

## 超高精細映像を対象とした高圧縮符号化における 最大マクロブロックサイズの決定方式に関する一検討

松村 篤志<sup>†</sup> 内藤 整<sup>†</sup>  
川田 亮一<sup>†</sup> 小池 淳<sup>†</sup>

4K システムに代表される超高精細映像は非圧縮時において膨大なデータ量となるため、同素材をビットレート制限下で伝送する際には高圧縮符号化が用いられる。これに対し筆者らは、H.264 High Profile を用いた符号化における最大マクロブロックサイズを拡張することによって、より高効率な圧縮符号化を実現させるための方式を提案した。本稿では、同方式における最適な最大マクロブロックサイズを、過去の符号化結果に基づいて決定するための方式を提案する。また、シミュレーション実験により提案方式の評価を行ったので、その結果を報告する。

### An Effective Maximum Macroblock Size Decision Scheme in High Compression Encoding for Super High Definition Images

ATSUSHI MATSUMURA,<sup>†</sup> SEI NAITO,<sup>†</sup> RYOICHI KAWADA<sup>†</sup>  
and ATSUSHI KOIKE<sup>†</sup>

High compression encoding is required in order to transmit super high definition videos, such as 4K systems, under the limited bit-rate because of their large data amounts. Therefore, the authors have proposed a scheme to improve the performance of H.264 High Profile encoding by extending maximum macroblock sizes. In this paper, a scheme to decide the appropriate maximum macroblock size by using the previously encoded results is proposed. The effectiveness of the proposed method is evaluated by several experimental results.

#### 1. はじめに

HD サイズ超の解像度を持つ超高精細映像アプリケーションに対する規格の1つとして、4096×2160の素材(以下4K素材と表記)を扱うシステムがDCIより提唱されている<sup>1)</sup>。同素材は非圧縮時において400Mbits/frameという膨大なデータ量となるため、伝送を行う際には、圧縮符号化の適用が現実的である。例えば、DCI規格ではMotion JPEG2000を用いて250Mbpsにて圧縮符号化を行う規定がなされている。しかしながら、同素材をインターネットを介した映像ストリーミングサービスなどの低ビットレート伝送に適用させるためには、更なる高圧縮符号化が必要といえる。

高圧縮符号化には、動き補償予測を用いた方式が一般に用いられており、特に標準符号化方式の1つであるH.264では、解像度の上限が4096×2304と定義されているため、4K素材への適用が可能となっている<sup>2)</sup>。しかしながら、同方式における最大マクロブロックサイズ(16×16)は、4K素材における画像解像度に対して小さすぎるため、サイド情報符号量が冗長となる場合がある<sup>3)</sup>。そこで筆者らは、先にH.264 High

Profileにおける最大マクロブロックサイズを拡張させることによって、4K素材に対する符号化効率を改善させるための方式を提案し、シミュレーション実験によりその有効性を示した<sup>4)</sup>。同方式における最適な最大マクロブロックサイズは素材やビットレートに依存するため、符号化効率の低下を抑止するためには、適応的に最大マクロブロックサイズを決定することが必要である。

そこで本稿では、過去の符号化結果をもとに、上記方式における最適な最大マクロブロックサイズを適応的に決定するための方式を提案する。また、シミュレーション実験により提案方式の評価を行ったので、その結果を報告する。

#### 2. 最大マクロブロックサイズ拡張型 H.264 High Profile 符号化方式

文献<sup>4)</sup>におけるマクロブロック(以下MBと表記)単位の符号化処理手順を図1に示す。同方式の特徴、および本稿においてこれを適用する際の特徴を以下に示す。

- 当該ピクチャ $t$ における最大MBサイズの水平方向および垂直方向の画素数(以下、単に最大MBサイズと表記)を $N(t)$ と定義する。ただし、 $N(t)$ は2のべき乗とし、最小値は16とする。また、

<sup>†</sup> 株式会社 KDDI 研究所 映像通信グループ  
KDDI R&D Labs Inc. Visual Communication Group.



