

フィルムグレインを考慮した H.264/AVC 符号化方式に関する一検討

後藤 崇行[†] 常松 祐一[†] 渡辺 裕[†]

[†] 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

E-mail: †{goto,tune}@tom.comm.waseda.ac.jp, ††hiroshi.watanabe@waseda.jp

あらまし フィルムグレインは、フィルムで撮影したコンテンツ特有の様相である。一般に雑音成分として捉えられがちだが、フィルムグレインによってフィルムコンテンツの質感を表現することができるため、質感を表現したい場合、除去されないことが望ましい。しかし、フィルムグレインを含んだ画像に圧縮符号化を施すと符号化による影響を受け失われてしまう。特に近年標準化された動画像符号化方式 H.264/AVC においては、デブロッキングフィルタや 4x4 整数変換により、フィルムグレインが失われやすくなっている。H.264/AVC の拡張方式 FRExt (Fidelity Range Extensions) では、映画コンテンツなどフィルムグレインを含む高解像度画像において画質改善を行う様々な符号化ツールが追加された。そこで筆者らは、更なるフィルムグレインの再現性を考慮した H.264/AVC 符号化方式について検討してきた。本稿では、符号化モード決定のコスト値について考察し、フィルムグレインの再現性を更に向上させる手法の検討を行う。

キーワード H.264/AVC, フィルムグレイン, モード決定

A Study on Coding Method in H.264/AVC Considering Film Grain

Takayuki GOTO[†], Yuichi TSUNEMATSU[†], and Hiroshi WATANABE[†]

[†] Graduate School of Global Information of Telecommunication Studies, Waseda University.

1011 Okuboyama Nishi-Tomida Honjo-shi Saitama 367-0035 Japan

E-mail: †{goto,tune}@tom.comm.waseda.ac.jp, ††hiroshi.watanabe@waseda.jp

Abstract Film grain is a specific texture of film contents. Though it generally tends to be recognized as a noise, it is preferable for film grain not to be removed to express the feeling of quality of film contents. However, film grain is influenced and lost easily by encoding. Especially, in H.264/AVC which is the state-of-the-art video coding standard, it is more easier for film grain to be lost by performing de-blocking filter and 4x4 integer-transform. In FRExt (Fidelity Range Extensions) which extended the conventional H.264/AVC, various coding tools were added to improve the quality of high definition size images such as movies containing film grain. We have been studying the H.264/AVC encoding method considering the fidelity of film grain. In this paper we consider the method to improve the fidelity of the film grain further by changing the cost value of the encoding mode decision.

Key words H.264/AVC, Film Grain, Mode Decision

1. はじめに

近年、蓄積分野において、ハードディスクドライブや次世代光ディスクに代表されるような、大容量化・低価格化がなされてきている。また、家庭用デジタル機器において、テレビやプロジェクトに代表されるように、高解像度 (HD) への対応が進んできている。これにより、一般家庭での高解像度・高画質な映像の視聴環境が整いつつある。一方、MPEG-1/2/4 や H.261, H.263, H.264/AVC [1] といった映像符号化技術が発達してき

ている。こうした背景を踏まえ、高解像度・高画質の映画コンテンツの流通が盛んになると予想される。高解像度・高画質の映画コンテンツにおいては、映画に特有な映像表現に対する再現性の要求が高まると考えられ、映画特有の映像表現には、フィルムグレインや色の再現性、ガンマ特性といったものがある。この中のフィルムグレインは、フィルムで撮影したコンテンツ特有の様相であり、映像をフィルムに記録し、現像する際の化学反応によって生じるものである。一般に雑音成分として捉えられがちであるが、フィルムグレインによってフィルムコンテン

