フレーム毎にMCと3次元予測を最適化する 動画像の可逆符号化方式

A Lossless Video Coding Scheme Using MC and 3D Prediction Optimized for Each Frame

正会員 塩寺 太一郎^{†*}, 梅津 有司[†], 学生会員 前田 弘樹[†],
正会員 松田 一朗[†], 正会員 森岡 一幸[†], 正会員 伊東 晋[†]

Taichiro Shiodera^{†*}, Yuji Umezumi[†], Hiroki Maeda[†],
Ichiro Matsuda[†], Kazuyuki Morioka[†] and Susumu Itoh[†]

Abstract A novel lossless coding scheme for video signals is described. The scheme uses a block-adaptive 3D prediction method that predicts a video signal based on both the current and motion-compensated previous frames. The resulting prediction errors are encoded by using a kind of context-adaptive arithmetic coder. Several coding parameters such as prediction coefficients and motion vectors are iteratively optimized for each frame so that the overall coding rate can be a minimum. Moreover, a quadtree-based variable block-size motion compensation technique is introduced for efficient representation of motion information. Experimental results show that the coding rates of our scheme are 11–17% lower than those of the H.264-based lossless coding scheme.

キーワード：可逆符号化, 動画像符号化, 適応 3 次元予測, 算術符号, 可変ブロックサイズ動き補償

1. ま え が き

動画像の可逆符号化は、映像情報のデジタル・アーカイブや放送用素材の記録など、高品質な動画像を扱う分野で必要とされる技術であり、近年様々な符号化アルゴリズムが開発されている。これらのアルゴリズムは、フレーム毎に独立に符号化を行うフレーム内（イントラフレーム）符号化方式と、隣接フレームとの相関を利用するフレーム間（インタフレーム）符号化方式に大別できる。これまでは、処理が簡便でランダムアクセスの容易なフレーム内符号化方式が主流であったが、最近では計算機の処理能力向上に伴い、動き補償（MC）の技術を駆使した高性能なフレーム間符号化方式に関する研究が盛んになっている^{1)~4)}。中でも、動き補償された前フレームの画素と当該フレームの符号化済み画素を同時に用いる 3 次元予測方式⁴⁾は、時間方向予測と空間方向予測を統合した画期的な手法であり、高効率な可逆動画像符号化を実現する上で有望な要素技術と

いえる。文献 4) では、予測の重み（予測係数）を符号化済みの画素を用いて逐次更新する、画素適応型の 3 次元予測方式が提案されており、比較的優れた符号化特性を達成している。しかし、この方式では動ベクトルの検出や予測係数の導出に平均 2 乗誤差最小（MMSE）規範を用いており、その結果が符号化レートを最小化するという意味で最適であるという保証はない⁵⁾。また、符号化済みの画素を用いて予測係数を導出しているため、適応予測に関する付加情報は一切不要であるものの、予測係数を求める処理を符号化側と復号側の双方で実行する必要があるため、復号処理が一般に複雑となってしまう。このことは、蓄積された動画像を実時間で再生する必要があるアプリケーションにおいて、実用性に直結する深刻な欠点といえる。

これらの問題に対処するため、本論文では復号処理が簡単なブロック適応 3 次元予測を用いた動画像の可逆符号化方式を提案する。ベースとなるブロック適応 3 次元予測は、先に静止画像用に開発したブロック適応 2 次元予測手法⁶⁾と動き補償の技術を組合せたものであり、予測係数だけではなく、ブロック毎の予測器の選択情報や動ベクトルといった復号処理に必要なパラメータを、フレーム毎に最適化することで高い予測性能を実現している。この最適化処理は、従来行われていたような MMSE 規範に基づいたものではなく、適算術符号化による符号量を表すコスト関数を最

2005 年 11 月 14 日受付, 2006 年 1 月 11 日採録

[†] 東京理科大学 理工学部

(〒 278-8510 千葉県野田市山崎 2641, TEL 04-7124-1501)

* 現在, (株) 東芝研究開発センター

[†] Faculty of Science and Technology, Tokyo University of Science
(2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan)

* Toshiba Corporate Research & Development Center

3. 予測誤差の符号化

上記のブロック適応予測によって得られる予測誤差 $e = s_t(\mathbf{p}_0) - \hat{s}_t(\mathbf{p}_0)$ は、コンテキストモデリングと呼ばれる手法を用いて適応算術符号化される。コンテキストモデリングは、符号化済みの周辺画素群の状態（コンテキスト）から注目画素における予測誤差の確率密度関数 $P(e)$ を推定する手法であり、提案方式ではコンテキストを近傍画素の予測誤差から算出される特徴量 U に基づいて定義している。

$$U = \sum_{k=1}^6 |s_t(\mathbf{p}_k) - \hat{s}_t(\mathbf{p}_k)| + \sum_{k=0}^4 |s_{t-1}(\mathbf{q}_k) - \hat{s}_{t-1}(\mathbf{q}_k)| \quad (2)$$

