

フレーム間相関を利用した動き補償方式の基礎検討

坂東幸浩, 高村誠之, 八島由幸

日本電信電話株式会社 NTT サイバースペース研究所

〒 239-0847 神奈川県横須賀市光の丘 1-1

E-mail : bandou.yukihiro@lab.ntt.co.jp

動き補償を用いた動画像符号化において、符号化対象フレームと予測フレームのフレーム間相関に基づく修正係数を予測フレームの領域毎に乗じることにより、予測誤差電力を低減できることが知られている。この際、問題となるのは、修正係数を乗じる領域の大きさである。小さな領域を単位として修正係数を乗じれば、それだけ予測誤差電力を低減できるため、変換係数の符号量低減につながる。一方、これは、小さな領域毎に付加情報としての修正係数を持つことになるため、修正係数の総数が増加し、修正係数を表現するのに必要な情報量は増大する。つまり、領域サイズと予測誤差の符号量にはトレードオフの関係がある。しかし、従来の技術では、こうしたトレードオフに関して最適な領域サイズの設定方法を与えていない。本稿では、予測誤差信号の符号量と符号化歪みの推定を通して、修正係数を乗じる最適な領域サイズの決定法を提案する。計算機シミュレーションにより、領域サイズを固定した場合と比較して提案法ではSNRが約2 [dB] 向上することを確認した。

A study on motion compensation using inter frame correlation

Yukihiro BANDO, Seishi TAKAMURA, and Yoshiyuki Yashima

NTT Cyber Space Laboratories, NTT Corporation

1-1 Hikarino-oka, Yokosuka, Kanagawa 239-0847, JAPAN

E-mail : bandou.yukihiro@lab.ntt.co.jp

Motion compensation (MC) is one of the core techniques in video coding. To reduce the power of MC error, it is effective to divide a frame into segments, and then to multiply each segment by a coefficient. The coefficient is computed based on correlation between current frame and MC frame. This coefficient is called modify-coefficient. In this modification, the segment size has to be chosen suitably. The smaller the segment size becomes, the less prediction error becomes, so coding rate can be reduced. On the other hand, because of the increase of the number of modify-coefficients, the amount of information for modify-coefficients increases. However, conventional techniques did not give any solutions to achieve a trade-off between the reduction of prediction error and increase of the amount of information for modify-coefficients. Therefore, to obtain the optimal segment size, it is necessary to perform encoding operation for all possible segment sizes. In this paper, we propose a method to determine the optimal segment size through the estimation of prediction error and coding rate. Results of computer simulation show that our method can achieve about 2 dB improvement in SNR over a method using fixed segment size

1 はじめに

動画像符号化における重要な技術の一つに動き補償 (MC) ブロック直交変換がある [1, 2]。この方法では、まず、符号化対象フレーム $f_t(\mathbf{p})$ ($\mathbf{p} \in \{(x, y) | 0 \leq x \leq X-1, 0 \leq y \leq Y-1\}$) を複数の領域に分割し、領域毎に動き補償 (MC) を用いた次のようなフレーム間予測を行う [3]。

$$\tilde{f}_t(\mathbf{p}) = f_{t-1}(\mathbf{p} - \mathbf{v}_i), \quad \mathbf{p} \in B_i$$

ここで、 B_i は第 i 番目の領域であり、 \mathbf{v}_i は次式を満たす動きベクトルである。

$$\mathbf{v}_i = \arg \min_{\mathbf{v} \in R_i} \sum_{\mathbf{p} \in B_i} \{f_t(\mathbf{p}) - f_{t-1}(\mathbf{p} - \mathbf{v})\}^2 \quad (1)$$

すなわち、探索範囲 R_i において、領域 B_i 内での予測誤差を最小化する \mathbf{v} が動きベクトル \mathbf{v}_i として選ばれる。次に、予測誤差に直交変換が施され、量子化後に、得られた変換係数が可変長符号化される。

予測誤差の低減が変換係数の符号量低減につながることから、予測誤差の低減を目的として、次のようなフレーム間予測の改良が検討されている [4]。

$$\hat{f}_t(\mathbf{p}) = \gamma_i f_{t-1}(\mathbf{p} - \mathbf{v}_i), \quad \mathbf{p} \in B_i \quad (2)$$

ここで、 γ_i は

$$\gamma_i = \frac{\sum_{\mathbf{p} \in B_i} f_t(\mathbf{p}) f_{t-1}(\mathbf{p} - \mathbf{v}_i)}{\sum_{\mathbf{p} \in B_i} f_{t-1}^2(\mathbf{p} - \mathbf{v}_i)}$$

であり、式 (1) の \mathbf{v}_i に対して

$$\sum_{\mathbf{p} \in B_i} \{f_t(\mathbf{p}) - \gamma_i f_{t-1}(\mathbf{p} - \mathbf{v}_i)\}^2$$

