

## H.264|AVC におけるイントラ予測高速化の一検討

高木 幸一<sup>†</sup> 滝嶋 康弘<sup>†</sup> 中島 康之<sup>†</sup>

<sup>†</sup>株式会社 KDDI 研究所 〒356-8502 埼玉県上福岡市大原 2-1-15

E-mail: <sup>†</sup>{ko-takagi, takisima, nakajima}@kddilabs.jp

あらまし 動画画像圧縮の新勧告 | 標準である H.264|AVC において、フレーム内符号化の効率を向上させるために、隣接した既符号化ブロックから画素値を予測する方式が取られている。しかし、その予測のためのモードが複数存在するため、全モードに対し評価を行う場合と比較し、符号化効率を損なわず、可能な限り高速に、モードを選択する方式が必要となる。そこで、本稿では、新しいモード選択方法を提案し、計算複雑度、符号化効率（レート歪特性）の観点から従来方式との比較を行う。

キーワード H.264|AVC, フレーム内符号化, イントラ予測, 予測誤差

### A Study on Fast Intra Prediction Scheme for H.264|AVC

Koichi TAKAGI<sup>†</sup> Yasuhiro TAKISHIMA<sup>†</sup> and Yasuyuki NAKAJIMA<sup>†</sup>

<sup>†</sup>KDDI R&D Labs. Inc. 2-1-15 Ohara, Kamifukuoka-shi, Saitama, 356-8502 Japan

E-mail: <sup>†</sup>{ko-takagi, takisima, nakajima}@kddilabs.jp

**Abstract** ITU-T H.264|ISO MPEG-4 AVC, which is a new video coding recommendation | standardization, adopts an intra prediction method to predict pixel values from adjacent coded pixel values in order to improve intra frame coding efficiency. However, we need to select the optimal mode as early as possible because there are several intra prediction modes. Therefore, we propose that new mode selection scheme and compare the conventional method in complexity and coding efficiency.

**Keyword** H.264|AVC, Intra-frame coding, Intra prediction, Prediction error

#### 1. はじめに

ITU-T と ISO MPEG が共同で検討を進めてきた新動画画像符号化方式 H.264|AVC（以下 JVT と書く）が、ついに勧告化/標準化された[1]。本方式は従来の動画画像圧縮標準である MPEG-4 や勧告 H.263 等と比べ、50%の符号量削減を目指したものであったが、あらゆる団体からさまざまな方式が提案され、検討を重ねた結果、この目標はほぼ実現されたと考えられる。

従来標準と比較した際の本標準の変更点としては、VCL/NAL 構造、動き補償予測単位の精細化、1/4-画素精度動き補償、複数参照フレーム、インループフィルタなどさまざまなものが挙げられるが、大きく異なる点の一つとしてイントラ予測がある。これは予測を空間（画素）領域で行うものであり、従来の符号化標準で用いられている周波数領域における予測方式と比較し、符号化効率の向上に貢献していることがわかっている[6]。

一方、イントラ予測では、複数の予測方向が存在するため、これらの中から最適な方向をいかに高速に見つけ出すかが極めて重要になる。従来法としては、各予測方向に対して実際に使用するビットと誤差を利用

して、レート歪特性の観点から最適なモードを探索する方法[2]があるが、これはすべてのモードに対し実際に符号化を行うため、非常に時間を要する。一方、[3]で提案されている方法は、一度画像全体を読み込まないといけないため、それだけメモリが必要になる。

そこで、本稿では、従来法と比較して、とりわけメモリを必要とせず、さらに符号化効率を保ったままイントラ予測を高速に行う方法を提案する。まず2節において本稿で検討対象とする JVT のイントラ予測方式についてその概要を述べる。続いて3節でそのイントラ符号化のためのモードの予測法に関する従来法について述べ、4節でそれに対する提案手法を述べる。これらの比較を行うため5節にてシミュレーション実験を行い、最後にまとめを行う。

#### 2. イントラ符号化とイントラ予測

JVT では、各マクロブロック (MB) がスライス符号化モード (I, P, B など) に従い、さまざまな符号化タイプがとられる。すべてのスライス符号化モードで以下に述べるイントラ符号化モードがサポートされている。これには Intra\_4x4 と Intra\_16x16 が存在する。

