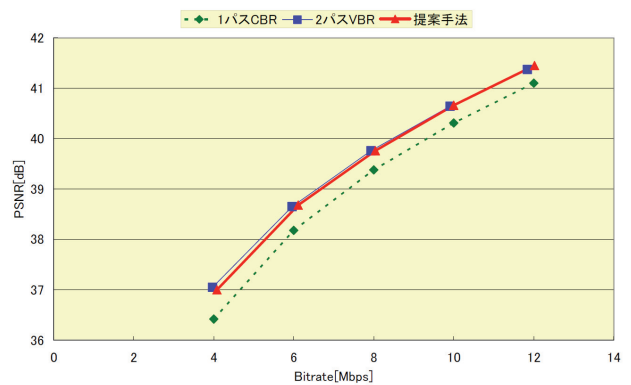


Fig.3 各手法における R-D 曲線

5. まとめ

本稿では、MPEG-2 ストリームを H.264 ストリームへトランスコードするためのビットレート制御手法として、MPEG-2 ストリーム情報に基づく VBR を提案した。本手法は、MPEG-2 ストリームをフルデコードせずにヘッダ情報を利用することで H.264 の符号化特性をフレーム毎に推定し、それを GOP 毎に累計した値を用いて、符号量割り当てを行う。これにより推定による誤差の影響を低減し、シーケンスを通じて適応的な符号量割り当てを実現した。本手法の性能評価を行うため、従来の 2 回符号化を行う 2パス VBR と比較した結果、同等の符号化性能をより少ない処理量で実現できることを確認した。

参考文献

[1] 古藤, 尾高, "MPEG-2 MP@ML 準拠ビデオエンコーダ-パス可変ビットレート符号化-", 信学会総合大会, pp.70, 1995
 [2] ISO-IEC/JTC1/SC29/WG11/N0400, Test Model 5, April, 1993