

H.264/AVCにおける符号量制御に関する考察

秋間 佑輔[†] 永吉 功^{††} 花村 剛^{††} 富永 英義[†]

[†] 早稲田大学大学院 国際情報通信研究科 〒 367-0035 埼玉県本庄市西富田大久保山 1011

^{††} 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 新宿ラムダックスビル 8階

E-mail: [†]{aky,tominaga}@tom.comm.waseda.ac.jp, ^{††}{isao,hana}@mediaglue.co.jp

あらまし 本稿では H.264/AVC にはおける符号量制御方式について議論する。参照ソフトウェアによる従来方式では入力画像先頭フレームは符号量制御されず、その発生符号量による影響が後続数十フレームに影響を与え、目標ビットレートへの収束性が不十分となることを指摘する。そこで本稿では、割り当て符号量を入力画像の先頭フレームから考慮することによって量子化パラメータを制御する方式を提案する。先頭フレームの割り当て符号量は符号化複雑度の初期値を設定することによって算出する。予備実験で発生符号量と量子化パラメータの特性を測定し、モデル化することで割り当て符号量から量子化パラメータを制御する。出力ビットレートと画質の観点で従来方式との比較評価実験を行うことで提案方式の妥当性を述べる。

キーワード H.264/AVC, 符号量制御, 割り当て符号量, 量子化パラメータ

A Study on Rate Control for H.264/AVC

Yusuke AKIMA[†], Isao NAGAYOSHI^{††}, Tsuyoshi HANAMURA^{††}, and Hideyoshi

TOMINAGA[†]

[†] Graduate School of GITS, WASEDA University, Okuboyama 1011, Nishitomida, Honjo-shi, Saitama, 367-0035 JAPAN

^{††} Media Glue Corporation, Ramdax Bldg, 8th floor, 2-4-12 Okubo, Shinjuku-ku, Tokyo, 169-0072 JAPAN

E-mail: [†]{aky,tominaga}@tom.comm.waseda.ac.jp, ^{††}{isao,hana}@mediaglue.co.jp

Abstract Rate control scheme is a very important technology of video transmission, especially in CBR. H.264/AVC standard as a joint development of ITU-T and ISO-MPEG supports rate control. In H.264/AVC reference software JM, initial I and P frame of video sequence can not be encoded with rate control. In this paper, we propose QP(quantization parameter) control scheme of initial I and P frame. We compute target bits of initial I and P frame by setting initial value of coding complexity, and decide the value of QP using the relation between actual bits and QP. We evaluate the proposed method by the value of actual bitrate and PSNR. Our method provide better results than JM.

Key words H.264/AVC, rate control, target bits, quantization parameter

1. はじめに

本稿では H.264/AVC [1] における符号量制御方式について議論する。その概要は、目標ビットレートから入力画像の各 GOP 内の各フレームに対する割り当て符号量の算出、その割り当て符号量を基にしたフレーム内各 MB の量子化パラメータ (QP 値) 制御の 2 段階により構成される。

H.264/AVC では、MB ごとに符号化モードを切り換えることができ、符号化ブロックごとのヘッダ情報量やテキスト符号量の変動が激しい。そのため符号量の割り当てがフレーム精度では不十分となる [2]。そこで、文献 [2] では、符号化フレーム内に任意の複数 MB 群として Basic Unit (以下、BU と呼ぶ) を定義し、BU 単位で符号量の割り当てを行っている。過去の発生符号量や割り当て符号量による R-D (Rate-Distortion) モデルの更新に伴い QP を算出する。

この BU レイヤの追加により厳密な符号量の制御が実現されているが、目標ビットレートへの収束性が、未だ不十分であることを予備実験により確認する。この問題は、入力画像の先頭フレームが符号量制御の対象とならないことが、後続フレームに影響することに起因する。特に、変化の激しいシーケンスやシーンチェンジが発生するシーケンスにおいて顕著に現れる。

そこで、本稿では、この問題を解消するための初期検討として割り当て符号量を入力画像の先頭フレームから考慮することによって QP を制御する方式を提案する。先頭フレームの割り当て符号量は符号化複雑度の初期値を設定することにより算出する。出力ビットレートと画質の観点から H.264/AVC の符号化参照ソフトウェア JM [4] との比較評価実験を行い、提案方式の妥当性を示す。

2. 従来方式

H.264/AVC の符号量制御方式には、従来の標準化方式 [3] より用いられている GOP レイヤ、フレームレイヤ、MB レイヤの演算に加えて BU レイヤの演算が定義されている [2]。BU は N_{bu} 個の MB で構成される複数の MB の集合体として定義され、任意に設定可能である。ただし、符号化フレームが N_{pic} 個の MB で構成されるとき、 N_{pic}/N_{bu} は整数値でなければならない。BU レイヤを加えた H.264/AVC の符号量制御方式は次の 2 段階で構成される。