Intra\_4x4 は 4x4 輝度ブロックごとに予測する方法で、

重要な detail を持った部分の符号化に非常に適している。一方で Intra\_16x16 は 16x16 画素の輝度ブロック全体の予測を行うモードで、平坦な部分の予測に適している。これらの2つのタイプの輝度予測に加え、色差の予測も別に行われる。

これまでのビデオ符号化標準ではイントラ予測は変換係数領域で行われていたのに対し、JVT ではすでに符号化された左または上のブロックの近傍の画素値から空間領域で行われる。ただし、インター符号化された MB において、動き補償のために伝播された伝送エラーがイントラ符号化モードを取るブロックにも伝播する可能性を秘めている。したがって、イントラ符号化を制限するモード(ConstrainedIntraPred)を選択することが可能となっている。

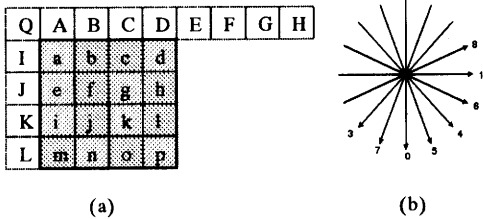


図 1. (a) Intra\_4x4 モードにおける 4x4 ブロック a-p と被予測画素 A-Q. (b) Intra\_4x4 モードにおける 8 方向の予測モード (DC 予測を除く).

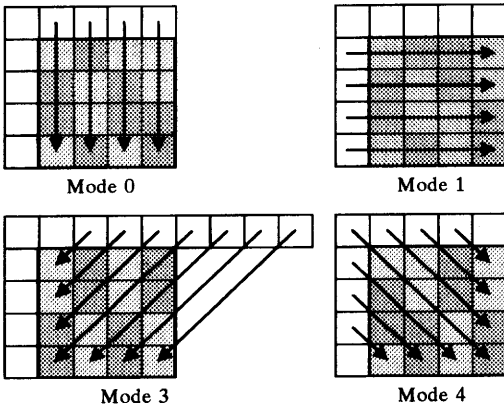


図 2. Intra\_4x4 予測モード (例)

Intra\_4x4 を利用する際、各 4x4 ブロックは別々に近傍の画素が予測される (図 1)。4x4 ブロックの 16 画素は直前にデコードされた隣接ブロックの画素から予測が行われる。各 4x4 ブロックにおいて 9 つの予測モードのうちから任意の一つが選択される。DC 予測に加えて同図に示すような 8 つの予測方向が存在する。

図 2 は 9 個の Intra\_4x4 のうちの 4 つを示したものである。モード 0 / 1 は上 / 左のブロックの隣接画素

から垂直 / 水平方向にそれぞれ予測を行い、モード 2 は上と左のブロックの隣接画素の平均値 (DC 成分) を予測値として使う。他の 6 個のモードは対角方向、あるいは対角方向と水平、垂直方向の間の方向への予測となる。例としてモード 3, 4 のみ示す。

なお、復号の際には、上と左のブロックそれぞれに対し、Intra\_4x4 が用いられていない場合、あるいは DC 予測を強制する場合に DC 予測が選択される。それ以外の場合は、左上のモード番号のうち小さい方に等しい場合はフラグ `prev_Intra_4x4_pred_mode_flag` = 1 (UVLC で 1bit) とし、さもなければ同フラグを 0 とした上で、モード番号をそのまま (あるいは 1 加えて) 変数 `rem_Intra_4x4_pred_mode` (UVLC で 3bit) で伝送する。このモードを利用しイントラ予測誤差の補償が行われるが、被予測画素が利用可能か否かにより選択できるモードが異なる。

Intra\_16x16 を扱う際には、MB の全輝度成分が予測される。予測モードは 4 方向サポートされている。予測モード 0~2 (垂直 / 水平方向 / DC 予測) は Intra\_4x4 で述べたものと非常に良く似ている (隣接画素として 4 画素を利用する点を 16 画素を利用すると置き換えればよい)。また、plane 予測モードと呼ばれるモードが存在する。なお、これらについても被予測画素が存在するか (イントラ予測制限フラグなどにも起因) により選択できるモードが異なる。