ツの質感を表現することができるため、質感を表現したい場合、除去されないことが望ましい。現在、映像を配布する場合には、非圧縮状態ではデータ量が大きすぎるため圧縮符号化処理は必要不可欠であるが、フィルムグレインを含んだ画像・映像に圧縮符号化処理を施すと、影響を受けやすく簡単に失われてしまうことが知られている [2]。これは、動き補償と DCT (Discrete Cosine Transform) という符号化方式において、フィルムグレインは時間的に無相関な信号として表現されるために動き補償による予測がうまく働かず、予測差分信号として現れることによる。また、フィルムグレインの値自体が小さいため、その後の量子化処理によって失われやすいことも影響している。近年標準化された H.264/AVC では新たに追加されたデブロッキングフィルタや 4x4 サイズの整数変換といったツールのため、よりフィルムグレインが失われやすい問題点がある。実際にフィルムグレインを含んだ写真をスキャナで取り込み、H.264/AVC の参照ソフトウェア JM (Joint Model) [6] で符号化した画像を図 1 に示す。H.264/AVC の拡張方式 FRExt (Fidelity Range Extensions) では、高解像度 (HD) サイズでのフィルムグレインを含んだ映画コンテンツの符号化における画質改善を目的として、8x8 ブロックサイズ整数変換や量子化マトリクスといった画質を改善するツールが加えられた。FRExt によって主観画質の改善がなされ、次世代光ディスクの映像圧縮符号化方式で利用される可能性が高い。

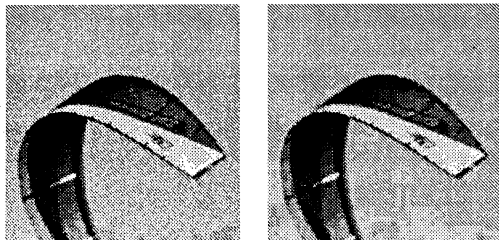
筆者らは、FRExt を使用したフィルムグレインの再現性に着目した H.264/AVC 符号化方式について検討してきた [3]。本稿では、さらに符号化モード決定のコスト値について考察することでフィルムグレインの再現性を更に向上させる手法の検討を行う。

2. フィルムグレイン再現性向上のツール・手法

H.264/AVC 符号化方式におけるフィルムグレイン再現性向上のためのツール・手法について以下で説明する。

2.1 デブロッキングフィルタ

ブロック分割ベース符号化の低ビットレート時に発生するブロックノイズを低減するためのツールである。H.264/AVC では、デコード画像を参照フレームバッファに挿入する直前において作用し、閉ループ内のツール群の一つとして定義されている。符号化モードや動きの特性によって、マクロブロックお



(a) 原画像 (b) H.264/AVC で符号化した画像

図 1 フィルムグレインの劣化

びブロック境界に適応的にフィルタを施したり、過度のフィルタを避けるために、量子化パラメータやオフセット値による強度の制御が可能である。しかし、フィルタを施すことでフィルムグレインが失われやすいのは確かであり、フィルムグレイン再現性の向上を目的とする場合、作用させないほうが望ましいと考えられる。

2.2 8x8 ブロックサイズ整数変換

FRExt でない H.264/AVC は、4x4 サイズの整数変換を使用する。ブロックノイズが目立ちにくいことや計算の容易性といった利点があるが、フィルムグレインのような細部を表現するにはより大きなサイズでの変換を使用したほうがよいことが知られている [4]。但し、H.264/AVC 符号化ツールの特徴である 8x8 ブロックサイズ未満の細かな動き補償を行えないという欠点もあるため、4x4 整数変換と 8x8 整数変換を適応的に用いることが望ましいと考えられる。

2.3 量子化マトリクス

入力画像に沿った適応的な量子化を行うことを目的としたツールであり、FRExt では、4x4・8x8 ブロックサイズ、インター・イントラ、輝度信号・色差信号、それぞれに対して定義でき、これらはピクチャ単位またはシーケンス単位で定義できる。量子化マトリクスの情報をデコーダに送らなければならないために発生符号量が増加するが、主観画質を向上させることが出来る。

2.4 量子化丸めオフセット

JM による量子化は、デッドゾーンが存在する量子化を採用しており、以下の式で表される。

$$Z = \text{int} \left(\frac{|W| + f}{\Delta} \right) \cdot \text{sign}(W) \quad (1)$$

$$W' = \Delta \cdot Z \quad (2)$$

$$f = \begin{cases} \Delta/3 & \text{Intra} \\ \Delta/6 & \text{Inter} \end{cases} \quad (3)$$

