

新しい動き予測モード選択方式を用いた動画像符号化効率改善 Improvement of Video Coding Efficiency Applying New Prediction Mode Selection

軽部 勲† 村上 智一† 鈴木 芳典†
Isao Karube Tomokazu Murakami Yoshinori Suzuki

1. まえがき

近年の動画像符号化では、効率的な符号化を実現するため、動き補償予測を行うブロックのサイズや分割方法によって複数の動き予測モードが用意されている。例えばフレーム間の予測について MPEG-4 ASP(Advanced Simple Profile)[1]では 8 種類、H.264/AVC[2]では 44 種類の動き予測モードがある。

一般的な動き予測モード選択方法では、動きベクトルと予測誤差が符号化データとして出力され、画質が一定となる場合は符号量が最も小さいモードが選択される。

動きベクトルは予測動きベクトルとの差分をそのまま符号化するため、各モードに対して符号量を実測しても演算量はそれほど大きくなりません。一方、予測誤差は符号化の際に各画素の予測値との差分に対して DCT(Discrete Cosine Transform)などの周波数変換や量子化する処理が必要となり、複数のモードについて符号量を測定すると演算量が増大してしまうため、各モードの符号量をリアルタイムで実測することは非常に困難となっている。そこで、符号量推定において、動きベクトル符号量については実測値を利用し、予測誤差符号量に関しては適切な推定関数を用いて評価することによって動き予測モードを選択する研究が行われてきた。この予測誤差符号量の推定関数については、少ない演算量で処理を行う方式が考案されている[3]。

[3]の方式では符号量推定のために 1 つの推定関数を定義しているが、予測誤差と符号量の関係は動きの大きさなどの画像の特徴によってばらつきがあるため、用意した推定関数が正確な符号量を推定できず、適切な動き予測モードを選択できない場合があった。

本稿では、推定関数を複数化することによって画像特徴の変化に対応した、より正確な符号量推定を可能にする動き予測モード選択方式を提案する。

2. 提案動き予測モード選択方式

本方式では、予測誤差の符号量が動きの大きい画像では大きく、動きの小さい画像では小さくなるという傾向に着目して動き予測モード選択に利用する推定関数を複数化した。以下 2.1 では、関数の複数化の方法について、また 2.2 では、複数化に伴って必要となる推定関数の選択方法について説明する。

2.1 予測誤差符号量推定関数の複数化

提案方式では、動き予測モード選択時に利用する予測誤差符号量の推定関数を以下のように設定した。

$$P_k = A_k(QP/D) + B_k \quad (1)$$

ここで、QP は量子化パラメータ、D はフレームの SAD(Sum of Absolute Difference, 予測誤差の絶対値)、また、k は推定関数のインデックスを表している。また A_k および B_k は k によって変化するパラメータとする。[3]の方式では、予測誤差符号量推定関数として QP/D の 3 次関数を導入していたが、ここでは簡単のため、1 次関数とした。

動画像符号化では、動きがない場合には符号化を行わないブロックが増加することなどの理由から、予測誤差の符号量が動きの大きい画像では大きくなり、動きの小さい画像では小さくなる傾向がある。本方式ではこのような特徴に着目し、図 1 のように、3 通りの推定関数を用意した。 P_1 、 P_2 、 P_3 はそれぞれ、動きの大きい画像、平均的な動きの大きさの画像、動きの小さい画像に対する予測誤差符号量推定関数を表しており、各関数において

$$A_1 > A_2 > A_3 \quad (2)$$

となっている。 A_1 、 A_2 、 A_3 、および B_1 、 B_2 、 B_3 はそれぞれの特徴を持つ複数の画像を実際に符号化して求めることとする。このように動きの大きさに応じて 3 種類の推定関数を用いることで、従来の問題であった実測値と推定値の差を減らし、推定精度を向上させることができる。

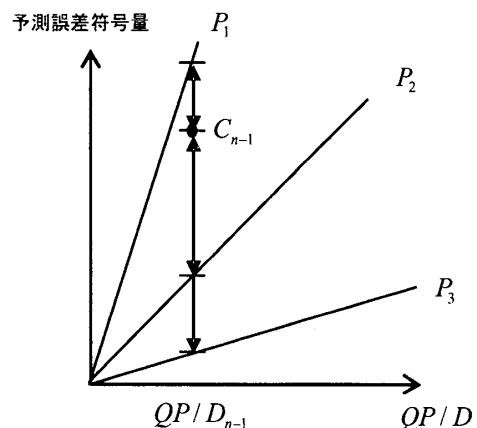


図 1. 複数の予測誤差符号量推定関数

† (株) 日立製作所, 中央研究所

2.2 誤差符号量推定関数選択方式

2.1 で述べたように推定関数を複数化した場合、適切な推定関数を選択する必要がある。提案方式では、

- ①符号化済みの画像では予測誤差の符号量は実測可能
- ②時間的に連続したフレーム間における性質の類似といった特徴を利用して、符号化が終了したフレームの予測誤差とその符号量の実測値を基に予測誤差符号量推定関数の評価を行い符号化対象のフレームに対する推定関数の選択することとした。