上式の右辺第1項、第2項はそれぞれ当該フレーム、前フレームの符号化済み画素における予測誤差の絶対値和を表している。ただし、2次元予測を適用するイントラフレームでは第2項を0として扱う。特徴量 U は周辺画素における予測の当たりやすさを示す指標とみなせ、予測誤差 e の確率密度関数 $P(e)$ は U の値に大きく依存していると考えられる。そこで、特徴量 U を閾値 $\{Th_m(1), Th_m(2), \dots, Th_m(15)\}$ を用いて16レベルに量子化し、それぞれの量子化レベルに対応したコンテキスト ($n = 1, 2, \dots, 16$) 毎に、予測誤差の条件付確率密度関数を次式の一般化ガウス関数でモデル化する⁶⁾。

$$\tilde{P}(e|n) = \frac{c_n \cdot \eta(c_n, \sigma_n)}{2\Gamma(1/c_n)} \cdot \exp\left\{-|\eta(c_n, \sigma_n) \cdot e|^{c_n}\right\},$$

$$\eta(c_n, \sigma_n) = \frac{1}{\sigma_n} \sqrt{\frac{\Gamma(3/c_n)}{\Gamma(1/c_n)}} \quad (3)$$

ただし、 $\Gamma(\cdot)$ はガンマ関数、 σ_n は標準偏差、 c_n は分布形を制御するパラメータ（形状パラメータ）である。本論文では、実際の画像の性質に適した確率密度関数モデル $\tilde{P}(e|n)$ を得るため、コンテキスト毎の c_n の値を0.2~3.2の間（0.2刻み）で可変としている。また、標準偏差 σ_n についてはコンテキスト毎に決められた固定値を採用し、代わりに特徴量 U の量子化に用いる閾値 $\{Th_m(1), Th_m(2), \dots, Th_m(15)\}$ を可変としている。

ここで、8ビット精度の画像では輝度値が0~255の整数で表現されることに着目すると、予測値 $\hat{s}_t(\mathbf{p}_0)$ が既知の場合の予測誤差 e のとり得る値は以下に示す256通りとなる。

$$e \in \{s - \hat{s}_t(\mathbf{p}_0) \mid s = 0, 1, \dots, 255\} \quad (4)$$

さらに、予測処理における演算精度を保証するため予測値 $\hat{s}_t(\mathbf{p}_0)$ が一定の精度（ステップ幅 $h_s < 1$ ）で意図的に丸められているものとする、予測値 $\hat{s}_t(\mathbf{p}_0)$ とコンテキスト n が同時に与えられた際の離散的な予測誤差 e の条件付生起確率を、式(3)の確率密度関数モデル $\tilde{P}(e|n)$ を用いて次のように表すことができる。

図2 予測誤差 e の条件付生起確率
Conditional probability of occurrence
of a prediction error e .

$$\Pr(e|\hat{s}_t(\mathbf{p}_0), n) = \frac{\Pr(e|n)}{\sum_{s=0}^{255} \Pr(s - \hat{s}_t(\mathbf{p}_0)|n)}, \quad (5)$$

$$\Pr(e|n) = \int_{-h_s/2}^{h_s/2} \tilde{P}(e + \varepsilon|n) d\varepsilon \quad (6)$$

提案方式では、上式の条件付確率 $\Pr(e|\hat{s}_t(\mathbf{p}_0), n)$ に基づいて各画素の予測誤差 e の値を適応算術符号化している。ここで、式(5)の分子は確率密度関数モデル $\tilde{P}(e|n)$ を e を中心に区間幅 h_s で積分した値（図2の濃い灰色部分の面積）、分母は式(4)に列挙した値について同様に求めた積分値の和（同図の灰色部分の総面積）をそれぞれ表している。したがって、これらの値を予めサンプリング間隔 h_s で全て求めておき、符号化処理の際にテーブル参照することで、算術符号化に必要な確率の計算を高速化できる。図2からわかるように、パラメータ h_s の値を小さく設定することは、予測値の小数点以下の値 ($\hat{s}_t(\mathbf{p}_0) - \lfloor \hat{s}_t(\mathbf{p}_0) \rfloor$) を考慮して高精度な確率推定を行うことを意味するが、同時に上記のテーブル参照に必要なメモリー容量も増加してしまう。本論文では、 h_s の値と符号化効率および必要なメモリー容量について調査した結果に基づき、その設定値を $h_s = 1/8$ とした⁷⁾。また、算術符号器の実装に関しては、高速化の観点から整数演算のみで構成されるレンジコード⁸⁾を採用している。レンジコードの符号化効率はエントロピーに充分漸近していることが知られており、提案方式による予測誤差の符号化ビット数は、次式によって見積ることができる。

$$L(e|\hat{s}_t(\mathbf{p}_0), n) = -\log_2 \Pr(e|\hat{s}_t(\mathbf{p}_0), n) \quad (7)$$