W は変換係数、 Z は量子化後の変換係数、 f は量子化丸めオフセット、 $\text{int}(x)$ は x を整数値化する処理、 $\text{sign}(x)$ は x の符号、 Δ は量子化ステップサイズ、 W' は再構成された変換係数を表している。 f によりデッドゾーンのサイズが変わるが、式 (3) からわかるようにインターとイントラで値が異なる。先に述べたように、フィルムグレインは微小係数として現れやすいため、デッドゾーンが小さいほうが再現性が向上するが、一方で統計モデル (ラプラス分布) から反れることとなり符号化効率の低下を招く [5]。JM version 9.4 から f の値を適応的に更新する量子化手法が実装されている。

2.5 フィルムグレインの SEI メッセージ化

符号化効率の低下を招く原因となるフィルムグレインをエンコーダ側で分離し、分離したフィルムグレインの統計的な性質を解析してパラメータ化する手法である [2]。このパラメータは H.264/AVC で定められた SEI メッセージ化されデコーダに付加情報として送られる。デコーダ側では、フィルムグレインを含んでいない画像デコードし、表示する直前に SEI メッセージを用いて擬似的にフィルムグレインを生成して合成する。映像

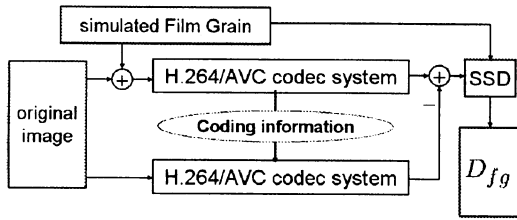


図2 フィルムグレイン差分電力 (D_{fg}) の定義

配信などの低ビットレート向けアプリケーションに有効であり、主観画質の改善がなされている。欠点としては、フィルムグレインのSEIメッセージを解釈できないデコーダでフィルムグレインをデコードできないことや、デコーダに負荷がかかることが挙げられる。

3. 予備実験

2. 節で述べた様々なフィルムグレイン再現性を向上させるツールに対し、符号化モード決定を変更することでフィルムグレイン再現性の向上を目指す。このために、まず以下でフィルムグレイン再現性の指標を定義し、この指標の妥当性について検討する。検討は整数変換のサイズの違いから行う。

3.1 フィルムグレイン再現性の定義

今までに、符号化によって劣化した映像に対するフィルムグレイン再現性に関する客観的な指標は定義されていない。そこで、フィルムグレイン再現性の指標について定義することを考える。算出の流れを図2に示す。本検討においては、フィルムグレインを含んでいない画像に擬似的にフィルムグレインを付加するという方式を取っている。実際は、フィルムグレインを含んだ画像をフィルムグレインとフィルムグレインを含まない画像とに分離するプロセスが必要となるが、本検討においては考慮しない。図2のH.264/AVCコーデックシステムを $F()$ とし、 $F(x)$ は信号 x をH.264/AVC符号化・復号する関数である。本検討ではH.264/AVCコーデックシステムの線形性を仮定する。即ち以下のように表現可能であるとす。

$$F(A+B) = F(A) + F(B) \quad (4)$$

$$F(B) = F(A+B) - F(A) \quad (5)$$

$$D_{fg} \equiv \sum \{F(B) - B\}^2 \quad (6)$$

A がフィルムグレインを含んでいない画像、 B がフィルムグレインである。但しこの際、 $F()$ に関して以下の条件を課すこととする。

- a) 同じものを使用するもの
 - 符号化モード
 - 動きベクトル
 - 参照フレーム番号
 - 整数変換ブロックサイズ
 - 量子化パラメータ
- b) それぞれに対して定義・生成するもの
 - 参照フレームバッファ

表1 符号化条件

Version	9.3
Profile	High
Image Size	1920x1072(Progressive)
Number of Coded Frames	13
Frame Structure	M=3, N=12
QP	24,26,28,30,32
Loop Filter	off
Number of Reference Frames	1
Search Range	48
Symbol Mode	CABAC

• イントラ予測画像

よって、式(6)において $B=0$ とした場合 $F(B)=0$ となり、 $D_{fg}=0$ である。フィルムグレインを含んだ画像の再構成とフィルムグレイン含んでいない画像の再構成の画素領域での差分を計算することで、フィルムグレイン成分の再構成が得られると考え、フィルムグレイン成分の再構成とフィルムグレイン成分とのSSD (Sum of Squared Difference) を計算することで、フィルムグレイン差分電力 D_{fg} が算出できると考えた。