図 1 符号化処理手順

$N(t)$  はピクチャ単位での更新を可能とし、その際の値は、 $N(t')$ 、 $2N(t')$  および  $N(t')/2$  のいずれかとする。ただし、 $t'$ 、 $N(t')$  は、それぞれ同一ピクチャタイプの直前ピクチャ、同ピクチャにおける最大 MB サイズを表す。

- 選択可能な MB サイズは、Intra MB においては  $N(t) \times N(t)$ 、 $N(t)/2 \times N(t)/2$  および  $N(t)/4 \times N(t)/4$  の 3 種類とする。Inter MB においては  $N(t) \times N(t)$ 、 $N(t)/2 \times N(t)$ 、 $N(t) \times N(t)/2$ 、 $N(t)/2 \times N(t)/2$  の 4 種類の中から選択した上で、 $N(t)/2 \times N(t)/2$  が選択された際には、各 MB におけるサブ MB サイズを  $N(t)/2 \times N(t)/2$ 、 $N(t)/4 \times N(t)/2$ 、 $N(t)/2 \times N(t)/4$  および  $N(t)/4 \times N(t)/4$  の 4 種類の中から選択する。
- Intra MB における予測誤差信号の DCT は、表 1 に従って行われる。また、Inter MB においては、MB をサイズ  $4 \times 4$  ないし  $8 \times 8$  のブロックに分割 (符号化効率に応じて適応的に選択) した上で、各ブロックに対し DCT を適用する。
- CBP フラグによる DCT 係数符号量の記述判定は、 $N(t)/2 \times N(t)/2$  単位で行う。
- 各ピクチャにおける最大 MB サイズ情報は表 2 に従ってスライスヘッダ部に記述する。また、 $N(t')$  が存在しない場合においては、 $\log_2(N(t)) - 4$  を算術符号化した上で記述する。

上記に記述がない事項に関しては、H.264 High Profile の規定に従う。

表 1 Intra MB における DCT 変換方式

MB サイズ	DCT 変換方式
4×4	ブロックサイズ $4 \times 4$ にて DCT を適用した上で、各係数に対する符号化を行う。
8×8	ブロックサイズ $8 \times 8$ にて DCT を適用した上で、各係数に対する符号化を行う。
16×16 以上	ブロックサイズ $4 \times 4$ にて DCT を適用した上で、DCT 係数を AC 成分と DC 成分に分離した後、各成分を個別に符号化する。

表 2 各最大 MB に対する記述データ

最大 MB サイズ	記述データ
$N(t) = N(t')$	0
$N(t) = 2N(t')$	10
$N(t) = N(t')/2$	11

### 3. 予備実験

最大 MB サイズ決定方式の検討に際し、予備実験を行った。符号化シミュレーション条件を表 3 に示す。実験素材は、DCI の提供するフレーム周波数 24Hz、プログレッシブスキャンの標準動画画像 StEM<sup>5)</sup> より、1865 フレームから 1892 フレームを抽出した上で、画面中央部より 2048 画素  $\times$  1088 ラインを切り出したものを用いた。ただし、同素材に対し、YCbCr4:2:0、深度各 8 ビットの色調変換 (以下、単に色調変換と表記) を行った。

まず、最大 MB サイズ拡張に対する評価を行うため、最大 MB サイズを 16、32 および 64 に固定して符号化を行った際の、PSNR とビットレートとの関係を図 2 に示す。ただし、同実験における量子化パラメータ (以下 QP と表記) は 25、35 および 45 とした。同結果より、ビットレートによって最適な最大 MB サイズが異なり、特に低ビットレートにおいて最大 MB サイズ拡張による符号量削減効果が現れていることがわかる。最大 MB サイズが 16 である従来方式と比較し、QP45 にて約 51.6% の符号量の削減が見られた。

次に、以下に示す 3 種類の値のうち、値 1 と値 3、および値 2 と値 3 の関係を調査した。

値 1: サイズ  $N \times N$  の Inter (I ピクチャでは In-

表 3 符号化シミュレーション条件

設定項目	設定値
符号化方式	CABAC
アダマール変換	なし
動き予測	UMHexagon Search
R-D 最適化	あり
GOP 構成	IBBPBBPBBPBBPBB
参照ピクチャ	I および P ピクチャのみ
参照ピクチャ数	2
動ベクトル探索範囲	参照ピクチャに依らず $\pm 128 \times \pm 128$