Step1 各レイヤでの割り当て符号量算出

Step2 BU レイヤでの QP 制御

以下に詳細を示す。

2.1 割り当て符号量算出

まず、GOP レイヤで符号化 GOP の割り当て符号量 T を算出する。次に、フレームレイヤで GOP 内 i 番目フレームの割り当て符号量 T_i を式 (1) のように算出する。

$$T_i = \frac{T - S_{i-1}}{N_{GOP} - i + 1}, \quad 1 \leq i \leq N_{GOP} \quad (1)$$

ここで S_{i-1} 、 N_{GOP} は、それぞれ $i-1$ 番目フレームまでの発生符号量の総和、GOP 内フレーム数を示す。

式 (1) によって算出された T_i を基に i 番目フレーム内 j 番目 BU の割り当て符号量 $T_{i,j}$ は式 (2) のように算出される。

$$T_{i,j} = \frac{T_i - S_{i,j-1}}{N_{unit} - j + 1}, \quad 1 \leq j \leq N_{unit} \quad (2)$$

ここで $S_{i,j-1}$ 、 N_{unit} は、それぞれ $j-1$ 番目 BU までの発生符号量の総和、フレーム内 BU 数を示す。

2.2 BU レイヤにおける QP 制御

2 次 R-D モデルから式 (3) に示す R-Q (Rate-QStep) モデルを導く。

$$T_{i,j} = a \times \frac{MAD_{cb}}{Q_{step}} + b \times \frac{MAD_{cb}^2}{Q_{step}^2} \quad (3)$$

ここで、 Q_{step} は量子化ステップを示す。また、 MAD_{cb} は符号化 BU の $MAD^{(cb)}$ を示し、直前フレームの同位置 BU における MAD である MAD_{pb} を用いて式 (4) のように予測される。

$$MAD_{cb} = c_1 \times MAD_{pb} + c_2 \quad (4)$$

ここで c_1 、 c_2 、および式 (3) の a 、 b は、それぞれ定数として各 BU で定義される。

ただし、入力画像先頭フレームでは参照する過去の情報がないため QP の制御ができない。そこで、先頭 I、P フレームにおいては、式 (5) のように 1 画素あたりの符号量 bpp [bit/pixel] を用いた簡単な制御を行う。

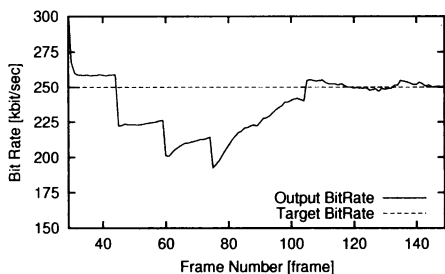
$$QP_i(1) = \begin{cases} 35 & bpp \leq l1 \\ 25 & l1 < bpp \leq l2 \\ 20 & l2 < bpp \leq l3 \\ 10 & bpp > l3 \end{cases} \quad (5)$$

ただし、 $bpp = T_r / (F_r \times N_{pixel})$ である。ここで F_r 、 N_{pixel} は、それぞれフレームレート、フレーム内画素数を示す。また、 l は定数であり、画像サイズが QCIF であれば $l1 = 0.1$ 、 $l2 = 0.3$ 、 $l3 = 0.6$ 、CIF であれば $l1 = 0.2$ 、 $l2 = 0.6$ 、 $l3 = 1.2$ 、CIF より大きいサイズでは $l1 = 0.6$ 、 $l2 = 1.4$ 、 $l3 = 2.4$ である。

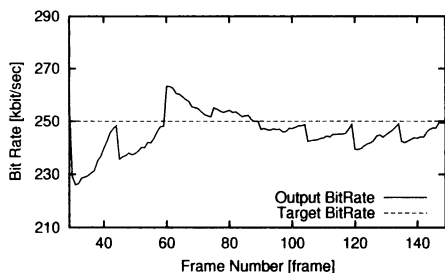
2.3 課題点

従来方式による出力ビットレートの調査結果を図 1 に示す。図 1 は、目標ビットレートと出力ビットレートの関係を示したものである。テストシーケンスとして、CIF サイズの Mobile & Calendar と Tempete を使用し、目標ビット

(注1) : Mean Absolute Difference : BU 内の各画素値の絶対差分和を表し符号化複雑度として用いられる



(a) Mobile & Calendar



(b) Tempete

図 1 目標ビットレートと出力ビットレートの関係

レートと 250[kbit/sec] として評価している。図 1 より 80~100[frame] ほど経過するまでは目標ビットレートへ収束しないことを確認できる。

