

周波数特性に基づく H.264/AVC イントラ予測モード決定手法に関する検討

筑波 健史[†] 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 〒169-0051 東京都新宿区西早稲田 1-3-10

^{††} 早稲田大学 国際情報通信研究センター 〒169-0051 東京都新宿区西早稲田 1-3-10

^{†††} 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: †{tsukuba,isao,hana,tominaga}@tom.comm.waseda.ac.jp

あらまし 本稿では、H.264/AVC の符号化における処理量を低減するために、 4×4 イントラ予測モードの効果的な選択手法を提案する。イントラ予測において、原信号を周波数領域で見たときに、エネルギーが多く集中した領域を低減させることが予測効率の向上に有効であると考えられる。そこで提案手法では、変換係数の低周波成分の係数レベルを特徴量として、周波数分布の性質を 8 種類のパターンに分類し、分類結果に基づき評価対象の予測モードを限定することで処理量を低減する。提案手法を参照ソフトウェアに実装し、PSNR、符号量、処理量の観点から評価を行った。実験結果より、提案手法は参照ソフトウェアと比較して符号化効率をほとんど低下させずに、処理量を 6 割程度削減できることが確認できた。

キーワード H.264/AVC、フレーム内符号化、イントラ予測、モード決定

H.264/AVC Intra-prediction Mode Decision based on Frequency Characteristic

Takeshi TSUKUBA[†], Isao NAGAYOSHI^{††,†††}, Tsuyoshi HANAMURA^{††,†††}, and Hideyoshi

TOMINAGA[†]

[†] GITS, WASEDA University, 29-7 building 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo 169-0051 Japan

^{††} GITI Waseda University 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan

^{†††} Media Glue Corp. Ramdax Bldg. 8th floor, 2-4-12 Ohkubo, Shinjuku-ku, Tokyo, 169-0072 Japan

E-mail: †{tsukuba,isao,hana,tominaga}@tom.comm.waseda.ac.jp

Abstract In this paper, we propose efficient mode decision schemes for intra 4×4 prediction mode to reduce the complexity of the H.264/AVC encoder. We consider that reducing the area where much more energy exists on the DCT domain may improve the efficiency of intra-prediction. In our proposed schemes, the encoder complexity is reduced by limiting the prediction modes with result that the frequency distribution is categorized into 8 parts according to the features extracted from the low-frequency components on the DCT domain. Our proposed schemes are evaluated from viewpoints of PSNR, Bitrate, and the encoder complexity. Experimental results show that the proposed schemes reduce the complexity by nearly 60 % while maintaining similar PSNR and Bitrate.

Key words H.264/AVC, intra-frame coding, intra-prediction, mode decision, encoder complexity

1. はじめに

H.264/AVC のイントラ符号化では、新たに符号化ツールとして空間予測が加わり、9 種類の 4×4 ブロックサイズの予測方法 (以下、 4×4 イントラ予測) と 4 種類の 16×16 ブロックサイズの予測方法 (以下、 16×16 イントラ予測) が採用されている。 4×4 イントラ予測では、図 1(a) に示す現ブロックの各画素 a~p は、隣接画素 A~M を用いて予測符号化される。予測方法は、図 1(b) に示す方法が定義されている。ただし、矢印は予測方向のモード番号であり、モード 2 は DC 予測である。

参照ソフトウェア JM [1] では、全てのモードに対して、レート歪み最適化方式 (以下、RDO 方式と呼ぶ) を適用し最適なモードの評価を行う。これはすべてのモードについて符号化処理を行うため、多大な処理量を必要とする。イントラ符号化処理は全体の約 20% を占めるため、効率的にイントラ予測モードを決定する手法を開発することが望まれる。