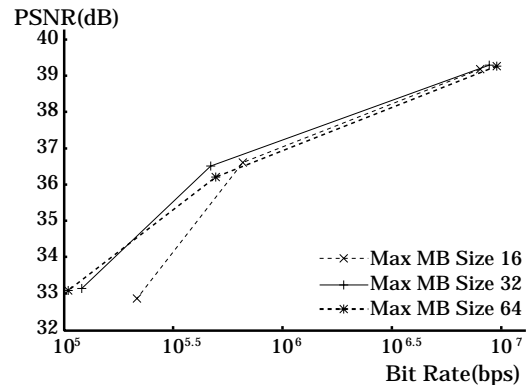


図 2 PSNR とビットレートの結果

tra)MBのみを用いて符号化を行った際に発生する画素単位の DCT 係数符号量

値 2: 符号化に用いられる量子化ステップ

値 3: 最大 MB サイズを  $M_2$  とした際の発生符号量に対する, 最大 MB サイズを  $M_1$  とした際の発生符号量の比率. ただし, 上記における発生符号量はいずれもサイド情報を含むこととし, 以下の記述においても断りがない限り同様とする. また, 符号化には適用可能な全 MB タイプを用いることとする.

ただし, ピクチャ間での QP の変動は行わないこととした.  $N = 16$ ,  $M_1 = 16$ ,  $M_2 = 32$  とし, QP を 25 から 35 に順次変動させて符号化を行った際の結果を図 3 および図 4 に示す. ただし, I ピクチャに関しては, 3 ピクチャ毎に Intra 符号化した結果を示す. 結果より, いずれのピクチャタイプにおいてもこれらの関係に強い相関があることがわかる. 素材や QP を上記以外の条件とした場合, および以下に示すパラメータ条件を用いた場合の結果においても同様の傾向が見られた.

条件 1:  $N = 32$ ,  $M_1 = 16$ ,  $M_2 = 32$

条件 2:  $N = 32$ ,  $M_1 = 32$ ,  $M_2 = 64$

条件 3:  $N = 64$ ,  $M_1 = 32$ ,  $M_2 = 64$

#### 4. 提案方式

3 章の結果より, 各最大 MB サイズを適用した発生符号量の比率は, 量子化パラメータおよび DCT 係数符号量と強い相関があることがわかった. そこで提案方式では, 同比率を符号化済ピクチャにおける量子化パラメータおよび DCT 係数符号量から推定し, その結果に基づき最大 MB サイズを決定する. 処理の流れを図 5 に示す. ただし,  $N(init)$ ,  $N(max)$  および  $N(min)$  はそれぞれ最大 MB サイズの初期値, 上限および下限とし,  $E(N(t'), 2N(t'))$  および  $E(N(t')/2, N(t'))$  は発生符号量比率の推定値とする. 処理の詳細を以下に示す.

Step 1: 初期値代入判定

$t'$  が存在しない場合においては  $N(t) = N(init)$  とした上で処理を終了する. それ以外の場合においては Step 2 に進む.

Step 2: MB サイズ上限判定

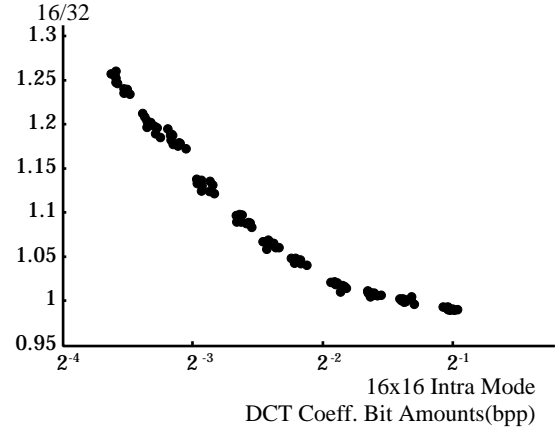
$N(t') = N(max)$  であった場合は Step 4 へ, それ以外の場合は Step 3 へ進む.