一方、色差成分についても Cb, Cr それぞれの成分について Intra\_16x16 とまったく同様な方法で予測が行われる (細かい点を除けば 16 画素を 8 画素に置き換えるだけである。)

### 3. 従来のイントラ予測モード選択法

前節に述べたとおり、イントラ予測には数多くのモードが存在するため、このモードをいかに符号化効率を落とさずに高速に求めるかが重要となる。ここで、JVT のテストモデル等で利用されているイントラ予測モード選択法について特に Intra\_4x4 に焦点をしほり概略を述べることにする。

#### 3.1. Rate-Distortion 最適法

JVT のテストモデルでは、各 MB における符号化モードや動きベクトル等を符号化効率の観点で最適な形で選択するために Rate-Distortion 最適化法 (以下「RDO 方式」と書く。) がとられている。これは選択可能な符号化モードおよび動きベクトルに対し、実際に必要となるビット R とその際の歪 D (あるいはそれらを模擬したもの) をラグランジェ乗数  $\lambda$  を用いて線形結合し、これが最小になるケースを選択するものである。Intra\_4x4 において予測モードを選択する場合にもこれが用いられており、評価関数  $J = D + \lambda R$  (以下、「RDCost 関数」) を最小とするモードが選択される。

なお、この値を求めるために各モードに対して、予測誤差の計算、変換、量子化、ビット数計算、逆量子化、逆変換、量子化誤差計算が必要となる。

なお、本方法は JVT のテストモデルである JM[2] でも用いられている。

### 3.2. Edge 成分に基づく方法

最適なイントラ予測モードの予測方向は当該ブロックの edge 方向にほぼ一致するとの考えに基づく方法である（文献[3]：以下「edge 方式」と書く）。

まず、画面全体の各画素に Sobel オペレータを適用した Edge map と呼ばれるものを作成する。これは、各画素に水平・垂直方向にそれぞれ適用されて得られた Sobel オペレータ値(Dx,Dy)に対し  $A=|Dx|+|Dy|$ 、 $\theta=\arctan(Dx/Dy)$  とすることにより得られる。

その後、イントラ予測方向の候補数 X (DC 予測は除く) に基づき、全角度を X 個に分け、 $\theta$  が属するクラスのヒストグラムとして A を追加する。これを着目するブロック内の各画素に対し行い、最終的にもっともヒストグラムが大きくなるクラスを当該ブロックの予測方向として取り扱うこととする。なお、Intra\_4x4 の場合はその両近傍の方向も候補として扱う。

最後に、上記で候補となったモード (Intra\_4x4 の場合は 3 つ、その他は 1 つ) と DC 予測モード全てに対し、RDO 方式を適用、すなわち、RDCost の計算を行い、それが最小になるモードを選択する。なお、両近傍の方向を候補として取り扱わない場合、符号化効率率が低下することが同文献においてわかっている。

### 4. 提案手法

以下の 2 つの手法を提案する。

1. JVT においても、従来方式同様、予測誤差が小さい方が一般的に符号量が少なくすむ。そこで、すべてのイントラ予測モードについて予測誤差を求め、その予測誤差の絶対値和が最小になるモードを選択すればよい（以下「SAD 方式」と書く）。本来であれば、変換係数の大きさをそのまま反映している誤差電力が最小になるモードを選択すべきである（これを「SSD 方式」と書く）が、計算量の観点から SAD 方式のほうが簡易であるため、両方の方式について性能を確認することとする。
2. 2 節で述べたとおり、あるブロックの予測モードとして上あるいは左のブロックのうち小さいモード番号がそのまま継承された場合は、UVLC の場合 3bit の符号量削減となる。一般的に、特に低レートになればなるほど、このようなオーバーヘッド部のビットが大きな比重を持つようになるので、このモードも選択肢の一つに加えることが好ましいと考えられる。それゆえ、上または左のブロックのモード番号のうち、小さい方を予測モー