を最小化する値として求めたものである。以下、 γ_i を修正係数と呼び、式 (2) に従う動き補償予測を修正動き補償 (Modified MC, MMC) と呼ぶ。

上述のように修正係数を予測フレームに乗じる場合、問題となるのは、同一の修正係数を乗じる領域のサイズ (同領域内に含まれる画素数) である。領域サイズを小さくとれば、それだけ予測誤差電力を低減できるため、変換係数の符号量低減につながる。

一方、これは、小さな領域毎に付加情報としての修正係数を持つことになるため、修正係数の総数が増加し、修正係数を表現するのに必要な情報量は増大する。つまり、領域サイズと予測誤差の符号量にはトレードオフの関係がある。しかし、従来の技術では、こうしたトレードオフに関して最適な領域サイズの設定方法を与えていない。このため、最適な領域サイズを求めようとした場合、候補となる全ての領域サイズに対して符号化を実行する必要があり、処理時間の面で実用に供さなかった。

そこで、本稿では、修正係数を乗じる最適な領域サイズの高速な決定法を提案する。別途求めた 2 つの領域サイズに対する符号化結果から、その中間の領域サイズにおける動き補償信号の歪み、及び符号量を推定する。具体的には、直交変換の電力保存性に基づいて変換係数毎に内挿を行う。こうして得られる符号化歪みの推定値に基づき、修正係数を乗じる最適な領域サイズを与える。

2 MMC における領域サイズの最適化

2.1 領域サイズと MMC 誤差電力の関係

複数の領域 B_i を結合した領域における MMC では、領域 B_i 毎の MMC と比較して、予測誤差が増加する。以下では、その予測誤差の増加量を解析的に示す。 B_i ($i = s+1, \dots, s+\ell$) を結合した領域を

$$C(s, \ell) = B_{s+1} \cup \dots \cup B_{s+\ell}$$

と表わす。このとき、領域 $C(s, \ell)$ において、

$$\sum_{\mathbf{p} \in C(s, \ell)} \{f_t(\mathbf{p}) - \Gamma f_{t-1}(\mathbf{p} - \mathbf{v}_i)\}^2$$

を最小化する値 Γ は γ_i ($i = s+1, \dots, s+\ell$) を用いて次のように表わせる。

$$\Gamma = \frac{\sum_{i=s+1}^{s+\ell} \gamma_i \sigma_i^2}{\sum_{i=s+1}^{s+\ell} \sigma_i^2}$$

ここで、 σ_i^2 は

$$\sigma_i^2 = \sum_{\mathbf{p} \in \mathbf{B}_i} \mathbf{f}_{t-1}^2(\mathbf{p} - \mathbf{v}_i)$$

である。この Γ を用いて、領域 $\mathbf{C}(s, \ell)$ 内の MC 後の値を次のように修正する。

$$\hat{\mathbf{f}}_t(\mathbf{p}) = \Gamma \sum_{i=s+1}^{s+\ell} \mathbf{f}_{t-1}(\mathbf{p} - \mathbf{v}_i), \quad \mathbf{p} \in \mathbf{C}(s, \ell) \quad (3)$$

このとき、領域を結合したことによる予測誤差の増加量は

$$\delta(s, \ell) = \sum_{i=s+1}^{s+\ell} \gamma_i^2 \sigma_i^2 - \frac{\left(\sum_{i=s+1}^{s+\ell} \gamma_i \sigma_i^2 \right)^2}{\sum_{i=s+1}^{s+\ell} \sigma_i^2} \quad (4)$$

となる。

動き補償ブロック直交変換では、予測誤差信号に対して直交変換が適用される。以下では、領域サイズに対して変換係数の符号化歪みを推定する方法を示す。MMC を行う領域サイズとして、次の3種類を考える。

- (i) 領域 \mathbf{B}_i ($i = 1, \dots, I$),
- (ii) 領域 $\mathbf{C}(s_j, \ell)$ ($j = 1, \dots, I/\ell$),
- (iii) 領域 $\mathbf{C}(s_{j'}, \ell')$ ($j' = 1, \dots, I/\ell'$).