Step 3: MB サイズ拡大判定

まず, 式 (1) を用いて, 最大 MB サイズが  $2N(t')$  の場合の発生符号量に対する, 同サイズが  $N(t')$  の場合の発生符号量の比率を推定する.

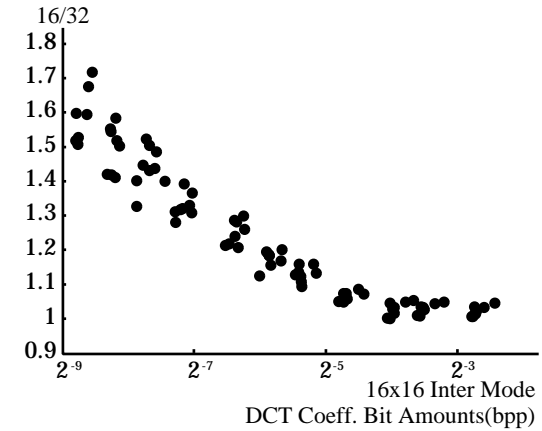
$$\begin{aligned}
 E(N(t'), 2N(t')) = & \\
 & k_1(N(t'), p(t')) \times \log_2(Q(t'))^2 + \\
 & k_2(N(t'), p(t')) \times \log_2(Q(t')) + \\
 & k_3(N(t'), p(t')) \times \log_2(B(t'))^2 + \\
 & k_4(N(t'), p(t')) \times \log_2(B(t')) + \\
 & k_5(N(t'), p(t')) \quad (1)
 \end{aligned}$$