符号量制御は QP 値の予測算出によって行われることから符号化フレームでは符号化済みフレームの QP 値を参照する。つまり、目標ビットレートへの十分な収束性を実現するためには、先頭フレームの適節な QP 制御が必要となる。

3. 提案方式

本稿では、この問題を解決するための初期検討として先頭フレームの割り当て符号量を考慮した QP を制御する方式を提案する。

符号量制御は 2Step で実行される。本稿では、Step1 にて先頭フレームの割り当て符号量を算出し、Step2 では QP の制御を行う。以下に詳細を示す。

3.1 先頭フレームの割り当て符号量

本稿では、MPEG-2 TM5 方式 [5] における符号化複雑度 X を用いた初期化演算によって先頭の I, P フレームの割り当て符号量 T_{initI} , T_{initP} を式 (6) のように算出する。

$$\begin{aligned} T_{initI} &= \frac{T}{1 + N_p X_p / X_i} \\ T_{initP} &= \frac{T - T_{initI}}{N_p} \end{aligned} \quad (6)$$

ここで T , N_p は、それぞれ先頭 GOP の割り当て符号量、GOP 内未符号化 P フレーム数を示す。また X_i , X_p は I, P

表 1 符号化条件

JM version	JM8.5
Profile	Baseline
Test Sequence	Mobile & Calendar Tempete
Number of frames	150 [frame]
GOP Structure	N=15, M=1
Frame Rate	30[frame/sec]
Target Bit Rate	250 [kbit/sec]
Basic Unit	22 [MB]

フレームの符号化複雑度を示し、初期値は目標ビットレート T_R [bit/sec] を用いて以下のように設定する。

$$\begin{aligned} X_i &= 160 \times T_R / 115 \\ X_p &= 60 \times T_R / 115 \end{aligned} \quad (7)$$

後続の各フレーム及び各 BU に対する割り当て符号量算出は式 (1)(2) に従う。

3.2 QP 制御方式

算出された割り当て符号量を基に、発生符号量と QP の関係を用いて各 BU の QP を制御する。また、本制御方式は I フレームには適用せず、フレーム内で固定 QP を用いる。以降の記述では P フレームに焦点を当てる。

図 2 に予備実験から求めた各入力画像における P フレームの発生符号量と QP の関係を示す。符号化条件は表 1 に従う。

図 2 の特性から式 (8) の近似が成立する。算出された i 番目フレーム内 j 番目 BU の割り当て符号量 $T_{i,j}$ を入力パラメータとして QP 値 $QP_{i,j}$ を式 (8) によって算出する。

$$QP_{i,j} = f^{-1}(T_{i,j}) = a \times T_{i,j} + b \quad (8)$$

ここで a , b は定数とし、各入力画像において実験的に求める。その値を表 2 にまとめる。

各 BU で $T_{i,j}$ が算出されたら表 2 をもとに定数 a , b を定めることで各 P フレーム内各 BU の QP 値を制御する。

ただし、先頭 GOP を含めた各 GOP 内 I フレームでは、各 GOP 内先頭 P フレームと均一の QP 値を用いて制御する。動き補償による P フレームの符号量や画質への影響を考慮し、均一の QP 値を用いることが妥当な結果を生み出すという考察に至ったためである。

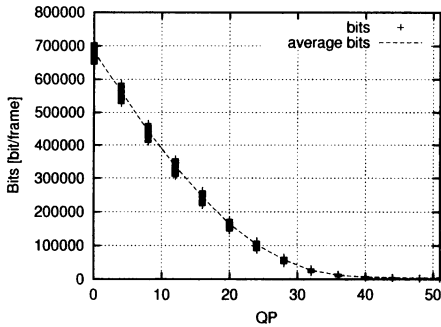
表 2 に示す定数条件に関する検討が必要であるが、本稿では初期検討として表 2 の値を用いる。

4. シミュレーション実験

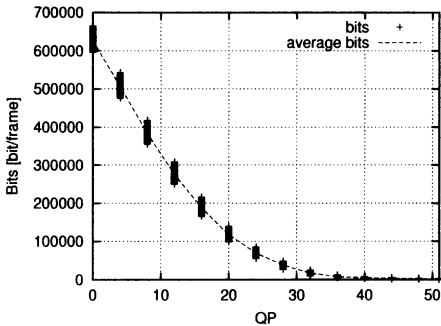
提案方式の妥当性を確認するためにシミュレーション実験を行う。本実験では、従来方式として JM を用いて各フレームの出力ビットレートと SNR 値を比較する。符号化条件は表 1 に示す通りである。出力ビットレートの特性を図 3 に、画質の特性を図 4 にそれぞれ示す。

