

H.264 の Intra 予測モードにおける計算量低減 Complexity reduction of Intra prediction mode decision in H.264

野口 達也[†]
Tatsuya Noguchi

金田 海渡[†]
Kaito Kaneda

平澤 茂一[†]
Shigeichi Hirasawa

1. はじめに

近年、デジタル放送の普及やインターネットのブロードバンド化により、高画質な動画を扱う機会が増えてきた。これらの動画のほとんどは圧縮されており、MPEG-2 や MPEG-4 等、現在まで様々な圧縮規格が開発、研究されている。その中でも最新の規格である MPEG-4 part10 Advanced Video Coding/H.264 [1](以下 H.264) に注目が集まっている。H.264 は高い圧縮効率が特徴であり、同画質ならば MPEG-2 の約半分の符号量に抑えることができる。しかし、高い符号化効率の反面、符号化過程において膨大な計算量を必要としており、特に、エントロピー符号化前における画面内、画面間予測がその大半をしめている。[‡]

そこで本稿では、画面内予測の計算量削減のために、画面内予測 (以下 Intra 予測) の決定に際し、隣接マクロブロック (以下 MB) の予測モードにモード間の単純マルコフ性を利用した手法を提案し、実験によりその有効性を示す。

2. 画面内符号化方式

2.1 H.264 の画面内符号化方式

H.264 の画面内符号化では 16×16 の MB を 4×4 の MB に分割し、それぞれの MB について全ての予測モードを試す。 4×4 MB の予測モードの番号とその方向を図 1 に示す。全ての予測モードに対して Rate-Distortion 最適化 (以下 RD 最適化) を行い、最適な予測モードを決定する。さらにここで得られた 4×4 MB の予測モードによるコストと 16×16 の予測モードから得られたコストを比較し、より最適なモードを選択している。

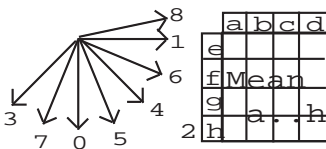


図 1: H.264 の 4×4 予測モード番号とその方向

しかし、前述したように全てのモードに対して符号化を行うため、画面内予測だけでも膨大な計算量がかかってしまうという問題がある。

2.2 従来研究とその問題点

画面内予測の計算量削減は広く研究が行われおり、その主な手法は、予測元となる 4×4 MB の画像特徴を検出し、それを指数として予測モードを RD 最適化に適用する前に絞り込む手法をとっている。文献 [2] では 4×4 MB 内の画像の濃度共起から変位を検出し、その変位から予測モードを決定する手法を提案し、約 30% の計算量の削減に成功している。文献 [3] では 4×4 MB 内から画像の濃度勾配による簡単なエッジを検出することで、予測モード候補の絞り込みを行い、約 60% の計算量の削減に成功している。

また、RD 最適化のアルゴリズムを簡略化することで計算量の削減を行っている研究もある。文献 [4] では RD

最適化に適用する予測モードを全探索ではなく、2 分探索を用いるアルゴリズムを提案し、約 30% の計算量の削減を行っている。

しかし、これらの手法は 4×4 MB 内のみの画像特徴を検出したり、単純なアルゴリズムの書き換えから計算量の削減を行っているだけであり、本来画像の持つ隣接 MB 同士の高い相関を利用していない。

そこで、MB 内ではなく、MB 間の予測モードの関係を考慮した予測モード決定方法を新たに提案する。そのため、MB 間の予測モードにどのような関係があるか調べる予備実験を行った。

2.3 予備実験

予備実験を行うにあたり、次のような仮定をした。

1. スキャン順の隣接する 4×4 MB の予測モードにどれくらいの相関があるか。
2. シーンカットがないとすれば、時間的に遠く離れた画面間で 4×4 MB の予測モードの遷移はどれくらい類似するか。

これらを検証するため、予備実験として 4×4 MB 番号の当該番号と 1 つ前の番号の MB 間の予測モードの遷移を取った。なお、予備実験には、H.264 のリファレンスソフトウェアである JM10.1[5] を利用し、予備実験用の動画に映像情報メディア学会の SIF 標準動画 [6] を用いた。また、符号化条件として、フレーム構成は全て I フレーム、 8×8 予測の使用なし、RD 最適化使用とする。“Cognac and Fruit”を用いた場合の結果を図 2、図 3 に示す。(上記の仮定 2 を検証するために 0 枚目と 108 枚目の画像を示した) 横軸が現在の MB の予測モードであり、縦軸がその次の MB の予測モードである。図 2 の一番左端の棒グラフを例にとれば、現在の MB の予測モードが 0 のとき、次の MB の予測モードが 0 である確率は約 0.8 である事を示している。

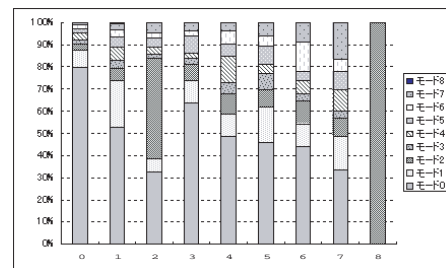


図 2: 0 枚目での隣接ブロックの予測モードの遷移

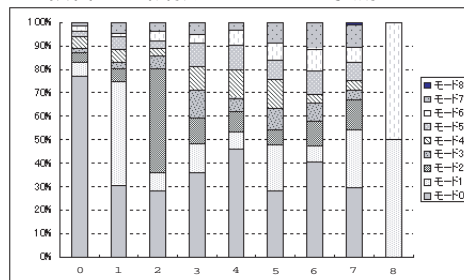


図 3: 108 枚目での隣接ブロックの予測モードの遷移

[†]早稲田大学理工学部経営システム工学科

[‡]文献 [1] によると、H.264 の符号化過程において画面内予測が 20%、画面間予測が 57% を占めているとの報告がある。

予備実験より、若干の差はあるものの予測モードの遷移確率には強い偏りがあり、100 フレーム程であれば遷移確率に大きな乖離がないことがわかった。また、遷移確率 0.1 以下のものは数が多い割りに、ひとつひとつの確率はあまり高くないことがわかる。