Ratio of Bit Amounts  
bet. Max MB Size



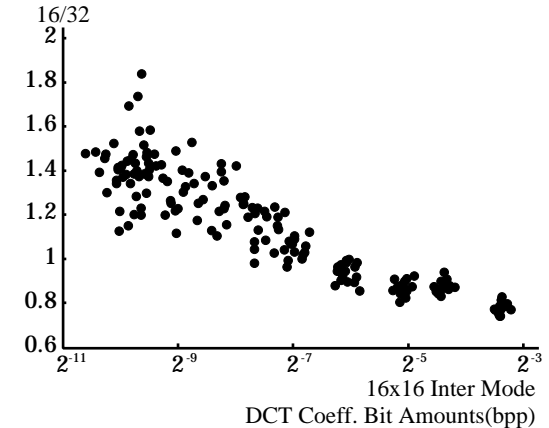
(a) I ピクチャ

Ratio of Bit Amounts  
bet. Max MB Size



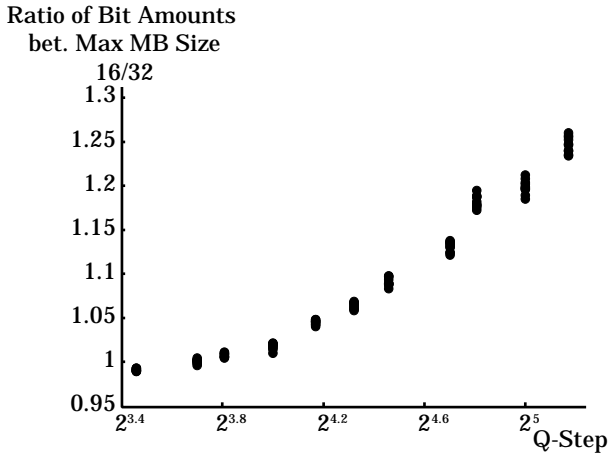
(b) P ピクチャ

Ratio of Bit Amounts  
bet. Max MB Size

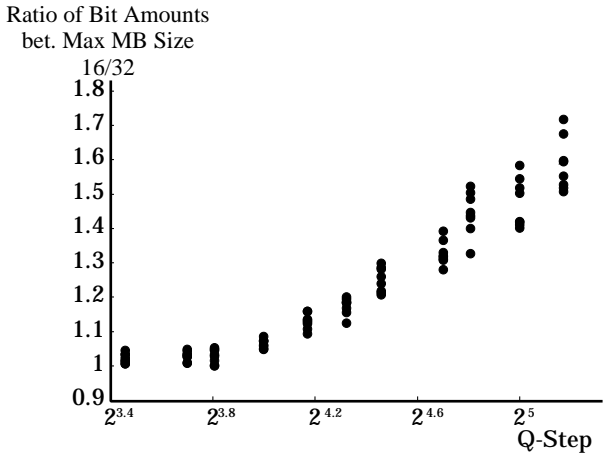


(c) B ピクチャ

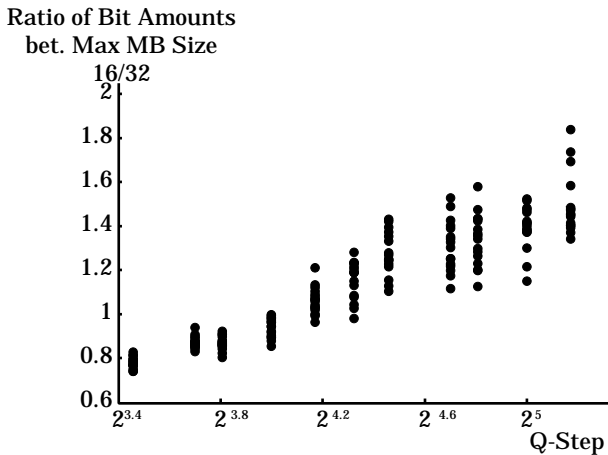
図 3 DCT 係数符号量と発生符号量の比率の結果



(a) I ピクチャ



(b) P ピクチャ



(c) B ピクチャ

図 4 量子化ステップと発生符号量の比率の結果

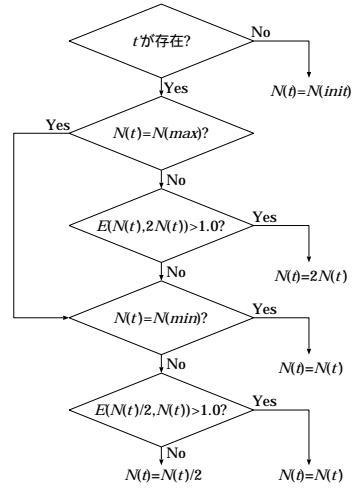


図 5 提案方式の流れ

ただし、 $Q(t')$ 、 $B(t')$  は、それぞれピクチャ番号  $t'$  における量子化ステップ、最大 MB サイズにて符号化を行った際の画素単位の平均発生符号量を表し、 $k_n(N(t'), p(t'))$  ( $n = 1, 2, 3, 4, 5$ ) は、 $N(t')$ 、および  $t'$  におけるピクチャタイプ  $p(t')$  によって決定される定数とする。また、 $B(t')$  の算出には、R-D コスト算出処理において求められた値を用いる。

$E(N(t'), 2N(t'))$  が 1.0 よりも大きい場合においては、 $N(t) = 2N(t')$  とした上で処理を終了する。それ以外の場合は Step 4 へ進む。

Step 4: MB サイズ下限判定

$N(t') = N(\min)$  であった場合は、 $N(t) = N(t')$  とした上で処理を終了する。それ以外の場合は Step 5 へ進む。

Step 5: MB サイズ縮小判定

まず、式 (2) を用いて、最大 MB サイズが  $N(t')$  の場合の発生符号量に対する、同サイズが  $N(t')/2$  の場合の発生符号量の比率を推定する。

$$E(N(t')/2, N(t')) = k'_1(N(t'), p(t')) \times \log_2(Q(t'))^2 + k'_2(N(t'), p(t')) \times \log_2(Q(t')) + k'_3(N(t'), p(t')) \times \log_2(B(t'))^2 + k'_4(N(t'), p(t')) \times \log_2(B(t')) + k'_5(N(t'), p(t')) \quad (2)$$

ただし、 $k'_n(N(t'), p(t'))$  ( $n = 1, 2, 3, 4, 5$ ) は、 $N(t')$  および  $p(t')$  によって決定される定数とする。