そこで本稿では、H.264/AVC の符号化における処理量低減を目的として、周波数特性に着目した 4×4 イントラ予測モード決定手法を提案する。

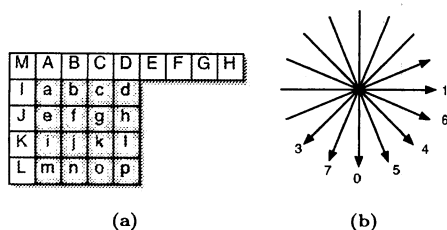


図 1 4×4 イントラ予測

2. 従来手法

4×4 イントラ予測のモード選択手法は、大別して以下の 2 種類に分類される。

- 簡易なコスト計算によるモード決定手法
- 信号特性に基づくモード決定手法

前者は、コスト計算の簡略化により高速化を実現する [2]。文献 [3] では、予測誤差電力が最小になる予測モードと、および隣接ブロックの予測モード番号が小さいモード (*most_probable_mode*) の 2 つのモードに対して RDO 方式を適用する手法が提案されている。同手法では 2 種のモードが一致する場合、RDO 方式を適用せずにモードが決定可能である。また 2 種のモードが一致する確率は、シーケンスによるが、30~50% であると報告されている。

後者は、最適な予測方向は対象ブロックのエッジ成分

と相関があるという考えに基づいた高速化手法である。文献 [4] [5] では、ブロック内の各ピクセルに対してエッジ成分 (dx, dy) を算出し、度数分布から支配的なエッジ成分を求める。その結果から場合分けを行い、RDO 方式を適用する候補を限定することで処理量を低減している。同手法は、ピクセル単位でフィルタ処理を行うため、特徴量を算出するのに必要な計算処理量が多い。一方、文献 [6] の手法は、ピクセル単位ではなく、 4×4 ブロックを単位として勾配を算出し、その結果からモードを限定する。同手法は、垂直・水平方向の勾配度を表す F_v, F_h に基づき、数種類の代表的な分布パターンに分類する。場合分けに対応した候補に対して RDO 方式を行うことで処理量を低減している。

本稿では、原画像信号を 4×4 DCT 変換した結果を利用したモード選択手法を提案する。原信号を周波数領域でみたときに、エネルギーが多く集中した領域を低減させることで効率的な予測が可能であると考えられる。そこで提案手法では、低周波成分の係数レベルを用いて特徴量を算出することで周波数分布の性質を 8 種類のパターンに分類し、分布パターンによって評価対象の予測モードの候補数を削減することで処理量を低減する。

3. 周波数特性に基づく特徴量の算出法

本章では、提案手法において、原信号を 8 種類の分布パターンに分類するときに利用する特徴量 E_v, E_h, C_v, C_h について述べる。

3.1 整数変換

H.264/AVC では、従来の動画像符号化方式で採用されている 8×8 ブロックサイズによる DCT 変換の代わりに、 4×4 ブロックサイズによる DCT 変換が使われる。H.264/AVC で採用された DCT 変換は次式で与えられる [7]。

$$\mathbf{Y} = \mathbf{C}_f \mathbf{X} \mathbf{C}_f^t \quad (1)$$

$$\mathbf{C}_f = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{pmatrix} \quad (2)$$

$$\mathbf{X} = [X_{ij}], \mathbf{Y} = [Y_{ij}], (i, j = 0, 1, 2, 3) \quad (3)$$

ここで、 $\mathbf{X}, \mathbf{Y}$ はそれぞれ、原信号、変換係数である。提案手法では、原信号の変換係数の低周波成分に着目して、特徴量を算出する。

3.2 特徴量 E_h, E_v, C_h, C_v の定義

周波数特性に基づく特徴量として、次の 4 変数を定義する。

- C_h : 水平方向の周波数成分
- C_v : 垂直方向の周波数成分
- E_h : 水平方向のエネルギー強度
- E_v : 垂直方向のエネルギー強度

C_h, C_v はそれぞれ DC 成分に隣接する水平・垂直方向の低周波数成分を表し、式 (4) で与えられる。

$$C_h = \left\lfloor \frac{Y_{01}}{S} \right\rfloor, C_v = \left\lfloor \frac{Y_{10}}{S} \right\rfloor \quad (4)$$

また E_h, E_v はそれぞれ水平・垂直方向の低周波数成分のエネルギー強度を表し、式 (5) で与えられる。

$$E_h = \left\lfloor \frac{\sum_{j=1}^3 Y_{0j}^2}{S^2} \right\rfloor, E_v = \left\lfloor \frac{\sum_{i=1}^3 Y_{i0}^2}{S^2} \right\rfloor \quad (5)$$