2.3の結果から、次の要件を満たす手法を3に提案する

1. MB間の予測モードに単純マルコフモデルを適用し、予測モードの遷移確率から確率が0.1以上のモードのみを候補としてRD最適化を行い、予測モードを決定する。
2. 単純マルコフによる不正解予測モードによるPSNRの低下を防ぐため、連続する4つのMBのうち1つは通常通り9つのモードでのRD最適化を行う。
3. 100フレームを越えたら、通常通り9つの予測モードで予測をし、遷移確率を作り直す。

3. 提案手法のアルゴリズム

提案手法のアルゴリズムを示す。また、本アルゴリズムを用いた場合の予測の流れを図4に示す。

提案手法のアルゴリズム

1. 0枚目のIフレームは通常通り9つの予測モードを用い、RD最適化を行う。
2. このときの当該4×4MBとその1つ前(ここでの順番はH.264のスキャン順とする)の4×4MBの予測モードの遷移を集計し、連続するMB間の予測モードの遷移確率表(辞書)を作成する。
3. B,Pフレームは通常通りの符号化を行う。(仮に画面内符号化が選択されても本手法は適用せず、通常通り9つのモードで予測するものとする)
4. 次のIフレームでは8×8MB内の左上の4×4MBに関しては9つのモードでRD最適化を行い、残りの3つの4×4MBに関しては2で作成した遷移確率表からモードの候補を選択する。このとき、選択する候補は遷移確率が閾値0.1以上のものとする。
5. 上記3.4.を繰り返し、全体の動画像が100フレームを超えた場合、その次のIフレームは遷移確率表を作り直す。1.に戻り、通常通りのRD最適化を行う。

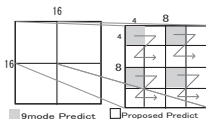


図4: 提案手法を用いた場合の予測順序と流れ

4. 実験とその結果

提案手法を検証するため、本手法をH.264のリファレンスソフトウェアであるJM10.1[5]に実装した。実験用の動画像は映像情報メディア学会のSIF標準動画像[6]を用いた。今回の手法では、シーンカットを考慮していないので、シーンカットなしの動画像を用いた。詳細な実験条件を表1に、実験結果を図5、表2に示す。また、実験にはCPU:Pentium 4HT 3.3GHz、メモリ:2GB、OS:WindowsXPのPCを用いた。

表1: 符号化条件

画像サイズ	SIF(352×240ピクセル)
実験用動画像	Cognac and Fruit, Flower Basket Yachting, Whale Show
フレーム構成	IBBPBBIBB...
符号化枚数	87枚(Iフレームは15枚)
エントロピー符号化	CABAC
R-D最適化	使用あり

表2: Iフレーム符号化における計算時間/フレーム

動画像	提案 (ms)	従来 (ms)	計算量削減比
Cognac and Fruit	846	2284	63% 削減
Flower Basket	956	2637	64% 削減
Yachting	809	2399	66% 削減
Whale Show	1055	3302	68% 削減

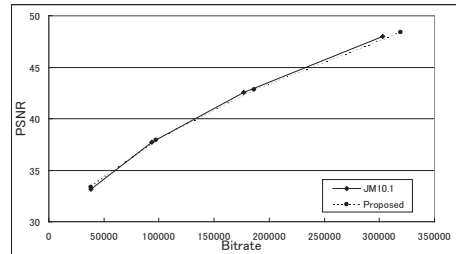


図5: Flower BasketにおけるIフレームのみのPSNR

5. 考察

1. 従来手法とほぼ同等の画質で、Iフレーム画面内の計算時間を平均して65%削減することに成功した。
2. 主観評価による評価も行った。結果は視覚的な違いは見られなかった。
3. 辞書作成の計算時間は、50ms-500msとばらつきはあるものの、符号化全体の計算時間から考えればさほど問題はないと思われる。
4. 本実験のフレーム構成であれば、本手法により全符号化時間のうち約10%を削減できた。

6. まとめと今後の課題

H.264の画面内符号化の予測符号化において、遷移確率表を用いることで計算量の低減を行い、実験によりその有効性を示すことができた。

今後の課題は、画面間符号化の計算量削減を検討し、H.264の符号化過程をさらに高速化することと、シーンカットにどう対処するかが考えられる。

参考文献

- [1] Yu-Wen-Huang, Bing-Yu Heish, Tung-Chien Chen, and Liang-Gee Chen, "Analysis, Fast Algorithm, and VLSI Architecture Design for H.264/AVC Intra Frame Coder", IEEE Transaction on Circuits and System, 2005.
- [2] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, "A DIRECTIONAL FIELD BASED FAST INTRA MODE DECISION ALGORITHM FOR H.264 VIDEO CODING", IEEE International Conference, 2004.
- [3] Zhang Yong-dong, Dai Feng, Lin Shou-xun, "Fast 4 × 4 Intra-prediction Mode Selection for H.264", IEEE International Conference, 2004.
- [4] Chao-Chung Cheng, Tian-Sheuan Chang, "Fast Three Step Intra Prediction Algorithm for 4 × 4 blocks in H.264", IEEE 2005.
- [5] JM10.1
<http://iphome.hhi.de/suehring/tml/>
- [6] SIF 画像・システム評価用標準動画像(社)映像情報メディア学会, 2003.