$E(N(t')/2, N(t'))$  が 1.0 よりも大きい場合は、 $N(t) = N(t')$  とし、それ以外の場合は、 $N(t) = N(t')/2$  とする。

## 5. 性能評価

提案方式の性能を評価するため、シミュレーション実験を行った。事前実験の結果より取得した式 (1) お

よび式 (2) における定数を表 4 に示す．符号化シミュレーション条件は表 3 に従うこととした．

### 5.1 発生符号量比率の推定精度の評価

まず，式 (1) および式 (2) の推定精度を評価するため，これらの式により推定した発生符号量の比率と実測値の比較を行った．実験では，StEM より 3 シーケンス (1865 フレーム，1913 フレーム，2157 フレームより各 28 フレーム) を抽出した上で，3 章と同様の切り出し処理および色調変換を行った素材を用いた．QP を 25 から 35 に順次変動させた際の各ピクチャタイプにおける結果を図 6 に示す．結果より，いずれのピクチャタイプにおいても良好な推定結果となっていることがわかる．従って，これらの式を用いて最適な最大 MB サイズの決定を高精度に行うことができると考えられる．

### 5.2 符号化性能の評価

提案方式における符号化性能評価実験を行った．実験素材として，5.1 節にて用いた 3 シーケンスを抽出した上で色調変換を行い，さらに上 230 ライン，下 232 ラインに黒色画素を埋め込んだ素材 (各素材 4096 画素  $\times$  2176 ライン) を用いた．また， $N(init)$ ， $N(max)$  および  $N(min)$  は 32，64 および 16 とした．

QP を 25，35 および 45 として符号化を行った際の

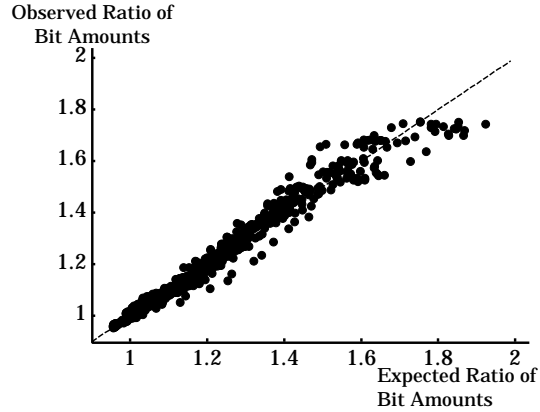
表 4 式 (1) および式 (2) における定数  
(a)  $k_n(N(t'), p(t'))$  ( $n = 1, 2, 3, 4, 5$ ) の値

	$N(t') = 16$	$N(t') = 32$
$p(t')=I$	$k_1(16, I) = 0.0427$ $k_2(16, I) = -0.4148$ $k_3(16, I) = 0.0121$ $k_4(16, I) = -0.0604$ $k_5(16, I) = 1.8594$	$k_1(32, I) = 0.0005$ $k_2(32, I) = -0.0521$ $k_3(32, I) = 0.0233$ $k_4(32, I) = 0.0917$ $k_5(32, I) = 1.1063$
$p(t')=P$	$k_1(16, P) = 0.0151$ $k_2(16, P) = -0.0610$ $k_3(16, P) = 0.0075$ $k_4(16, P) = 0.0162$ $k_5(16, P) = 1.0274$	$k_1(32, P) = 0.0240$ $k_2(32, P) = -0.2574$ $k_3(32, P) = 0.0088$ $k_4(32, P) = 0.0503$ $k_5(32, P) = 1.6046$
$p(t')=B$	$k_1(16, B) = 0.0017$ $k_2(16, B) = 0.0818$ $k_3(16, B) = 0.0072$ $k_4(16, B) = 0.0339$ $k_5(16, B) = 0.5961$	$k_1(32, B) = 0.0729$ $k_2(32, B) = -0.7949$ $k_3(32, B) = 0.0025$ $k_4(32, B) = -0.0219$ $k_5(32, B) = 2.7035$