ここで、 S はスケール因子であり、 $\lfloor \cdot \rfloor$ は丸め込み操作を表す。 C_h, C_v により、水平・垂直それぞれの方向の勾配を推定し、 E_h, E_v により水平・垂直方向のどちらかへのエネルギーの偏りを評価し、両者を合わせて分布パターンを推定する。

4. 提案手法

4.1 アルゴリズム

提案手法のフローを図 2 に示す。

A. 周波数特性に基づくモード選択

- (1) E_v, E_h および、 C_v, C_h を計算する。
- (2) E_v, E_h および、 C_v, C_h によって、表 1 に示す 8 種類の分布パターンに場合分けを行う。
- (3) もし $Case = 6, 7$ なら (B) へ進む。それ以外ならば、(A)-(4) へ進む。
- (4) もし $Case = 2, 5$ なら DC 予測 (モード番号 2) を選択して終了する。それ以外ならば、(C) へ進む。

B. 簡易コスト計算によるモード選択

- (1) 各モードにおけるコストを SAD あるいは SATD により求める。
- (2) コストの低い順に予測モードをソートする。
- (3) コストの低い上位 1~N 番目までのモードを選択する。
- (4) 閾値処理により候補数を適応的に 1~N 個に削減する。

C. レート歪み最適化手法によるモード決定

- (1) 各モードにおけるコストを RDO 方式により算出する。
- (2) コストが最小となるモードを選択する。

以下に、上記 B, C の詳細を述べる。

表 1 分布パターンと予測モードの対応表

Case	分布パターン	予測モードの候補
0	$E_v = E_h = 0$	2
1	$E_h = 0, E_v > 0$	0, 2
2	$E_v = 0, E_h > 0$	1, 2
3	$E_v = E_h > 0, C_v \times C_h < 0$	3, 2
4	$E_v = E_h > 0, C_v \times C_h > 0$	4, 2
5	$E_v = E_h > 0, C_v \times C_h = 0$	2
6	$E_v > E_h > 0$	0, 3, 4, 5, 7, 2
7	$E_h > E_v > 0$	1, 3, 4, 6, 8, 2

4.2 簡易コスト計算によるモード選択

(A) において、分布パターンが Case 6, 7 に選択された場合、(B) の簡易コスト計算によって予測モードの候補数 1~N 個に限定する。以下に詳細を示す。

各モードに対して、コスト関数を式 (6) のように定義する [1]。

$$Cost_{Mode} = SA(T)D + Const. \quad (6)$$

$SAD^{(註1)}$ 、および $SATD^{(註2)}$ は次式で与えられる。

$$SAD = \sum_{i=0}^3 \sum_{j=0}^3 |X_{ij} - P_{ij}| \quad (7)$$

$$SATD = \sum_{i=0}^3 \sum_{j=0}^3 |T\{X_{ij} - P_{ij}\}| \quad (8)$$

ここで、 P_{ij} は予測信号、 $T\{\cdot\}$ はアダマール変換を表す。ただし $Const.$ は、ヘッダビット分のコストを考慮したものであり、式 (9) で与えられる。

$$Const. = \begin{cases} 0 & (mode = mpm) \\ \lfloor 4 \times \lambda \rfloor & (mode \neq mpm) \end{cases} \quad (9)$$

$$\lambda = 0.85 \times 2^{(QP-12)/3} \quad (10)$$

ここで、 mpm は、左または上ブロックのモード番号が小さいモード (*most_probable_mode*) を表す。

本ステップでは、表 1 に示す Case 6, 7 の各候補について、式 (6) によりコスト評価を行い、値の低い順に並び変える。並べ替え後のコストを $Cost(i)$ 、対応する予測モード番号を $ipmode(i)$ として、コストの低い上位 N 個のモードに対して、 $ipmode(0)$ とおよび、

$$Cost(i) < Cost(0) + Cost_{thr}, (1 \leq i \leq N-1) \quad (11)$$

を満たすモード $ipmode(i)$ を候補とする。ここで、 $Cost(0)$ は最小コストのモードである。

以上のような制御によって、候補となるモード数を適応的に 1~N 個に限定する。

(注1) : the sum of absolute differences

(注2) : the sum of absolute transform coefficient differences

4.3 レート歪み最適化手法によるモード決定

Cでは、候補のモードに対して、RDO方式を適用して、最終的な予測モードを決定する。以下では、参照ソフトウェア[1]で実装されているRDO方式について簡単に述べる。