4. 符号化パラメータの最適化

提案方式では予測誤差以外の付加情報として以下のパラメータが復号処理に必要となる。

- 動ベクトル $\mathbf{v}$
- 予測係数 $a_m(k)$
- 予測器（クラス）の選択情報 m
- コンテキストモデリングの閾値 $Th_m(n)$
- 確率密度関数モデルの形状パラメータ c_n

図 1 予測に用いる画素の配置
Disposition of pels used in prediction.

小化するように実行される。また、得られたパラメータはフレーム毎の付加情報としてビットストリームに多重化されるため、復号側では上記の最適化処理が一切不要であり、原理的に高速な再生が可能である。

以下では、動き補償に基づいたブロック適応 3 次元予測およびそれを用いた動画画像信号の可逆符号化アルゴリズムについて説明する。また、MC に必要な動ベクトルの情報を効率良く記述するため、四分木に基づいた可変ブロックサイズ MC の導入に関しても検討を加える。最後に、CIF サイズのモノクローム動画画像を対象とした符号化実験を通して提案方式の有効性を明らかにする。

2. 動き補償に基づいたブロック適応 3 次元予測

提案方式では、MPEG などの国際標準方式と同様に、連続する動画画像信号を数フレーム毎に区切った GOP (Group Of Pictures) 単位で符号化を行う。GOP 内の先頭フレーム (イントラフレーム) は、基本的に先に報告した静止画像用の可逆符号化方式⁶⁾と同様な手順で符号化される。すなわち、符号化対象フレームは予め 8×8 画素のブロック単位で M 通りのクラスに分類されており、ラスタースキャン順に走査される画素に対して、その画素が属するブロックのクラスに対応した線形予測器が適応的に割り当てられる。このブロック適応予測で使用する M 組の予測器は、いずれも図 1 (a) のように当該フレーム内の符号化済み画素 p_k を参照画素としている。このように、注目画素と同一フレームの画素群を用いる予測方式を 2 次元予測と呼ぶ。

一方、2 枚目以降のフレーム (インタフレーム) には、図 1 (b) のように当該フレームの符号化済み画素 p_k だけではなく、前フレーム上の画素 q_k も同時に用いる 3 次元予測を適用する。ここで、画素 p_k, q_k における輝度値をそれぞれ $s_t(p_k), s_{t-1}(q_k)$ と表記すると、注目画素 p_0 に対し m 番

目 ($m = 1, 2, \dots, M$) の予測器を適用した際の予測値は次式によって算出される。

$$\hat{s}_t(p_0) = \sum_{k=1}^{K_p} a_m(k) \cdot s_t(p_k) + \sum_{k=1}^{K_q} a_m(k + K_p) \cdot s_{t-1}(q_{k-1}) \quad (1)$$

ただし、 K_p, K_q はそれぞれ当該フレームおよび前フレームに配置した参照画素の数、 $a_m(k)$ ($k = 1, 2, \dots, K_p + K_q$) は各参照画素に関する予測の重み (予測係数) である。また、前フレームに配置した参照画素群 q_k ($k = 0, 1, \dots, K_q - 1$) は、正方ブロック単位で検出された動ベクトル $v = (v_x, v_y)$ にしたがって動き補償されており、その中心画素 q_0 と符号化対象画素 p_0 との間には $q_0 = p_0 + v$ の関係が成り立つ。これにより、動きを有する領域においても時間方向の相関を効果的に利用することが可能となり、3 次元予測の効率を大幅に改善できる。本論文では、動ベクトルの検出を行う正方ブロック ($w \times w$ 画素) のことを MC ブロックと呼ぶ。なお、式 (1) で $K_q = 0$ とおくと図 1 (a) に示した 2 次元予測 (フレーム内符号化) を適用することに対応し、この場合の動ベクトル v は不要となる。提案方式では、 8×8 画素毎のクラス選択情報 (m)、MC ブロック毎の動ベクトル (v)、それぞれのクラスで用いる予測係数の組 ($a_m(k)$) といった適応予測に必要なパラメータをフレーム毎に最適化している。例えば、動きの激しい領域のようにフレーム間相関が低いブロックでは、前フレームに関する予測係数 $a_m(k)$ ($k > K_p$) の値が 0 に近いクラスが選択されることとなるため、2 次元/3 次元予測を明示的に切り替える文献 2) のような手法は基本的に不要である。

