

J-072

H.264 符号化における画質変動抑制を目的とした レート歪み予測型符号量制御方式

Appropriate Rate Control Scheme to Minimize Temporal Image Quality Fluctuations Based on Rate Distortion Approximation Effective for MPEG-4 AVC/H.264 Encoding.

松村 篤志[†]
Atsushi Matsumura

内藤 整[†]
Sei Naito

川田 亮一[†]
Ryoichi Kawada

小池 淳[†]
Atsushi Koike

1. はじめに

動画像標準符号化方式の1つである H.264 は、携帯端末向け地上波デジタル放送や ADSL ストリーミングなどの低ビットレート制限下での適用が期待されている。しかしながら、同方式における符号量制御を隣接ピクチャの画質を考慮せずに行った場合には、隣接ピクチャにおける画質の急激な変動（以下単に画質変動と表記）に起因し、フリッカによる主観画質の劣化を招く危険がある。そこで筆者らは、H.264 符号化時における画質変動の抑制を目的とした符号量制御方式を提案し、シミュレーション実験によってその有効性を確認した [1]。本稿では、同方式において用いられる符号量と歪みの関係を表す近似式の高精度化によって、性能改善を実現するための検討を行ったので、その結果を報告する。

2. 画質変動抑制型符号量制御方式

文献 [1] にて提案されている符号量制御方式のうち、参照ピクチャにおける量子化パラメータ（以下 QP と表記）決定手順を以下に示す。まず、式 (1) を用いて当該ピクチャ i における歪み量 $D(i)$ を決定する。ただし、本稿における歪み量とは、各画素における平均二乗誤差の合計を表す。

$$\begin{aligned} R_I(D(i)) + H(I) + \\ N_P \times (R_P(D(i)) + H(P)) + \\ N_B \times (R_B(D(i)/w_b) + H(B)) = R_G \end{aligned} \quad (1)$$

ここで、各係数の定義は表 1 に従うこととする。その後、 $D(i)$ に最も近い歪み量となる QP を当該ピクチャにおける QP として選択する。

また、非参照ピクチャにおいては、以下の歪み量 $D(i)$ を与える QP を選択する。

$$\min |D(i) - D'(i)/w_b| \quad (2)$$

ただし、 $D'(i)$ は前後の参照ピクチャの平均歪み量を表す。

3. レート歪み近似式の高精度化

文献 [1] の方式では、 $R_I(D(i))$ 、 $R_P(D(i))$ および $R_B(D(i)/w_b)$ を符号量と歪みの関係を表す近似式（以下レート歪み近似式と表記）によって算出している。そこで本章では、レート歪み近似式の精度についての検討を行った上で、より高精度な近似式を提案する。

これまで、いくつかの文献においてレート歪み近似式が提案されている。例えば、文献 [2] では式 (3) が提案されている。

$$R(i) = a(i) \times D(i)^{b(i)} \quad (3)$$

[†] (株) KDDI 研究所, KDDI R&D Inc.

表 1: 式 (1) における係数

係数	定義
$R_I(D(i)), R_P(D(i))$	I, P ピクチャにおいて、歪み量が $D(i)$ となる際の符号量
$R_B\left(\frac{D(i)}{w_b}\right)$	B ピクチャにおいて歪み量が $D(i)/w_b$ となる際の符号量
N_I, N_P, N_B	1GOP 内に含まれる I, P, B ピクチャの枚数
$H(I), H(P), H(B)$	I, P, B ピクチャにおけるサイド情報符号量
w_b	I, P ピクチャの歪み量 $D(Ref)$ と B ピクチャの歪み量 $D(B)$ の比率 $(D(Ref)/D(B))$
R_G	Coded Picture Buffer に基づいて決定される 1GOP 長の目標符号量

ただし、 $R(i)$ はピクチャ番号 i に対する DCT 係数部の符号量を表し、 $a(i)$ および $b(i)$ はピクチャ毎に決定される係数を表す。また、文献 [3] や文献 [4] では、式 (3) に対して $b = -1$ ないし $b = -1/\sqrt{2}$ を代入した式が用いられている。なお、文献 [1] では、式 (3) を用いて符号量を予測している。

式 (3) に対する精度を検証するため、Cognac and Fruit および Whale Show を H.264 参照ソフトである JM9.6[5] を用いて符号化した。B ピクチャのレート歪み曲線を対象とし、実測結果と式 (3) による近似結果を図 1 に示す。 $a(i)$ および $b(i)$ は、実測値をもとに最小二乗法を用いて算出した。結果より、Cognac and Fruit においては高精度な近似がなされている一方、Whale Show においては形状の類似性は見られるものの、値が大きく乖離していることがわかる。他のピクチャタイプにおいても同様の傾向が見られた。これらの結果より、式 (3) の近似精度は素材に大きく依存すると考えられる。

そこで本稿では、式 (3) における形状の類似性に着目し、これに対し係数 $c(i)$ を加えた以下のレート歪み近似式を提案する。

$$R(i) = a(i) \times D^{b(i)} + c(i) \quad (4)$$

式 (4) の近似精度を検証するため、式 (3) および (4) により求められた予測符号量と実測値の誤差の割合を表 2 に示す。ただし、表における誤差の単位は%とする。結果より、いずれのピクチャタイプにおいても式 (4) の適用によって、より高精度な近似がなされていることがわかる。式 (3) および (4) における平均誤差は、それぞれ 12.01% および 3.89% となり、また、最悪値の比較では、式 (4) を用いることによって約 25% の改善が見られた。