以下、図2を用いて、フレーム単位で関数選択を行う場合の例を説明する。

図2において、n-1を符号化済みのフレーム、nを符号化対象フレームとする。まず、フレームn-1全体における予測誤差とその符号量を測定し、それぞれ D_{n-1} 、 C_{n-1} とおく。次に推定関数 $P_1 \sim P_3$ に D_{n-1} を適用して推定値を求める。関数は図1に示すような3つの関数を用いることとし、推定値 $P_1(D_{n-1})$ 、 $P_2(D_{n-1})$ 、 $P_3(D_{n-1})$ をそれぞれ求める。そして、推定関数選択部において D_{n-1} と $P_1(D_{n-1})$ 、 $P_2(D_{n-1})$ 、 $P_3(D_{n-1})$ とを比較して推定関数を決め、フレームnに適用する。

D_{n-1} 、 C_{n-1} 、 $P_1(D_{n-1})$ 、 $P_2(D_{n-1})$ 、 $P_3(D_{n-1})$ の関係が図1のようになっていると仮定すると、

$$\begin{aligned} & \text{Min}(|P_1(D_{n-1}) - C_{n-1}|, |P_2(D_{n-1}) - C_{n-1}|, \\ & |P_3(D_{n-1}) - C_{n-1}|) \\ & = |P_1(D_{n-1}) - C_{n-1}| \end{aligned} \quad (3)$$

となっており、推定関数 P_1 がフレーム n-1 に対して最も推定精度が良い関数であることがわかる。ここで、上記②の性質から、 P_1 が符号化対象フレーム n に対しても推定精度の良い推定関数であると考えられるため、 P_1 をフレーム n に対して適用する。こうしてフレームの特徴に合わせた符号量の推定関数の選択が可能となる。

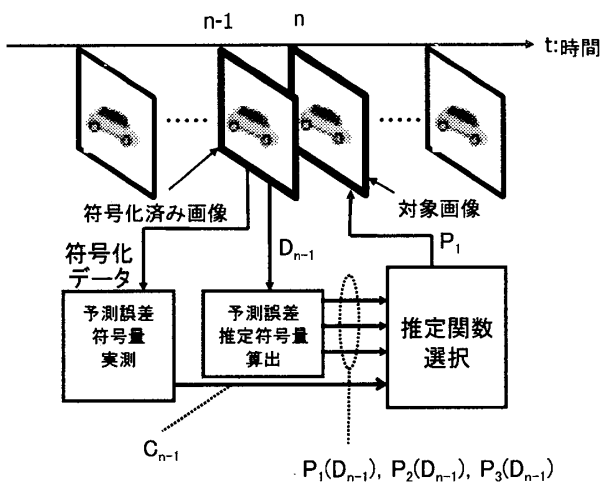


図2 予測誤差符号量推定関数選択の概要

3. 評価実験

上記提案方式のシミュレーション実験を行った。本実験には MPEG-4 ASP で規定された動き予測モードを対象として、提案方式を評価した。

一例として標準画像 Paris を用いた場合の、従来方式と提案方式における、ビットレート符号化性能比較結果を図3に示す。破線は推定関数が1つの場合(従来方式)、実線は推定関数が3つの場合(提案方式)の結果を示している。画像サイズは CIF、フレームレートは 30fps、フレーム数は 1064 枚、B フレーム間隔は 2 とした。

図3より、提案方式は従来方式と比較して同一の画質で約7%のビットレート削減を、同一ビットレートで約0.2dBのPSNRの向上が確認できる。

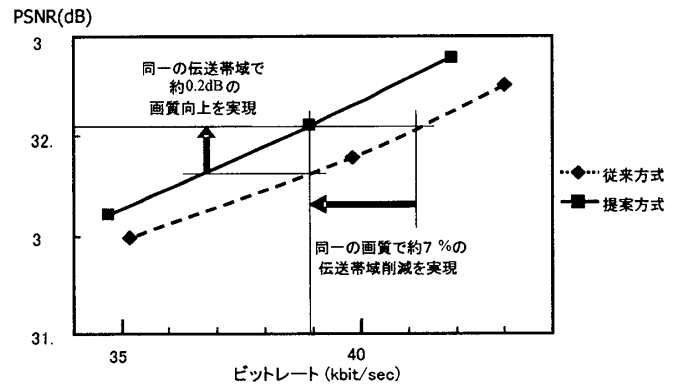


図3. ビットレートと画質の比較

4. まとめ

リアルタイム処理可能で、より高画質な符号化を実現する動き予測モード選択方式として、予測誤差符号量推定関数の複数化と、推定関数の選択アルゴリズムを提案した。また、本方式を従来方式と比較した結果、同一画質で約7%の符号量削減を実現した。本稿では実験に MPEG-4 ASP(Advanced Simple Profile)エンコーダを用いたが、MPEG-2 や H.264 などの他方式に対しても適用が可能である。今後は更なる精度向上に向け、より正確な推定関数の導出方法について検討する予定である。

参考文献

- [1] Jens-Rainer Ohm, et al. "Text of ISO/ICE 14496-2 Third Edition", March 2003.
- [2] T. Wiegand, et al. "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)", May, 2003.
- [3] Y. Suzuki and Y. Nakaya, "ON/OFF Decision of Global and Local Motion Compensation in Low-Bitrate Video Coding" in Processing of ICIP, Vol. 1, pp99-103, Oct. 1999.