3.2 実験条件

使用したシーケンスは、SpySorge [7](24[fps])、PedestrianArea [8](25[fps])、Photos(24[fps]) である。Photosは、Kodak ISO 800 フィルムで撮影した写真を、Digital ICE, Digital GEM フィルタ [9] でホコリ除去やフィルムグレイン除去をし、スキャナで取り込んだものである。ホコリ除去したものをフィルムグレインを含んだ画像として、ホコリ除去・フィルムグレイン除去をしたものをフィルムグレインを含まない画像としてそれぞれ使用した。符号化条件を表1に示す。フィルムグレインの擬似的なモデル式は様々なものがあるが、文献[10]で述べられているモデルは多くの論文で使用されており、本検討でもこのモデルを採用した。このフィルムグレインのモデル式を以下に示す。

$$g(x, y) = f(x, y) + n(x, y)$$

$$n(x, y) = \alpha \{f(x, y)\}^\gamma u_1(x, y) + u_2(x, y) \quad (7)$$

$f(x, y)$ は原画像、 $g(x, y)$ はフィルムグレインを含んだ画像、 $n(x, y)$ はフィルムグレインであり、 α は比例定数、 γ は $1/2 \sim 1/3$ の値をとる定数、 $u_1(x, y)$ と $u_2(x, y)$ は平均値ゼロの白色ガウス雑音である。本検討においては、実際のフィルムグレインを含んだ画像と比較して妥当と判断したときの値である $\alpha = 1.0$ 、 $\gamma = 0.5$ 、 $u_1(x, y)$ の分散 = 0.5、 $u_2(x, y) = 0$ を使用し、それぞれSpySorgeとPedestrianAreaに対してRGB領域に付加した。

3.3 予備実験結果

整数変換サイズの違いから、 D_{fg} の値の妥当性について検討する。フィルムグレインの再現性の違いは、整数変換サイズによって変わることが知られているためであり、具体的には、4x4 整数変換のみ、4x4 整数変換と8x8 整数変換、8x8 整数変換のみの3種類に分類した。検討結果を図3~5に示す。図の横軸はビットレート、縦軸は1画素当たりに平均化した D_{fg} で

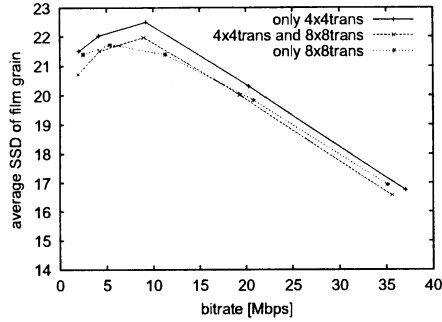


図3 D_{fg} の値 (SpySorge)

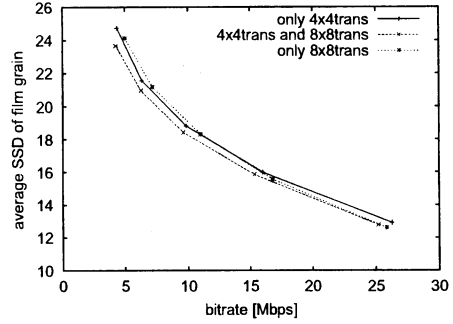


図5 D_{fg} の値 (Photos)

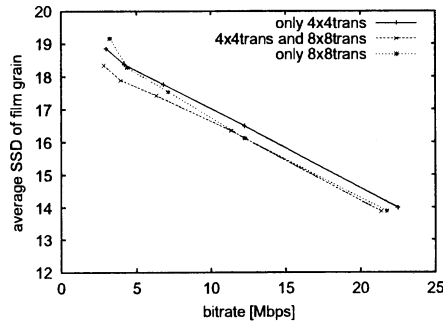


図4 D_{fg} の値 (PedestrianArea)