ドとする（以下、「MPM」方式と呼ぶ）。なお、本方式を単独で適用すると、最上端、最左端で適用されたモードがそのまま継承されることになるため、特定の場合を除き符号化効率は明らかに悪化する。よって、本方式を単独で用いることはしない。

上記 1, 2 で選択される 2 つの予測モードに RDO 方式を適用し、RDCost が小さくなる予測モードを選択する手法も検討に加える（「SSD+MPM 方式」、「SAD+MPM 方式」と書く）。なお、1, 2 で選択されるモードが重なる場合は RDO 方式を適用する必要がない。

### 5. シミュレーション実験

前節の提案法の性能を検証すべく計算機シミュレーション実験を行った。実験の条件は標準化の際に利用されていた文献[5]を利用した。実験に用いるパラメータを表 1 に示す。

表 1: 実験パラメータ

項目	内容
シーケンス	Foreman(QCIF,10Hz,100frame), News(QCIF,10Hz,100frame), Mobile&Calendar(CIF,30Hz,100frame)
MV 探索範囲	±32
Hadamard	Used
RDO optimization	On
QP	16,20,24,28(high bitrate)/ 28,32,26,40(low bitrate)選択
MultiReference	1 frame only
GOP 構造	IPPP...
エントロピー	UVLC/CABAC 選択
Constrained	Off
IntraPred	

まず、文献[4]の方法に基づき、3.4 節で述べた方式について RDO 方式(3.1 節)に対するビット増加率(BDBR)で比較を行った。これをエントロピー符号化別に表 2, 3 に示す(それぞれ、UVLC+CAVLC,CABAC の場合の結果である)。この値が大きいほど、RDO 方式の結果と比較して符号化効率が悪化していると言う意味であるが、表 2 から、SAD+MPM 方式は Edge, SSD 両方式と比較し平均 1%弱程度良く、かつ、SSD+MPM 方式と大差ないことがわかる。さらに、表 3 から UVLC を用いたときだけでなく、CABAC を用いたときも同様のことが言える。

また、最初の I フレームのみで同様の比較を行った(表 4, 5)。その結果、SAD+MPM 方式は、すべてのシーケンスで edge, SSD の両方式と比較し約 5%程度ビット削減できていることがわかる(5%は約 0.25dB に相当する)。一方で、SAD+MPM 方式と SSD+MPM 方式は大差がないことも同様に確認できる。

MPM を加えることによる効果であるが、表 4, 5 か

ら、ビットレートが低いときほどその差が顕著に現れることがわかる（低ビットレート：5~6%，高ビットレート：3%程度）。また、表4の Foreman について、ビットレート-PSNR 特性をグラフに書き出してみると図3のようになるが、ビットレートが高いときは RDO 方式と SAD+MPM 方式の間に差がある一方でビットレートが低くなるとその差はほとんど0になる。これは、前節で述べたとおり、特に低レートになればなるほど、オーバーヘッド部のビットが大きな比重を持つようになることから、MPM 方式により選択される予測モードも候補の一つすることによる各4x4ブロックでの3bit削減の効果が現れていると言えよう。

一方、計算時間の観点から、提案方式の有効性を検証してみる。RDO 法を適用する Intra 予測モード数を比較すると、RDO 方式：9モード、edge 方式：4モードに対し、SSD+MPM 方式、SAD+MPM 方式は2モードあるいは0モードである。Edge 方式では事前にエッジ成分を計算する必要がある、これは予測誤差(SSD,SAD)を計算するために要する時間とほとんど変わらない。

表2. BDBR(UVLC) [%]

Seq.	Foreman		News		Mobile&Cal.	
	Low	High	Low	High	Low	High
Edge	1.74	0.83	1.05	1.12	0.77	0.08
SSD	1.96	0.93	2.24	1.23	0.56	0.09
SAD+MPM	1.03	0.49	0.83	0.90	0.46	0.03
SSD+MPM	0.74	0.33	0.82	0.80	0.45	0.04

表3. BDBR(CABAC) [%]

