

I-056

# JPEG/H.264 トランスコーダにおけるイントラ予測モード決定手法

## Intra Prediction Mode Decision for JPEG/H.264 Transcoding

草野 勝大†  
Katsuhiko Kusano

工藤 大樹†  
Daiki Kudo

西川 博文†  
Hirofumi Nishikawa

和田 稔†  
Minoru Wada

加藤 嘉明†  
Yoshiaki Kato

### 1. はじめに

監視映像やデジタルカメラにおいて、JPEG を連続して表示することで動画とする Motion-JPEG 形式が採用されている。JPEG などの画像符号化方式では、離散コサイン変換(DCT)が用いられている。一方、動画画像符号化方式の一つである H.264[1]は、画像領域上で予測を行うイントラ予測が採用されており、JPEG に比べて高い符号化効率を実現可能である。上記の理由から、ストレージ量の削減、長時間コンテンツの長期間保存などのため、Motion-JPEG を H.264 へと変換したいという需要があると予想される。しかし、JPEG から H.264 へと変換する際に、イントラ予測モード決定のために多くの計算量が必要となってしまう。本稿では、JPEG の DCT 係数情報を利用した、イントラ予測モード決定手法を提案し、その性能を評価する。

### 2. JPEG/H.264 トランスコーダ

図1は、JPEG/H.264イントラトランスコーダの構成を示したものである。これは JPEG デコーダと H.264 イントラエンコーダを直列接続したものである。図1は最も単純な構成で、全ての予測モードに対してコストを計算することで、H.264 イントラエンコーダにおける、最も少ない符号量での符号化が可能となるが、処理量が膨大となってしまう。

イントラ予測は画像空間における予測方法であり、9種類の4×4ブロック単位の予測方式(以下、4×4イントラ予測)と16×16ブロック単位の予測方式(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2(a)に示すように符号化対象ブロックの各画素 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2(b)に示すように予測方向のモードに従って定義されている。モード2はDC予測である。

JPEG の DCT 係数は画像の空間的特徴を示しており、イントラ予測モード決定の一つの指標になると考えられる。[2]や[3]などでは、DCT 係数などの周波数成分を利用し、イントラ予測モードの決定における処理量を低減する手法が検討されている。16×16ブロック内の周波数成分のみを利用しており、ブロック周辺の被予測画素については検討されていない。本稿では、被予測画素の周波数成分を利用した、高精度かつ高速なイントラ予測モード決定手法を提案する。

### 3. 提案手法

本章では、JPEG デコード時に取得した DCT 係数及び画素値から、イントラ予測モードを決定する方法について述べる。図3は本手法の JPEG/H.264 イントラトランスコーダを示したもので、イントラ予測モード決定部において

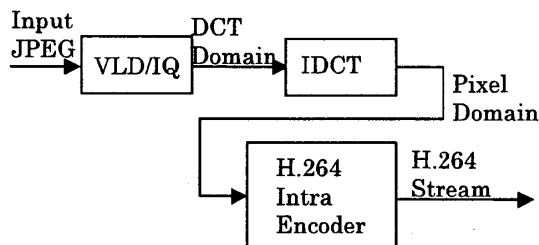


図1 JPEG/H.264イントラトランスコーダ

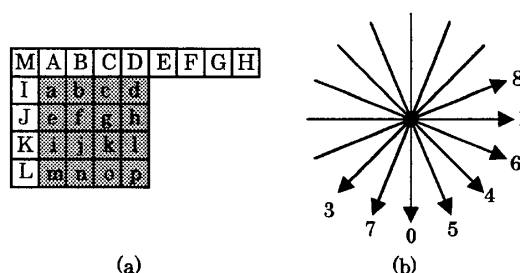


図2 4×4イントラ予測

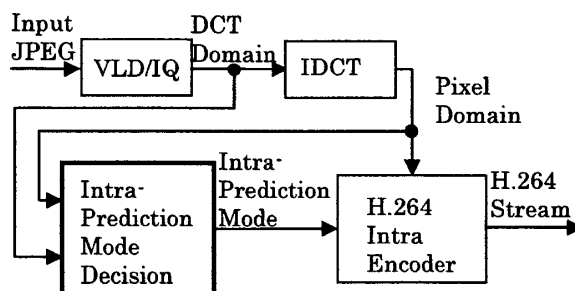


図3 JPEG/H.264イントラトランスコーダ(提案手法)

H.264エンコーダで用いる予測モードを決定する。

まず、図4に示すような、符号化対象16×16ブロックの周辺ブロックのうち、上部ブロックの最下部の垂直周波数成分と左部ブロックの最右部の水平周波数成分を算出する。

算出した垂直周波数成分、水平周波数成分と16×16ブロック内のDCT係数のDC成分、垂直周波数成分、水平周波数成分より直流成分コスト  $C_{DC16}$ 、垂直成分コスト  $C_{V16}$ 、水平成分コスト  $C_{H16}$  を算出する。

16×16イントラ予測モードの決定は以下の通りである。

- モード0:  $C_{V16} \leq T_{V16}$ ,  $C_{H16} > T_{H16}$
- モード1:  $C_{V16} > T_{V16}$ ,  $C_{H16} \leq T_{H16}$
- モード2:  $C_{DC16} \leq T_{DC16}$ ,  $C_{V16} \leq T_{V16}$ ,  $C_{H16} \leq T_{H16}$
- モード3:  $C_{DC16} > T_{DC16}$ ,  $C_{V16} \leq T_{V16}$ ,  $C_{H16} \leq T_{H16}$
- 4×4イントラ予測:  $C_{V16} > T_{V16}$ ,  $C_{H16} > T_{H16}$