表 2 割り当て符号量と定数 a , b の関係

Mobile & Calendar			Tempete			The value of QP
Target bits[bit]	a	b	Target bits[bit]	a	b	
$442196 \leq T_{i,j}$	-3.7×10^{-5}	22.8	$382911 \leq T_{i,j}$	-3.3×10^{-5}	20.6	$0 \leq QP_{i,j} < 8$
$335572 \leq T_{i,j} < 442196$	-3.7×10^{-5}	24.6	$274970 \leq T_{i,j} < 382911$	-3.7×10^{-5}	22.2	$8 \leq QP_{i,j} < 12$
$244266 \leq T_{i,j} < 335572$	-4.4×10^{-5}	26.7	$188159 \leq T_{i,j} < 274970$	-4.6×10^{-5}	24.7	$12 \leq QP_{i,j} < 16$
$163821 \leq T_{i,j} < 244266$	-5.0×10^{-5}	28.1	$117841 \leq T_{i,j} < 188159$	-5.7×10^{-5}	26.7	$16 \leq QP_{i,j} < 20$
$102974 \leq T_{i,j} < 163821$	-6.6×10^{-5}	30.7	$69661 \leq T_{i,j} < 117841$	-8.3×10^{-5}	29.8	$20 \leq QP_{i,j} < 24$
$58599 \leq T_{i,j} < 102974$	-9.0×10^{-5}	33.3	$37813 \leq T_{i,j} < 69661$	-1.1×10^{-4}	32.7	$24 \leq QP_{i,j} < 28$
$27654 \leq T_{i,j} < 58599$	-1.2×10^{-4}	35.5	$17504 \leq T_{i,j} < 37813$	-2.0×10^{-4}	35.4	$28 \leq QP_{i,j} < 32$
$12465 \leq T_{i,j} < 27654$	-2.6×10^{-4}	39.3	$8281 \leq T_{i,j} < 17504$	-4.3×10^{-4}	39.6	$32 \leq QP_{i,j} < 36$
$6644 \leq T_{i,j} < 12465$	-6.8×10^{-4}	44.6	$4584 \leq T_{i,j} < 8281$	-1.1×10^{-3}	45.0	$36 \leq QP_{i,j} < 40$
$4391 \leq T_{i,j} < 6644$	-1.8×10^{-3}	51.8	$2853 \leq T_{i,j} < 4584$	-2.3×10^{-3}	50.6	$40 \leq QP_{i,j} < 44$
$3167 \leq T_{i,j} < 4391$	-3.2×10^{-3}	58.4	$1743 \leq T_{i,j} < 2853$	-3.6×10^{-3}	54.3	$44 \leq QP_{i,j} < 48$
$T_{i,j} < 3167$	-8.2×10^{-3}	76.0	$T_{i,j} < 1743$	-9.9×10^{-3}	65.2	$48 \leq QP_{i,j} < 51$



(a) Mobile & Calendar



(b) Tempete

図 2 QP と発生符号量の関係

図 3 より提案方式では先頭フレームから割り当て符号量と QP の関係を考慮していることで適応的に QP の制御が行えていることから目標ビットレートへの収束性が大幅に改善されていることを確認できる。

図 4 より画質においても同様の改善を確認できるが、SNR 値が安定した以降で従来方式よりも SNR 値が下がっている箇所があり、今後の解決が必要となる。この課題点に対する

表 3 出力結果および標準偏差 σ

Method	Mobile & Calendar		Tempete	
	JM	Proposed	JM	Proposed
Bit Rate [kbit/sec]	251.06	249.39	250.45	249.02
Y SNR [dB]	23.49	23.5	27.27	27.21
σ_{rate}	21.30	3.948	7.688	4.465
σ_{SNR}	1.299	0.5040	0.8718	0.5827

考察として、従来方式と提案方式における QP の制御方式の違いが挙げられる。従来方式では、画質の変動を抑制するために QP 値の大幅な変動を抑制する。そのため図 4 から確認できるように各 GOP の先頭 I フレームから SNR 値の急激な減少が抑えられている。一方、提案方式では、割り当て符号量の制約より厳密に QP の制御を行っているため I フレームと P フレームの SNR 値に大きな差が生じている。

