

CABAC 圧縮率推定を用いた H.264 におけるレート - 歪み最適化モード判定の高速化

A Fast Rate-Distortion Optimized Mode Decision Method for H.264 using CABAC Compression Ratio Estimation

浅野 渉† 古藤 晋一郎†
Wataru Asano Shin-ichiro Koto

1. まえがき

動画画像符号化の国際標準である H.264/MPEG-4 AVC[1] (以下 H.264)は、従来の MPEG-2 や MPEG-4 と比較して、約 2 倍の符号化効率があり、注目を集めている。H.264 には、複数参照フレームや様々な予測ブロック形状があり、画像の特徴に応じて最適な符号化モードを選択することで符号化効率の向上を実現している。また、エントロピー符号化に算術符号化を採用することで、更なる符号化効率の向上を実現している。最適な符号化モードを選択する手法として、レート - 歪み最適化(Rate-Distortion Optimization: RDO)手法がある[2]。筆者らは、符号化モード選択を 2 段階で行う高速 RDO 手法を提案した[3]。本稿では、算術符号化における圧縮率推定を用いることで、符号化効率を低下させずに、さらに RDO を高速化する手法に関して検討を行った。

2. レート - 歪み最適化手法

RDO 手法は、マクロブロック毎に各符号化モードに対して仮符号化処理を行い、得られる符号量と符号化歪みから、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り、H.264 標準化リファレンスソフト(JM)にも組み込まれている[4]。これは、目標符号量 R を拘束条件として、符号化歪み D が最小となる符号化モードを選択するということであり、次式のラグランジュコスト J を最小化することに等しい。

$$J = D + \lambda R \quad (1)$$

ここで、 λ はラグランジュ定数であり、JM では量子化値 QP に依存し、次式が用いられている。

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

このように、ある QP が与えられると λ が決定され、一定下で J が最小であるモードを選択する。 J を求めるためには R と D が必要であり、仮符号化を行わなければならない。H.264 には、ブロック形状や予測モードが多数あるため、RDO を行うと仮符号化回数が増加し、演算量が膨大になってしまう。

3. 算術符号化

H.264 に採用されている算術符号化は CABAC(Context-based Adaptive Binary Arithmetic Coding)と呼ばれており、符号化要素(syntax element)を 2 値化する Bin 生成部、発生確率を表すコンテキストを選択するコンテキスト選択部、ビットストリームを生成する算術符号化器から成る(図 1)。syntax element が与えられると、まず Bin 生成部で多値から 2 値化される。それと同時にコンテキスト選択部は syntax

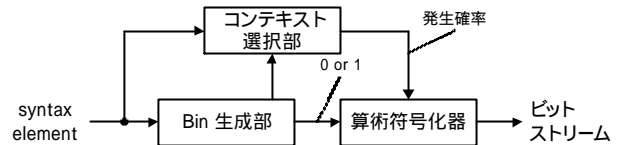


図1 CABACの構成

element の種類、周辺の状態、過去の符号化結果などに応じて 1Bin 入力毎に所定のコンテキストを選択する。算術符号化器は、選択されたコンテキストに従って、Bin 系列の符号化を行う。

このように、CABAC は、ビットレートを上回るレートでコンテキスト選択と算術符号化器での符号化処理が必要となる。また、1Bin 毎にコンテキストを切り替える必要があるため、並列化が非常に困難であり、演算負荷が大きい。このため、RDO 時の仮符号化に CABAC を使用すると演算量が大幅に増加してしまう。

4. 提案手法

以上で述べたように、RDO 方式でのモード判定は、膨大な演算量を必要とする。そこで、以下で RDO 方式の高速化手法を提案する。

4.1 Bin による RDO

RDO 時に CABAC を使用する場合、符号化歪み D の計算と発生符号量 R の計算では、 R の計算の負荷が大きい。そこで、 R を仮符号化によって求めるのではなく、Bin 系列の長さから R を推定することで RDO を行う。こうすることで、仮符号化でのコンテキスト選択、算術符号化処理が不要となり、演算量の削減が可能となる。Bin 数を N_{bin} 、発生符号量推定値を $R(N_{bin})$ とすると、ラグランジュコスト J

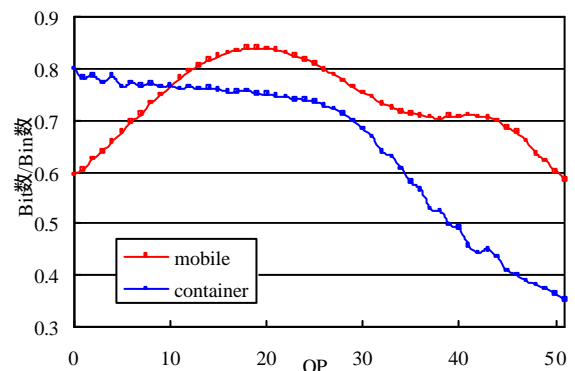


図2 Binの圧縮率の一例

表 1 実験条件

Profile: Main
QP: 固定 (16, 24, 32, 40)
予測構造 BBIBBPBBPBBPBBP...
参照フレーム枚数 2枚

は,

$$J = D + IR(N_{Bin}) \quad (3)$$

となる.