ある。 D_{fg} は差分電力なので値が小さければよく、すなわちグラフが左下にあるほど特性がよいと判断できる。どの図も基本的に4x4 整数変換のみが特性が悪く、4x4 整数変換と8x8 整数変換の両方及び8x8 整数変換のみはほとんど同じ特性であることがわかる。すなわち、 D_{fg} はフィルムグレインの再現性を表す指標として理論と一致した結果を表すが確認でき、以降の検討において、 D_{fg} をフィルムグレイン再現性の指標として使用する。

4. 提案手法

3. 節で求めた D_{fg} を符号化モード決定のためのコスト値に反映させることを考える。JM で使用されているレート歪み最適化手法による符号化モード決定コストは以下の通りである。

$$J = D + \lambda_{MODE} R \quad (8)$$

$$\lambda_{MODE,I,P} = 0.85 \times 2^{(QP-12)/3} \quad (9)$$

$$\lambda_{MODE,B} = \max\left\{2, \min\left(4, \frac{QP-12}{6}\right)\right\} \times \lambda_{MODE,I,P} \quad (10)$$

D は符号化歪み、 R は発生符号量、 λ_{MODE} はラグランジュ乗数であり、これらから計算されるラグランジュコスト J が最小となる符号化モードを選択することで符号化を行う。

提案手法は、式(8)に D_{fg} を埋め込む。フィルムグレインの値の範囲 (-255~255) を考慮して1/4 倍した D_{fg} を式(8)に

加算して符号化を行う。最適化という観点からは単に加算しただけではなし得ないが、以下の式を用いる。

$$J = D + (D_{fg} + 2) \gg 2 + \lambda_{MODE} R \quad (11)$$

式(11)によって求められるコスト値を以下の箇所で適用する。

- イントラ 4x4 のモード決定
- イントラ 8x8 のモード決定
- インター 8x8 以下のサブマクロブロックモード決定
- マクロブロックの符号化モード決定

イントラ 4x4 のモード決定、イントラ 8x8 のモード決定、インター 8x8 以下のサブマクロブロックモード決定は輝度値のみの D_{fg} を、マクロブロックの符号化モード決定は輝度値と色差値の D_{fg} をそれぞれ算出する。

5. 実験と考察

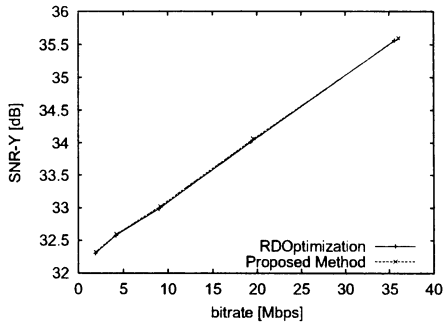
5.1 実験条件

提案コストを JM に組み込み、実験を行った。使用したシーケンスは、予備実験と同じく SpySorge, PedestrianArea, Photos であり、符号化条件も同じ表1の通りである。整数変換のサイズは、4x4 サイズと8x8 サイズの両方を使用可能とした。

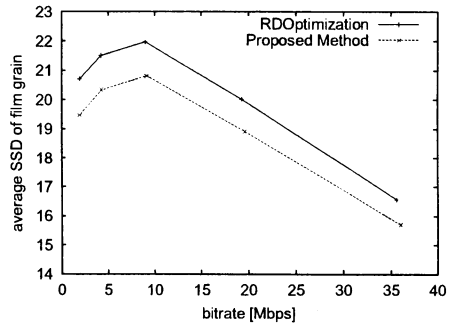
5.2 結果と考察

レート歪み最適化手法である式(8)を使用して符号化した場合と提案コストである式(11)を使用して符号化した場合の結果を図6~8に示す。SNRの比較の図(a)から、輝度成分のSNRの差はほとんど見受けられない。一方、図(b)の D_{fg} による差は見受けられ、提案手法のほうがすべての図において左下に曲線が描かれているため、予備実験の結果から、良好な特性が得られていると判断できる。

次に、選択された符号化モードの分布がどのように変わったかについて、各画像に対するモードの分布を表2~4に示す。これは、選択されたマクロブロック数を1フレームあたりに平均化した値である。これらの表からわかることとして、まず、QPの値が大きくなればなるほど、提案手法はレート歪み最適化手法のモード分布に近づくということである。これは、低ビットレート時は歪みが支配的であり、コスト値に対する D_{fg} の項の影響が少ないからであると考えられ、一方、高ビットレート時はレートの影響が大きくなり、それに伴い、 D_{fg} の項の影響が