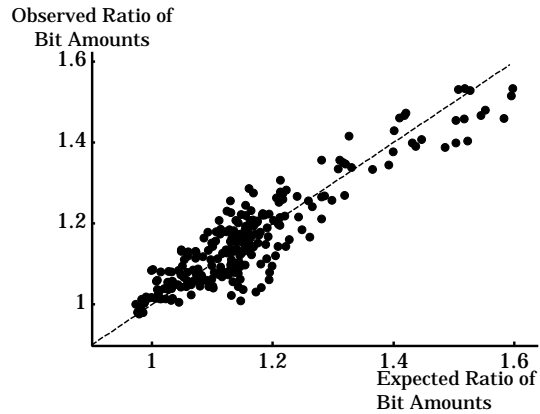
(b)  $k'_n(N(t'), p(t'))$  ( $n = 1, 2, 3, 4, 5$ ) の値

	$N(t') = 32$	$N(t') = 64$
$p(t')=I$	$k'_1(32, I) = 0.0456$ $k'_2(32, I) = -0.4392$ $k'_3(32, I) = 0.0144$ $k'_4(32, I) = -0.0563$ $k'_5(32, I) = 1.9155$	$k'_1(64, I) = 0.0038$ $k'_2(64, I) = -0.0827$ $k'_3(64, I) = 0.0228$ $k'_4(64, I) = 0.0850$ $k'_5(64, I) = 1.1672$
$p(t')=P$	$k'_1(32, P) = 0.0156$ $k'_2(32, P) = -0.0655$ $k'_3(32, P) = 0.0067$ $k'_4(32, P) = 0.0060$ $k'_5(32, P) = 1.0262$	$k'_1(64, P) = 0.0202$ $k'_2(64, P) = -0.2244$ $k'_3(64, P) = 0.0066$ $k'_4(64, P) = 0.0112$ $k'_5(64, P) = 1.4328$
$p(t')=B$	$k'_1(32, B) = 0.0111$ $k'_2(32, B) = 0.0022$ $k'_3(32, B) = 0.0047$ $k'_4(32, B) = 0.0034$ $k'_5(32, B) = 0.6986$	$k'_1(64, B) = 0.0716$ $k'_2(64, B) = -0.7774$ $k'_3(64, B) = 0.0008$ $k'_4(64, B) = -0.0452$ $k'_5(64, B) = 2.6124$

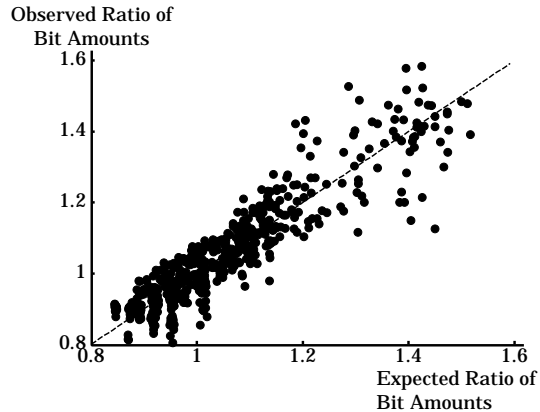
PSNR とビットレートとの関係を図 7 に示す．比較のため，最大 MB サイズを 16，32 および 64 に固定させて符号化した結果も併せて示す．結果より，提案方式を適用することによって，最大 MB サイズを固定とした場合における最良ケースの結果と同等の符号化効率を得られていることがわかる．