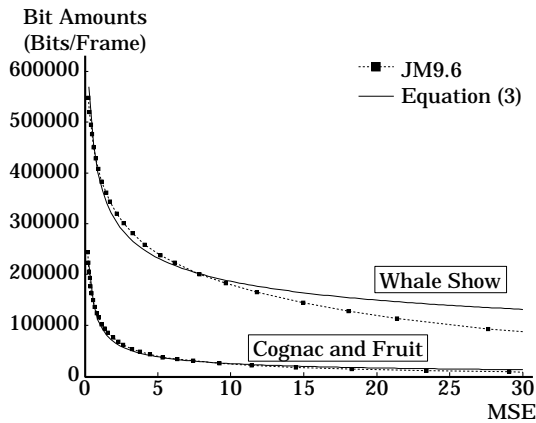


図 1: B ピクチャにおける符号量と歪み量

表 2: 各符号量に対する誤差

素材名	ピクチャタイプ	式 (3)	式 (4)
Cognac and Fruit	I ピクチャ	6.47	1.87
	P ピクチャ	8.40	3.22
	B ピクチャ	7.14	4.17
Flower Basket	I ピクチャ	9.94	5.45
	P ピクチャ	11.06	6.19
	B ピクチャ	8.28	7.86
Whale Show	I ピクチャ	10.15	3.35
	P ピクチャ	14.92	3.29
	B ピクチャ	32.62	5.95
Green Leaves	I ピクチャ	5.23	1.47
	P ピクチャ	12.10	1.52
	B ピクチャ	18.11	2.35

表 3: 符号化シミュレーション条件

設定項目	設定値
w_b	0.9
使用プロファイル	Main Profile
符号化レート	500kbps
GOP 構成	IBBPBBPBBPBBPBB
符号化方式	CABAC
参照ピクチャ数	2

表 4: MSE 分散の結果

素材名	式 (3)	式 (4)
Cognac and Fruit	21.09	16.71
Flower Basket	34.56	32.24
Woman with Bird Cage	31.32	29.69
European Market	1300.21	959.78
Street Car	85.15	84.91
Whale Show	5916.51	5907.94
Opening Celemony	318.90	187.47
Green Leaves	11856.01	2587.82
Crowded Crosswalk	335.62	186.86
Ice Hockey	49.29	39.30

の結果、従来提案されている近似式に対して定数項を追加することによって、同式の高精度化を実現させた。また、シミュレーション実験より、上記符号量制御方式に提案した近似式を適用することによって、同方式の性能が改善されることを確認した。

なお、本研究は独立行政法人情報通信研究機構による委託研究「ソフトウェア符号化技術に関する研究開発」として実施したものである。

4. シミュレーション実験

提案したレート歪み近似式を文献 [1] の方式に適用させた際の性能を検証するため、シミュレーション実験を行った。実験では、映像情報メディア学会 SIF 標準画像より 10 素材を用い、先頭ピクチャから 301 ピクチャの符号化を行った。符号化シミュレーション条件を表 3 に示す。初期 QP は JM9.6 の初期値に従い 35 とした。また、QP の設定はピクチャ単位でのみ行い、サイド情報符号量に関しては、直前の同一ピクチャタイプの符号量を代入することとした。

MSE 分散の結果を表 4 に示す。MSE 分散が小さいほど、画質の変動が小さく、主観画質の改善に寄与するものと期待される。結果より、いずれの素材においても式 (3) を用いた場合と比べ、式 (4) を用いることによって、MSE 分散が減少していることがわかる。また、特に Green Leaves や European Market などの MSE 分散の減少が顕著に現れた素材において、式 (4) の適用により、フリッカが抑止され、主観画質が改善されたことを目視によって確認した。

5. まとめ

本稿では、H.264 符号化における画質変動抑制を目的とした符号量制御方式に対し、レート歪み近似式の高精度化によって性能を改善させるための検討を行った。そ

参考文献

- [1] 松村 篤志, 内藤 整, 川田 亮一, 小池 淳, “MPEG-4 AVC/H.264 における画質変動の抑制を目的とした量子化パラメータ選択方式”, 電子情報通信学会 2005 年総合大会, pp. D-11-59, Mar. 2005.
- [2] Yucel Altunbasak and Nejat Kamaci, “An analysis of the DCT coefficient distribution with the H.264 video coder,” IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004, vol.3, pp.iii-177-180, May 2004.
- [3] S. Takamura, H. Watanabe, and N. Kobayashi, “One-pass VBR algorithm for an MPEG-2 encoder chip SuperENC,” Picture Coding Symposium Japan, pp. 39-40, Sep. 1999.
- [4] P. Carlsson, P. Pan, C. Liang-Tien, “Coefficient thresholding and optimized selection of the lagrangian multiplier for non-reference frames in H.264 video coding” ICIP '04. Vol. 2, pp.773 - 776, Oct. 2004.
- [5] Karsten Sühring, H.264/AVC Software Coordination <http://bs.hhi.de/~suehring/tml/download/>