4.3.1 Rate-Distortion Optimization

参照ソフトウェアでは、多数ある符号化ツールから最適なツールを選択するためにレート歪み最適化手法が使われている[8]。RDO方式は、MB毎に各符号化モードに対して仮符号化処理を行い、得られる符号量 R と符号化歪み D から、ラグランジュ未定乗数法を用いて、最適なモードを選択する手法である。これは、次式のコスト関数 J を最小化するモードを選択することである。

$$J = D + \lambda R \quad (12)$$

ここで、 λ はラグランジュ定数であり、式(10)に示す通りである。

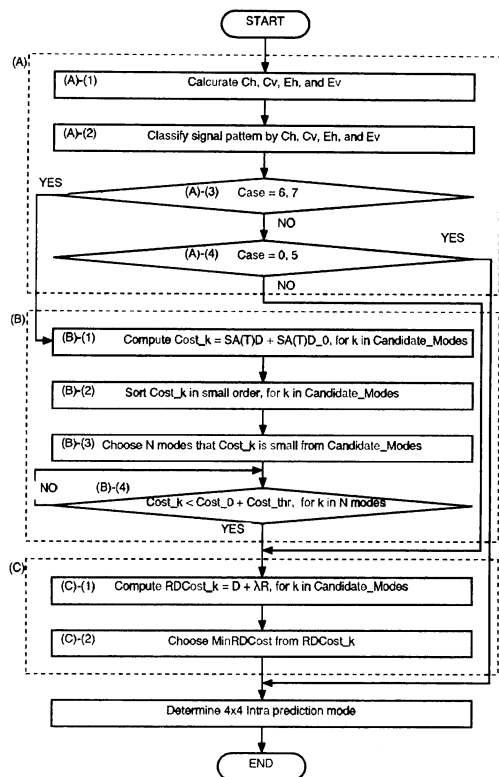


図2 フローチャート

5. シミュレーション実験

5.1 実験方法

提案手法の有効性を検証するため、実験を行った。画

質、発生符号量、符号化処理時間の3つの項目について評価する。比較方式として、JM方式(イントラ予測4×4のみ)と文献[6]の方式(FMS方式と呼ぶ)、および提案手法を用いる。比較方式を従来手法(JM, FMS)も含めて表2に示す。ここで、方式Cはフローチャート(図2)において(b)における簡易コスト計算によって候補数を削減する手法を適用せず、表1の全候補について、4.3を適用する。また符号化条件と使用したPC環境をそれぞれ、表3、表4に示す。

表2 比較方式

方式	内容	簡易コスト計算方法
A	提案方式	SATD
B	提案方式	SAD
C	提案方式	なし
D	JM方式[1]	-
E	FMS方式[6]	-

表3 符号化条件

項目	内容
シーケンス	Foreman(QCIF), News(QCIF), Mobile&Calender(CIF), Tempete(CIF)
フレーム数	150
フレームレート [frames/sec]	15(QCIF), 30(CIF)
QP	20,24,28,32(固定)
GOP 構造	All Intra-frame
RDOOptimization	On
エントロピー符号化	CAVLC
イントラ予測	4×4イントラ予測のみ
Codec	JM 8.5
パラメータ	$S = 16$, $Cost_{thr} = 256$, $N = 3$

表4 実験用PCの環境

CPU	Pentium 4 3.2[GHz]
メモリ	DDR SDRAM 1024[MByte]
OS	Vine Linex 2.6 CR(kernel 2.4.22)

5.2 実験結果

JM方式に対する各手法の符号化処理時間、PSNR、符号量の差分を Δ Complexity, Δ PSNR, Δ Bitrateと定義し、結果を表5に示す。また、各テストシーケンスにおける各手法のレート対歪み特性を図3に示す。

符号化時間について、表5を参照すると、シーケンスおよびQPによらず削減できた処理量は、方式Bが最も多く、続いて方式A、方式E、方式Cの順であることがわかる。提案手法に注目すると、方式Cは、全体の約