ここで、 $\ell < \ell'$ であり、領域サイズは (i), (ii), (iii) の順で大きくなっている。なお、 I は領域 \mathbf{B}_i の個数を表わす。(i), (iii) の領域を単位として MMC を行い、予測誤差信号に対して直交変換を適用する。こうして求めた2種類の MMC 誤差信号に対する直交変換の結果から、(i) と (iii) の中間サイズである (ii) について、符号化歪みを内挿により推定する。

2.2 変換係数の電力推定

まず、MMC 誤差電力の増加量をもとに、変換係数の各成分に対して電力の推定を行う。ここでは、変換係数の第 k 成分の電力を (i), (ii), (iii) の場合について、各々 $\zeta_1(k)$, $\zeta_2(k)$, $\zeta_3(k)$ ($k = 0, 1, \dots, K-1$) とする。なお、 K は直交変換の次元数を表わす。このうち、直交変換の実行結果として $\zeta_1(k)$ と $\zeta_3(k)$ は求まっている。さらに、(i) から (ii) へ領域を結合することに伴う MMC 誤差の増加量 Δ_1 および、(i) から (iii) へ領域を結合することに伴う MMC 誤差の増加量 Δ_2 は、

$$\Delta_1 = \sum_{j=1}^{I/\ell} \delta(s_j, \ell) \quad (5)$$

$$\Delta_2 = \sum_{j'=1}^{I/\ell'} \delta(s_{j'}, \ell') \quad (6)$$

と計算できる。ここで、 $\delta()$ に関しては式 (4) により計算する。直交変換では、変換の前後で信号の電力を不変に保つため、MMC 誤差の増加量が、そのまま変換係数の電力の増加量となる。そこで、(ii) の場合の $\zeta_2(k)$ の推定値を

$$\hat{\zeta}_2(k) = \frac{\Delta_1 \zeta_3(k) + (\Delta_2 - \Delta_1) \zeta_1(k)}{\Delta_2} \quad (7)$$

とする。これは、図 1 に示すように、(i) と (iii) の場合の係数電力を電力増加量で内分したものである。

次に、推定された係数の電力 $\hat{\zeta}_2(k)$ を用いて、領域サイズと変換係数の符号化歪みの関係を定式化する。MMC に要する総符号量 R は、上述の変換係数の符号量 r_e と修正係数の符号量 r_c との和となる。このうち、修正係数の符号量 r_c は領域サイズから一意に定まる。各修正係数を ρ [bits] で符号化するとすれば、(ii) の場合は I/ℓ 個の修正係数があるため、修正係数の符号量は

$$r_c = \rho I / \ell \quad [\text{bits}]$$

となる。このため、変換係数の符号化に割り当てられる符号量は

$$r_e = R - \rho I / \ell \quad [\text{bits}]$$

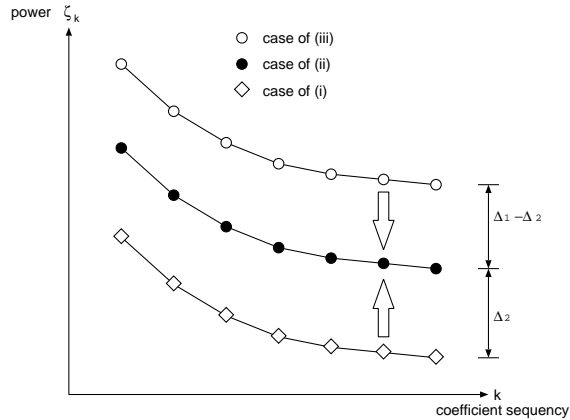


図 1: 変換係数の電力の内挿

となる。上式に従い、変換係数の符号化に用いる符号量を $R - \rho I / \ell$ ビットとした場合、達成しうる最小の符号化歪み（量子化誤差）は次のようになる [5]。

$$D = \beta^2 2^{-2(R - \rho I / \ell)} \left\{ \prod_{k=0}^{K-1} \hat{\zeta}_2(k) \right\}^{1/K} \quad (8)$$

ここで、 β は変換係数の分布に対応して定まる定数であり、ラプラス分布に対しては $\beta = 4.5$ となる。直交変換として離散コサイン変換 (DCT) を想定した場合、変換係数毎の分布はラプラス分布で近似できることが経験的に知られている。