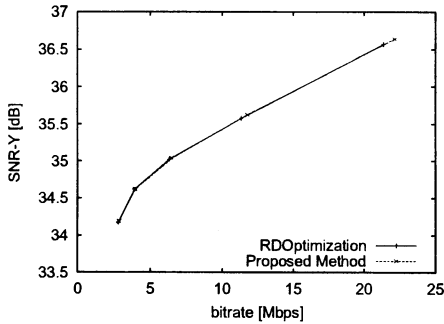


(a) SNR 比較

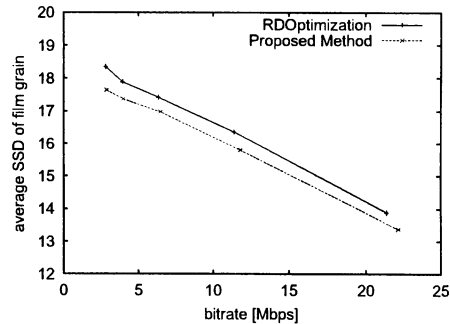


(b) D_{fg} 比較

図 6 レート歪み最適化手法と提案手法との比較 (SpySorge)

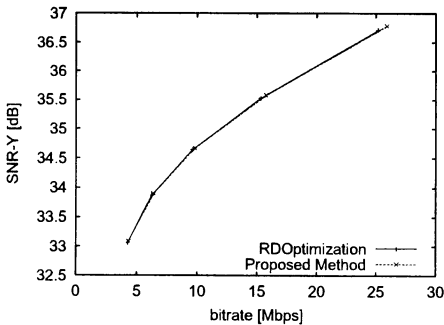


(a) SNR 比較

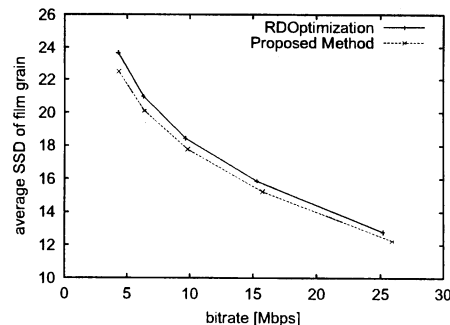


(b) D_{fg} 比較

図 7 レート歪み最適化手法と提案手法との比較 (PedestrianArea)



(a) SNR 比較



(b) D_{fg} 比較

図 8 レート歪み最適化手法と提案手法との比較 (Photos)

大きくなるからであると考えられる。次にわかることとして、提案手法の方がイントラ符号化の選択割合が高いということである。これは、 D_{fg} は定義上、歪みよりの傾向を持った値であるため、この値を足して重み付けすれば幾分歪みを重視したコストとなるからである。歪みを重視する傾向となっているので、イントラマクロブロックが選択される割合が増えていると考えられる。予備実験の結果から、 D_{fg} が低ければフィルムグレインの再現性が高いと判断されるため、適当にイントラマクロブロックが選択されるほうがフィルムグレインの再現性が高くな

ると結論付けられる。SpySorge を QP=28 で符号化したときの任意の P ピクチャ画像による比較を図 9 に示す。フィルムグレインが目立ちやすい均等色領域の一部 (200x150[pe]) を切り取ってきた画像である。提案手法のほうがわずかにフィルムグレインの再現性が高いのが確認できる。

6. まとめ

本稿では、H.264/AVC 符号化によるフィルムグレインの再現性を考慮した符号化方式について検討した。提案手法はフィ

表 2 モードの分布 (SpySorge)