4.2 発生符号量の適応推定

算術符号化の圧縮率を α とすると, 発生符号量の推定値は, 次式のように記述できる.

$$R(N_{Bin}) = \frac{1}{\alpha} \cdot N_{Bin} \quad (4)$$

圧縮率 α は, 図 2 に示す様に, 量子化値 QP や画像の性質などによって異なるものとなる. そこで, 直前ピクチャでの Bin の圧縮率を用いて, Bin 数から発生符号量の推定を行う.

5. 実験及び考察

本提案の有効性を確認するために, SDTV サイズの ITE 標準動画 4 種類に対して提案方式と RDO 方式, 非 RDO 方式で符号化を行い, SNR とエンコード速度の比較を行った. 符号化条件は表 1 の通りである. 表 2 に提案手法と RDO 方式の符号量増加率 (Bitrate), 輝度信号の PSNR 増加量 (PSNR), 平均速度比を示す. 平均速度比は, 総エンコード時間から動き補償 (ME) の処理時間を除いた時間の逆数の比である. この表より, 提案手法は従来の RDO 手法より約 2.5~4 倍高速を達成している. さらに, 本手法では PSNR が約 0.1~0.3dB 向上していることが分かる. 図 3 に画像 16SD_Whale_Show に対する R-D 曲線を示しているがこの図を見ると, 低ビットレートで PSNR が向上していることが分かる. これは, 本提案手法が近似的に R を推定していることを鑑みると, 式(2)に示した RDO 方式の α が必ずしも最適ではないことを示唆している. 図 4 には, 非 RDO 方式の $QP=40$ 時の符号化時間を 1 とした時の相対符号化時間を示している. この図によると RDO 方式ではビットレートが高くなると処理時間が大幅に増加してしまうが, 提案方式では処理時間がそれほど増加していない. このことより, 提案方式は特に高ビットレートで高速化の効果が大きいことが分かる.

以上より, 提案方式は RDO 方式より高速で, 符号化効率も若干向上することが確認できた.

6. まとめ

本稿では, RDO におけるレート算出を算術符号化前の Bin から推計することで仮符号化時の CABAC 符号化を省略し, 符号化効率を損なわずに RDO を高速化する手法を提案した. また, Bin からビットストリームへの圧縮率を考慮し発生符号量推定パラメータ α を適応的に更新することで, 量子化値 QP や画像の性質の変化に対してロバストな手法を提案した. 本手法は, 先に提案した高速 RDO 手法 [3] と組み合わせることで, さらに高速な RDO 符号化を実現することが可能である.

表 2 実験結果 RDO 方式 VS 提案方式

Seq.	Bitrate	PSNR [dB]	速度比
01SD_Cognac_and_Fruit	-4.58%	0.313	2.53
16SD_Whale_Show	-1.52%	0.123	3.49
23SD_Green_Leaves	-2.22%	0.205	4.05
31SD_Flamingoes	-2.50%	0.183	2.83

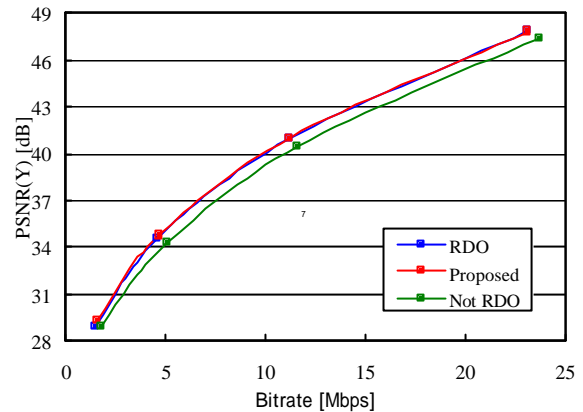


図 3 R-D 曲線 (16SD_Whale_Show)

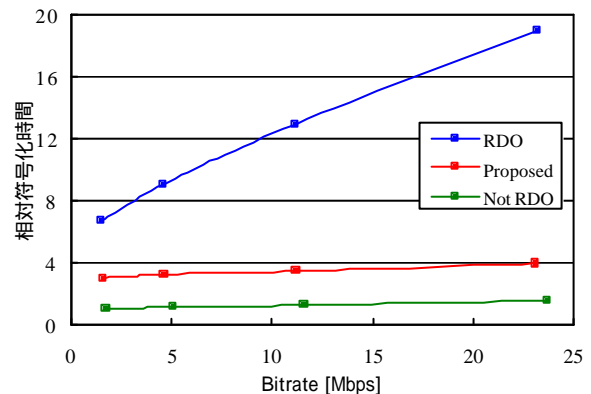


図 4 相対符号化時間 (16SD_Whale_Show)

参考文献

- [1] ISO/IEC 14496-10, "Advanced Video Coding" & ITU-T Rec. H.264, 2003.
- [2] Thomas Wiegand, and Bernd Girod, "Multi-frame motion compensated prediction for video transmission", Kluwer Academic Publishers, 2001.
- [3] 谷沢ほか, "H.264 におけるレート - 歪み最適化モード判定の高速化に関する一検討", 信学技報, IE2003-238, 2004.
- [4] JVT Reference Software, <http://bs.hhi.de/~suehring/tml/>