(a) I ピクチャ

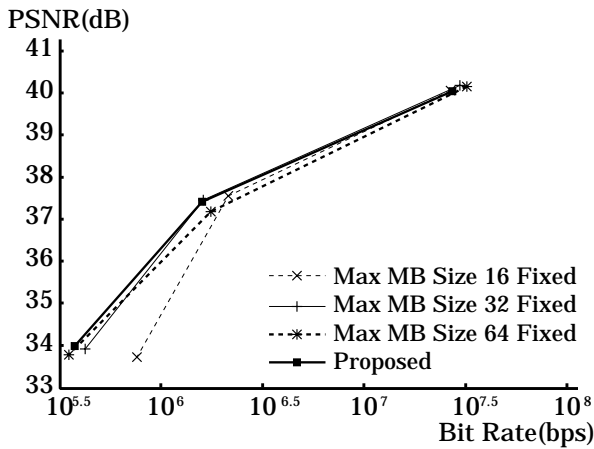


(b) P ピクチャ

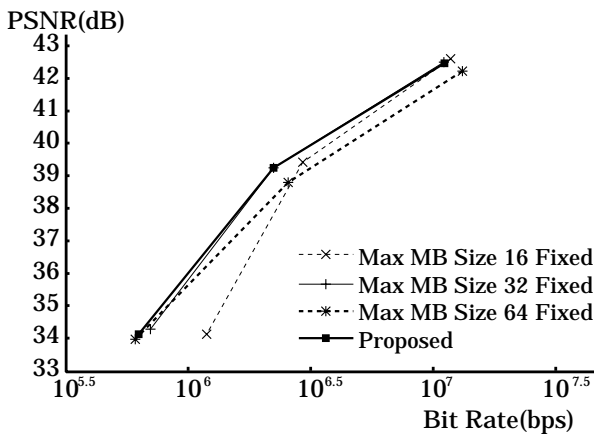


(c) B ピクチャ

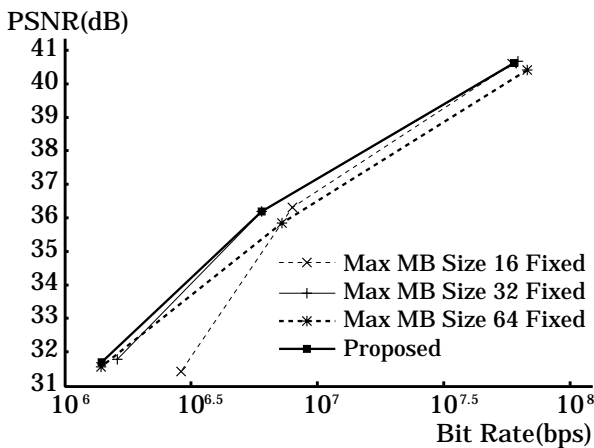
図 6 推定された発生符号量比率と実測値の結果



(a) 素材 1



(b) 素材 2



(c) 素材 3

図 7 PSNR とビットレートの結果

## 6. ま と め

本稿では、超高精細映像に対する符号化効率改善を目的とし、H.264 High Profile を用いた符号化における最大マクロブロックサイズを拡張した際の、最適な最大マクロブロックサイズを決定するための方式を提案した。提案方式では、最大マクロブロックサイズを過去の符号化結果に基づき決定しているため、当該ピクチャを予め符号化せずに最適な最大 MB サイズを決定させることができる。また、性能評価実験により、提案方式の適用によって、最適な最大 MB サイズを予め算出した場合と同等の符号化効率が得られることがわかった。

今後はレート制御に対応させるための検討、および他方式との組み合わせによる更なる符号化効率改善に向けた方式検討を行っていく予定である。

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。

## 参 考 文 献

- 1) Digital Cinema System Specification V1.0  
[http://www.dcinovies.com/DCI\\_Digital\\_Cinema\\_System\\_Spec.v1.pdf](http://www.dcinovies.com/DCI_Digital_Cinema_System_Spec.v1.pdf)
- 2) Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 - ISO/IEC 14496-10 AVC)  
<http://www.ncesd.org/vc/docs/JVT-G050.pdf>
- 3) 松村 篤志, 内藤 整, 川田 亮一, 小池 淳: “高精細動画画像を対象とした MPEG-4 AVC/H.264 FR-Ext 符号化処理の高速化方式に関する検討”, 信学技報, EMCJ2004-151, pp.25-28, Mar. 2005.
- 4) 松村 篤志, 内藤 整, 川田 亮一, 小池 淳: “超高精細映像の H.264 FR-Ext 符号化を対象とした最大マクロブロックサイズの拡張による符号化効率改善方式”, 映情学技報, AIT2005-96, Vol. 29, No. 42, pp.25-28, Jul. 2005.
- 5) StEM Information: Mini-movie Access Procedure  
<http://www.dcinovies.com/StEMAcce.pdf>