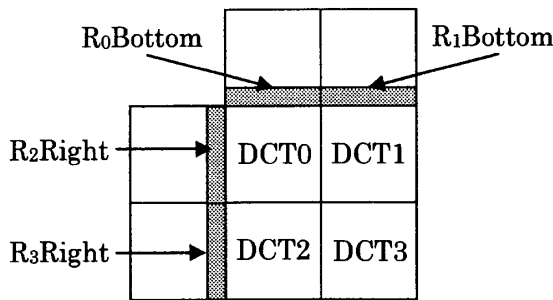


図4 16×16ブロックと周辺8×8ブロック

4×4 イントラ予測が選択された場合は、図2の a~p の画素に、周辺画素である A~D, I~M を加えた 5×5 ブロックを新しく構成する。この 5×5 ブロックに式(1)の 5×5 の DCT を施し、5×5 ブロックにおける DCT 係数を算出する。

$$F(u, v) = \frac{2}{5} C(u) C(v) \sum_{x=0}^4 \sum_{y=0}^4 f(x, y) \cos \frac{\pi(2x+1)u}{10} \cos \frac{\pi(2y+1)v}{10}$$

$$C(u), C(v) = \begin{cases} 1/\sqrt{2} & (u, v = 0) \\ 1 & (u, v \neq 0) \end{cases} \quad \dots(1)$$

算出した 5×5 DCT 係数よりブロック内のエッジの角度  $\theta$  を求める。

$$E_{V5} = \sum_{j=1}^4 F(0, j), \quad E_{H5} = \sum_{i=1}^4 F(i, 0) \quad \dots(2)$$

$$\theta = \arctan(E_{V5}/E_{H5}) \quad \dots(3)$$

$$\theta = \begin{cases} \theta & F(0,1)F(1,0) \geq 0 \\ -\theta & F(0,1)F(1,0) < 0 \end{cases} \quad \dots(4)$$

$E_{V5}$ ,  $E_{H5}$  が両者とも小さい場合は、モード 2 を選択し、それ以外の場合は、図 2 の 4×4 予測モードのうちエッジの角度  $\theta$  に最も近いモードとその両隣のモードを選択する (モード 1 が選択された場合は、1,6,8)。予測モードは、画像内に同じ方向のエッジがあったときに、そのモードにおいて高い予測効率が得られる可能性が高くなるので、上記処理により、高効率なイントラ予測モード決定が可能となる。選択された 3 モードの予測コストを計算し、最も予測コストが小さいものを 4×4 ブロックの予測モードとする。

#### 4. 評価実験

提案手法の有効性を確認するため、計算機シミュレーションを行い、画質、符号量、H.264 エンコード処理時間(予測モード決定時間を含む)の 3 つの項目について評価した。画質は、JPEG デコード後の画像と H.264 符号化後の画像との比較で評価する。入力 JPEG は CIF サイズの YUV 画像 "Foreman", "Football", "Mobile&Calender" を量子化レベル 50 で JPEG 符号化したもので、フレーム数は 100 である。H.264 の符号化条件は表 1 に示す。実験に用いた PC の環

表 1 H.264 符号化条件

プロファイル	Baseline Profile
QP	28, 32, 36, 40 (固定)
GOP 構造	ALL Intra-Frame
RDOptimization	On

表 2 シミュレーション結果

Sequence	$\Delta$ Bitrate [%]	$\Delta$ PSNR [dB]	$\Delta$ Time [%]
Foreman	8.66	-0.22	-55.4
Football	5.79	-0.21	-51.9
Mobile&calender	4.30	-0.18	-51.1

境は、Intel Core 2 Duo E6850 3.0GHz, WindowsXP, Visual Studio 2005 である。

図 1 において、H.264 イントラエンコーダに JM14.0[4] を用いた構成との比較を行った。結果を表 2 に示す。 $\Delta$ Bitrate,  $\Delta$ Time は、JM を基準とした場合の増加割合の平均値を、 $\Delta$ PSNR は変化差分の平均値を表している。

結果より、JM に対して 50%以上処理量を削減できていることがわかる。PSNR に関しても、0.2dB 前後とほとんど低下しておらず、提案手法が有効であることがわかる。しかし、"Foreman"においてビットレートの増加分が大きくなっている。これは、5×5 ブロックでの DCT に、図 2 の E~H までの画素が含まれておらず、モード 3 や 7 の選択が誤っていることが原因として挙げられる。例えば、E~H の画素値が a~p の画素値と大きく異なっているも、5×5 ブロック内のエッジがモード 3 や 7 の方向に存在する場合、モード 3 や 7 を選択してしまうためである。

#### 5. まとめ

本稿では、JPEG デコード時に取得した DCT 係数及び画素値から、それぞれ新たに周波数成分を算出し、これをもとにイントラ予測モードを決定する手法を提案した。提案手法において、従来方式とほぼ同等の符号化効率を保ちつつ、処理量の削減が可能であることを確認した。今後は更なる符号化効率改善及び高速化について検討していく。また、今回は JPEG/H.264 イントラへの変換を行ったが、同じく離散コサイン変換が用いられている MPEG-2 からの変換に関してもその効果を確認していきたい。

#### 参考文献

- [1] ITU-T Rec. H.264, "Advanced video coding for generic audiovisual services," Nov. 2007
- [2] H. Kato, Y. Takishima, Y. Kaji, "Fast Intra Mode Decision Method For MPEG to H.264 Transcoding," ICIP 06, pp.833-836
- [3] X. Tang, Q. Du, S. Shang, "Fast Intra Prediction Mode Decision for MPEG-2 to H.264 Transcoding," ICME 07, pp. 428-431
- [4] Joint Video Team(JVT), "Reference Software JM14.0," <http://iphome.hhi.de/suehring/tml/>