表 5 Results of Sequences

(a) Foreman

Method	Δ Complexity(%)	Δ PSNR(dB)	Δ Bitrate(%)
Method A	-57.377	-0.078	2.081
Method B	-57.786	-0.098	2.468
Method C	-31.638	-0.030	2.001
Method E	-41.718	-0.078	4.782

(b) News

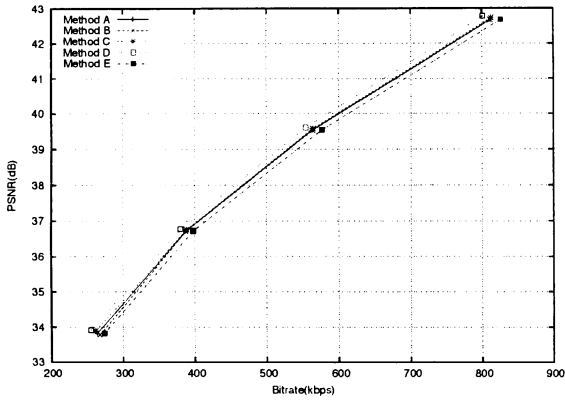
Method	Δ Complexity(%)	Δ PSNR(dB)	Δ Bitrate(%)
Method A	-59.352	-0.140	2.434
Method B	-59.445	-0.145	2.669
Method C	-35.947	-0.053	2.430
Method E	-43.021	-0.088	5.531

(c) Mobile & Calendar

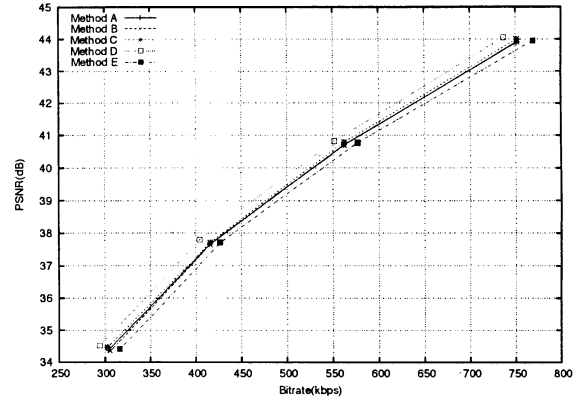
Method	Δ Complexity(%)	Δ PSNR(dB)	Δ Bitrate(%)
Method A	-57.141	-0.235	0.971
Method B	-57.383	-0.228	1.188
Method C	-28.520	-0.065	1.048
Method E	-36.752	-0.115	2.563

(d) Tempete

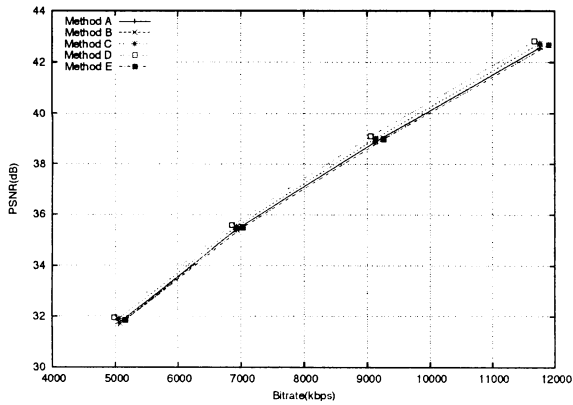
Method	Δ Complexity(%)	Δ PSNR(dB)	Δ Bitrate(%)
Method A	-56.491	-0.183	1.090
Method B	-57.264	-0.185	1.350
Method C	-29.833	-0.063	1.186
Method E	-38.228	-0.115	2.866