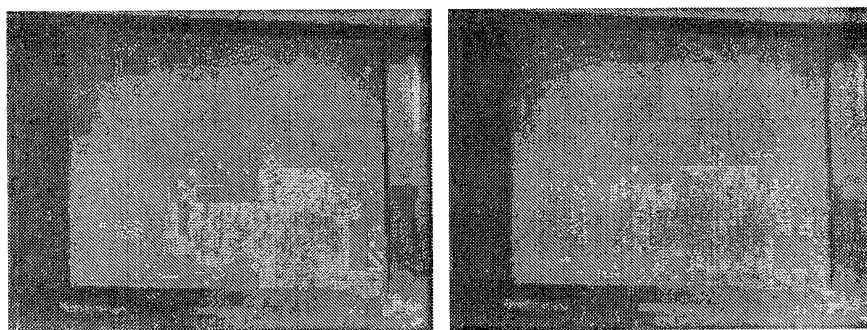
method	QP	Direct or Skip	Inter 16x16	Inter 16x8	Inter 8x16	Inter less than 8x8	Intra 4x4	Intra 8x8	Intra 16x16
RDOpt.	24	369.6	535.2	545.6	550.7	1704.2	598.5	2678.1	1058.0
Proposed	24	331.6	434.5	458.2	465.7	1545.1	780.2	2896.2	1128.4
RDOpt.	32	5804.5	716.8	108.6	137.5	60.6	49.6	804.4	358.1
Proposed	32	5781.8	706.2	122.3	136.8	74.9	51.2	805.8	360.8

表 3 モードの分布 (PedestrianArea)

method	QP	Direct or Skip	Inter 16x16	Inter 16x8	Inter 8x16	Inter less than 8x8	Intra 4x4	Intra 8x8	Intra 16x16
RDOpt.	24	1335.2	417.0	214.9	216.0	542.7	706.6	3270.8	1336.7
Proposed	24	1190.6	356.9	198.8	195.2	500.5	809.5	3368.1	1420.3
RDOpt.	32	3585.2	665.3	58.4	72.5	69.5	45.8	1937.2	1606.2
Proposed	32	3503.8	677.3	61.8	78.0	86.5	50.6	1935.2	1646.8

表 4 モードの分布 (Photos)

method	QP	Direct or Skip	Inter 16x16	Inter 16x8	Inter 8x16	Inter less than 8x8	Intra 4x4	Intra 8x8	Intra 16x16
RDOpt.	24	1775.7	1226.4	388.2	436.8	1087.5	794.4	1740.5	590.4
Proposed	24	1566.2	1092.8	367.1	408.4	1270.2	903.3	1795.8	636.3
RDOpt.	32	3740.0	1433.6	206.4	219.4	166.5	288.1	1119.7	866.4
Proposed	32	3767.4	1335.8	229.8	242.8	175.8	285.5	1130.3	872.5



(a) レート歪み最適化手法

(b) 提案手法

図 9 画像による比較 (P ピクチャの一部:200x150[pel])

フィルムグレイン再現性の指標を定義し、符号化モード決定コスト値に反映させることで実現した。PSNRの低下は見受けられないのに対し、わずかにフィルムグレインの再現性の向上が見受けられた。今後はコスト値と量子化の方式に関して検討していく予定である。

文 献

- [1] ISO/IEC 14496-10, "Advanced Video Coding," (ITU-T Rec. H.264) 2003.
- [2] C.Gomila et al., "SEI message for film grain encoding," JVT document, JVT-H022, May 2003.
- [3] 後藤他, "フィルムグレインを考慮した H.264/AVC モード決定手法に関する一検討," 信学総大 D-11-56, Mar 2005.
- [4] Mathias Wien, "Variable Block-Size Transforms for H.264/AVC," IEEE Trans. Circuits Syst. Video Technol., vol.13, pp.604-613, July 2003.
- [5] T.Wedi et al., "Quantization with an Adaptive Dead Zone Size for H.264/AVC FRExt," JVT document, JVT-K026, Mar 2004.
- [6] "JVT Reference Software version 9.3," <http://iphome.hhi.de/suehring/tml/>, Mar 2005.
- [7] 藤田正浩原作・制作・監督作品, "スパイ・ゾルゲ," 編集前源

情報 (GITI).

- [8] "MPEG-Testsequences," ftp://ftp.ldv.e-technik.tu-muenchen.de/pub/test_sequences/
- [9] "DIGITAL ICE3 Technology for Film Scanners," <http://www.asf.com/products/ICE3/Example.shtml>
- [10] D.T.Kuan et al, "Adaptive noise smoothing filter for images with signal-dependent noise," IEEE Trans. Pattern Anal. Machine Intell., vol.7, pp.165-177, Mar 1985.