さらに、シーケンス全体の出力変動を評価するために出力ビットレート、SNR 値の標準偏差 σ_{rate} , σ_{SNR} を用いる [6]。出力ビットレート、平均 SNR 値と合わせて結果を表 3 に示す。ただし、 σ_{rate} は以下のように定義する。

$$\sigma_{rate} = \sqrt{\frac{1}{N} \sum_{k=1}^N (S_k - T_R)^2} \quad (9)$$

ここで N , S_k はそれぞれシーケンスの入力フレーム数、 k 番目フレームにおける出力ビットレートを示す。また、 T_R は目標ビットレートを示す。式 (9) を用いることで、目標ビットレートに対する出力ビットレートの変動値をパラメータ化することができる。 σ_{SNR} では SNR の平均値に対する変動値を示す。

表 3 より提案方式によって σ_{rate} では最大約 80%, σ_{sn} では最大約 61%削減された。

以上より、目標ビットレートへの収束性、また出力 SNR 値の観点から本提案方式の妥当性が示された。

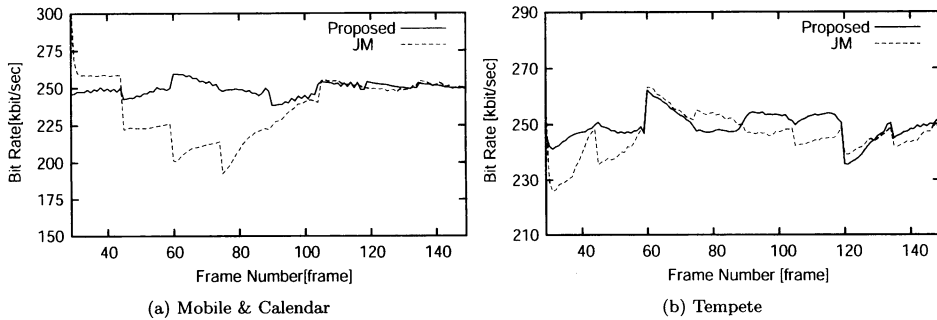


図3 出力ビットレート特性

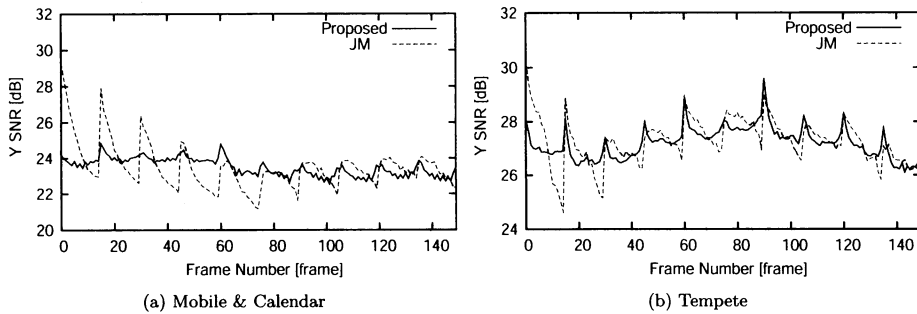


図4 画質特性

5. まとめ

本稿では、H.264/AVCにおける符号量制御方式に関する考察として議論した。従来方式として参照ソフトウェアによる方式の概要を述べ、目標ビットレートへの収束性に関する課題点を予備実験により示した。この問題を解消する初期検討として先頭フレームより割り当て符号量を考慮することによって、厳密にQPを制御する方式を提案した。実験の結果、出力ビットレートの収束性、画質変動の観点から提案方式の妥当性を確認できた。

本稿では議論できなかったシーケンスに依存する定数条件の解析、Iフレームにおける制御方式の検討が今後の課題といえる。

文 献

- [1] ISO/IEC 14496-10, "Advanced Video Coding," ITU-T Recommendation H.264, 2003.
- [2] Z. Li, F. Pan, K.P. Lim, G. Feng, X. Lin, S. Rahardja, "Adaptive Basic Unit Layer Rate Control for JVT," JVT-G012-r1, JVT 7th Meeting, Pattaya, Mar2003.
- [3] J.R. Corbera, S. Lei, "Rate Control in DCT Video Coding for Low-Delay Communications," *IEEE Trans. Circ. and Syst. Video Tech.* VOL.9, NO.1, Feb1999.
- [4] Joint Video Team(JVT), "Reference Software," <http://iphome.hhi.de/suehring/tml/>
- [5] MPEG-2 Test Model 5, Rev. 2, Section 10: Rate Control and Quantization Optimization, ISO/IEC/JTC1SC29WG11,

April 1993.

- [6] X. Yi, N. Ling, "Rate Control Using Enhanced Frame Complexity Measure for H.264 Video," *IEEE Workshop on SiPS*, pp.263-268, Austin, Texas, USA, Oct2004.