Seq.	Foreman		News		Mobile&Cal.	
	Low	High	Low	High	Low	High
Edge	1.05	0.76	1.32	1.37	0.38	0.10
SSD	1.02	0.82	1.45	1.23	0.22	0.07
SAD+MPM	0.57	0.48	1.20	1.00	0.06	0.05
SSD+MPM	0.41	0.46	1.16	0.98	0.00	0.02

表4. I frame の BDBR(UVLC) [%]

Seq.	Foreman		News		Mobile&Cal.	
	Low	High	Low	High	Low	High
Edge	7.77	5.07	5.81	5.11	5.38	3.58
SAD	9.06	5.91	6.99	5.34	5.66	5.06
SSD	7.90	5.36	5.92	4.93	4.88	4.82
SAD+MPM	2.45	2.56	2.17	2.39	2.19	2.67
SSD+MPM	2.03	1.86	1.32	2.07	1.87	2.53

表5. I frame の BDBR(CABAC) [%]

Seq.	Foreman		News		Mobile&Cal.	
	Low	High	Low	High	Low	High
Edge	8.82	5.91	5.79	4.98	5.48	3.56
SAD	9.36	6.02	6.80	4.87	5.66	4.34
SSD	8.35	5.38	5.73	4.52	4.88	3.84
SAD+MPM	3.66	3.02	1.79	1.79	2.19	2.37
SSD+MPM	3.06	2.62	1.58	1.43	1.87	2.10

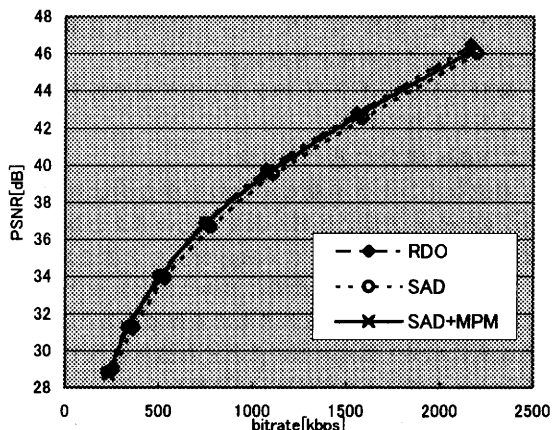


図3. Bitrate-PSNR 特性 (Foreman, I frame, UVLC)

なお、表4の実験で SAD+MPM 方式において予測誤差絶対値和が最小になる予測モードと MPM で求まるモードが重なる割合は、Foreman で約 28±2%，News で 45±2%，Mobile&Calendar で 33±2%と全体の 3,4 割を占めていることが確認できる。なお、この割合はエントロピー符号化の違いや QP の違いにより大差ないことも確認した。これは絵の性質のみに依存すると考えられる。

## 6. おわりに

本稿では、JVT のイントラ予測に着目し、その高速化に関する検討を行った。予測誤差の絶対値和、あるいは、電力が最小となるような予測モード、および、左か上のブロックから継承された予測モードに対し評価関数を適用することで、従来とほぼ同等の符号化効率を保ちつつ（ビット増加率は 1%程度）、より高速に予測モードを選択できることが確認できた。

今後は Intra\_16x16 と Intra\_4x4 の高速適応選択法、ならびに、全マクロブロックモードの高速適応選択法について検討を進めていく。

## 文献

- [1] ITU-T Rec. H.264 | ISO 14496-10 AVC, 2003.
- [2] Gary S., Thomas W. and Keng-Pang L. "Joint Model Reference Encoding Methods and Decoding Concealment Methods," JVT-1049, Sept.2003.
- [3] Feng P. et al, "Fast Mode Decision for Intra Prediction," JVT-G013, Mar.2003.
- [4] Gary Sullivan, "Recommended Simulation Common Conditions for H.26L Coding Efficiency Experiments on Low Resolution Progressive Scan Source Material," ITU-T VCEG-N81, September 24-27, 2001.
- [5] Gisle Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," ITU-T VCEG-M33, April, 2001.
- [6] 清水, 高木, 柳原, 中島, "JVT と MPEG-4 の圧縮符号化効率の比較について", PCSJ2002, 2002.11.