2.3 最適領域サイズの推定

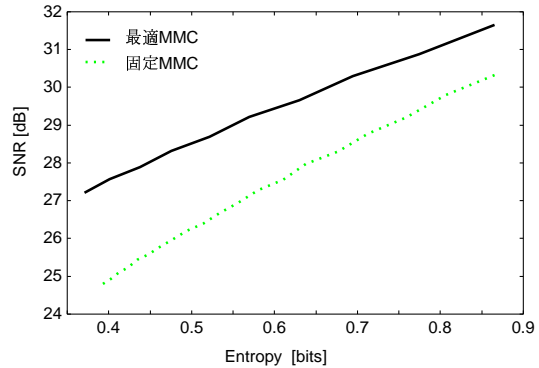
上述の結果を用いれば、符号化歪みを最小化する領域サイズを求める問題は、

$$\ell_{min} = \arg \min_{\ell} D$$

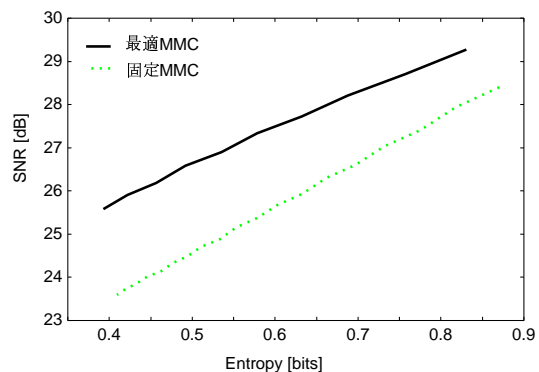
として、式 (7) を最小化する ℓ を求める問題に帰着される。

3 実験

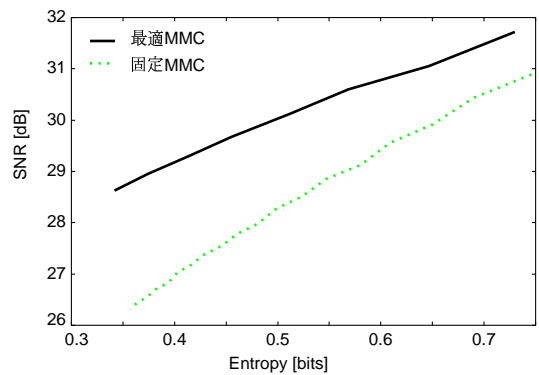
本手法の効果を検証するために計算機シミュレーションによる符号化実験を行った。実験において用



(a) Flower Garden



(b) Mobile & Calender



(c) Horse

図 2: エントロピーと SNR の関係

いた画像は CIF 画像 (352 × 288 [画素/フレーム]) “Flower Garden”, “Mobile & Calender”, “Horse”, である。MMC 誤差に対する直交変換は DCT を用いる。なお、符号量は DCT 係数の量子化ステップを用いて制御する。修正係数の符号量は 1 係数あたり 16 [bits] とする。また、イントラフレームの挿入は 14 フレーム毎に行い、フレーム間予測は順方向 MC を用いる。このとき、1 つの動きベクトルに対応するブロックのサイズは 16 × 16 [画素] とする。

図 2 に最適 MMC と固定 MMC に対するエントロピーと SNR の関係を示す。ここで、最適 MMC は修正係数を乗じる領域サイズとして 8 × 8, 16 × 16, 32 × 32, 88 × 72, 176 × 144, 352 × 288 [画素] の 6 種類を用意し、2.3 に基づきフレーム毎に最適な領域サイズを選択した MMC である。一方、固定 MMC は領域サイズを 8 × 8 [画素] に固定した MMC である。なお、縦軸の SNR は原画像を $f_t(x, y)$ 、復号画像を $g_t(x, y)$ ($0 \leq x \leq X - 1, 0 \leq y \leq Y - 1$) として、次式により求めた。

$$\text{SNR} = 10 \log_{10} \left(\frac{255^2}{\frac{1}{XY} \sum_{y,x} \{f_t(x, y) - g_t(x, y)\}^2} \right)$$

また、横軸は DCT 係数のエントロピーである。

同図において両手法を比較することにより、領域サイズの推定を導入した効果を確かめることができる。固定 MMC のように、領域サイズを 8 × 8 に固定した場合、修正係数に対する 1 画素あたりの符号量は 0.25 [bits] となり、全体の符号量が少ない場合では過剰な符号量が修正係数に割かれるため、符号化効率が低下する。同図において、エントロピーの減少に伴い本手法と従来法の差が大きくなっているのはこのためである。

4 おわりに

本稿では、MMC において修正係数を乗じる領域サイズの決定法を提案した。提案法は、予測誤差信号の符号量と符号化歪みの推定を通して、領域サイ

ズを定めるものである。このため、候補となる全ての領域サイズに対して符号化を実行することなしに、領域サイズを求めることが可能となった。これにより、フレーム毎に修正係数を乗じる領域サイズを適応的に切り替えることが可能となった。その結果、領域サイズを固定した MMC と比較して、SNR に関して約 2 [dB] の向上を達成した。

なお、今回の基礎検討では符号量をエントロピーで代用したが、今後は本手法をコーデックに実装し、その効果を検証する予定である。

参考文献

- [1] 安田浩, 渡辺裕. デジタル画像圧縮の基礎. 日経 BP 出版センタ, 1996.
- [2] 原島 博 (監修). 画像情報圧縮. オーム社, 1991.
- [3] J. R. Jain and A. K. Jain. Displacement measurement and its application in interframe image coding. *IEEE Trans. Computer*, Vol. COM-29, No. 1, pp. 1799–1806, 1981.
- [4] K. H. Tzou, T. R. Hsing, and N. A. Daly. Block-recursive matching algorithm (BRMA) for displacement estimation of video images. *Proc. of IEEE Int. Conf. Acoustics, Speech & Signal Processing*, pp. 359–362, 1985.
- [5] N. S. Jayant and P. Noll. *Digital coding of waveforms*. Prentice Hall, 1984.