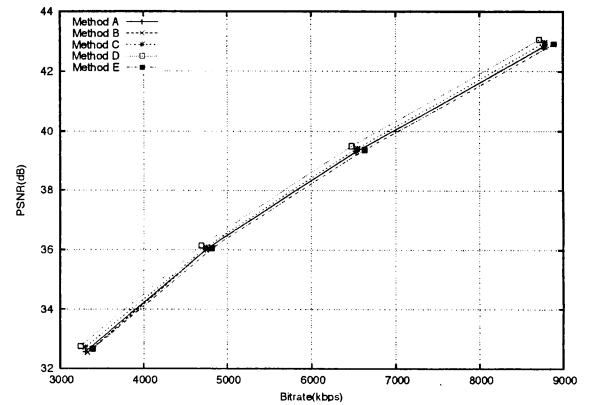
(a) Foreman



(b) News



(c) Mobile & Calendar



(d) Tempete

図 3 レート対歪み特性

30%削減し、方式 A, B は、全体の約 60%削減している。特徴量 C_h , C_v , E_h , E_v による分類を適用することで、Case0～5 では、1～2 個、Case6～7 では、6 個の候補数に評価対象が限定される。これにより 30%の処理量の削減が可能となる。また case6～7 の場合は、簡易コスト計算によるモード選択手法によって、候補数を 6 個から 1～3 個に減らしているため、さらに 30%の処理量の削減が可能となる。RDO 方式を適用する回数についても、提案手法によって約 75%削減できることがわかっていく。

図 3 のシーケンス毎のレート歪み特性を比較すると、方式 D に対する各手法の符号化効率の低下わずかである。中でも方式 C は、結果が一番良く、方式 A, 方式 B, 方式 E の順となる。特に QCIF サイズのシーケンス (Foreman や News) では、 $QP \leq 28$ では、方式 A と方式 C はほぼ一致し、 $QP > 28$ で差が見られる。平均して簡略化による PSNR の劣化は、方式 C では、約 $-0.05dB$ となり、方式 A, B では、約 $-0.1 \sim -0.2dB$ 、方式 E は $-0.1dB$ となる。また符号量の増加は、シーケンスによらず方式 A, B, C は 1.0～2.0%、方式 E は 2.0～5.0% となる。

以上の結果をまとめると、提案手法では PSNR は $-0.1dB$ 程度の低下、符号量は 1.0～2.0% 程度の増加で、処理量を 60% 削減可能となり、提案手法の有効性が示せた。また提案手法の中では、符号化効率と処理量のトレードオフから方式 A が最も結果が良いと考えられる。

6. ま と め

本稿では、H.264/AVC における 4×4 イントラ予測を対象として、エネルギー分布に着目したモード選択手法に関して検討を行った。提案手法では、変換係数の低周波成分の係数レベルに基づき、8 種類の分布パターンに分類し、分布パターンによって評価対象の予測モードを選択することで、従来方式とほぼ同等の符号化効率を保ちつつ、処理量を約 60%ほど削減することが可能であることが確認できた。今後は提案手法のさらなる符号化効率の改善および高速化を検討していく。

文 献

- [1] Joint Video Team(JVT), "Reference Software JM8.5", <http://iphome.hhi.de/suehring/tml/>
- [2] Meng,Wong, Lam,"Efficient intra-prediction mode selection for 4x4 blocks in H.264",IEEE Int. Conf. Multimedia & Expo, 2003
- [3] 高木, 滝嶋, 中島, "H.264/AVC におけるイントラ予測高速化の一検討," FIT, 2003
- [4] F. Pan, X. Lin, R. Susanto, "Fast Mode Decision

for Intra Prediction, " JVT-G013, 7th Meeting: Pattaya II,Thailand, 7-14 March, 2003

- [5] F. PAN, "A DIRECTIONAL FIELD BASED FAST INTRA MODE DECISION ALGORITHM FOR H.264 VIDEO CODING, " IEEE, 2004
- [6] Z. Yong-dong, D. Feng, L. Shou-xun, "Fast 4×4 Intra-prediction Mode Selection for H.264, " IEEE, 2004
- [7] Iain E. G. Richardson, "H.264/MPEG-4 Part 10 White Paper: Transform & Quantization, " www.vcodex.com, May 2003
- [8] G. Sullivan, T. Wiegand, "Rate-distortion optimization for video compression, " IEEE Signal Proc. Magazine, 1998
- [9] H. S. Malvar, "Low-Complexity Transform and Quantization, " IEEE Trans. on Circuit & Syst. for Video Tech., VOL. 13, NO. 7, JULY 2003