これらのパラメータは、フレーム毎に以下のコスト J が最小となるよう決定され、最終的にビットストリームに多重化される。

$$J = \sum_{p_0} L(e | \hat{s}_t(p_0), n) + B_{side} \quad (8)$$

ここで、右辺第1項は式(7)によって見積もられる予測誤差の総符号量、第2項は上に列挙した符号化パラメータの記述に必要な符号量である。このように、提案方式では予測誤差の符号量だけではなく、付加情報自身の符号量も含んだコスト関数の下でパラメータの最適化を図っている⁹⁾。これは大規模な組合せ最適化問題とみなせ、その最適解を一度に求めることは困難である。そこで、以下に示す手順によって個々のパラメータの修正処理を繰り返し、フレーム全体の符号量を徐々に低減させることとする。

- (1) $w \times w$ 画素の MC ブロック毎に平均 2 乗誤差最小 (MMSE) に基づいたブロックマッチングを実行し、整数画素精度の動ベクトル $\mathbf{v}$ の初期値を決定する (インタフレームのみ)。
- (2) パラメータ $m, Th_m(n), a_m(k)$ の初期値を決定する。この処理には、確率密度関数がガウス分布に従うという仮定の下で予測誤差の情報量を重み付き 2 乗誤差和で近似し、その最小化に基づいて適応予測を実現する簡易手法⁵⁾を利用する。
- (3) 二つの予測係数 $a_m(i), a_m(j)$ ($1 \leq i, j \leq K_p + K_q; i \neq j$) をランダムに選び、その組合せを探索範囲内でコスト J の値が最小となるよう更新する。ただし、上記の探索範囲は予測係数の量子化精度 h_a によって決まる離散点の集合とする⁶⁾。この処理を各クラス ($m = 1, 2, \dots, M$) について一定回数ずつ繰り返す。
- (4) 各クラスについて、特徴量 U の量子化に用いる閾値 $\{Th_m(1), Th_m(2), \dots, Th_m(15)\}$ の組合せを、動的計画法に基づいて決定する⁵⁾。
- (5) 各コンテキスト ($n = 1, 2, \dots, 16$) について、確率モデルのパラメータ c_n を 0.2~3.2 の範囲で 0.2 刻みで変化させ、コスト J が最小となる値に更新する。
- (6) 8×8 画素のブロック毎に M 種類の 3 次元予測器をそれぞれ試し、コスト J が最小となるクラス m を選択し直す。
- (7) MC ブロック毎に、動ベクトル $\mathbf{v}$ の終点を周囲 8 近傍の画素の中でコスト J が最小となる位置に修正する (インタフレームのみ)。
- (8) (3)~(7) の手順をフレーム全体で求めたコスト J が減少しなくなるか、予め設定した反復回数の上限 (100 回) に達するまで繰り返す。

5. 可変ブロックサイズ動き補償の導入

先に述べたように、提案方式では MC ブロック毎に整数

図3 差分動ベクトルの確率モデルの例
Example of a probability model for differential motion vectors.

画素精度の動ベクトル $\mathbf{v}$ を必要とする。これらの動ベクトルを符号化する際は、隣接ブロックからのメディアン予測によって求めた予測ベクトルとの差分値 $(\Delta v_x, \Delta v_y)$ を 2 次元のベクトルのまま算術符号化する。ただし差分動ベクトルの生起確率は、図3のような一般化ガウス関数を拡張した 2 次元の関数 $\tilde{P}_v(\Delta v_x, \Delta v_y)$ でモデル化できるものとし、その分布形を決めるパラメータ σ_v, c_v については、フレーム毎にその適切な値をダウンロードしている¹⁰⁾。

$$\tilde{P}_v(\Delta v_x, \Delta v_y) = \alpha \cdot \exp \left\{ - \left| \eta(c_v, \sigma_v) \cdot \sqrt{\Delta v_x^2 + \Delta v_y^2} \right|^{c_v} \right\} \quad (9)$$

ここで、 α は動ベクトルの探索範囲 $([-15, +15])$ で決まる差分動ベクトルの定義域について確率の和を 1 とする正規化係数である。

一般に、MC ブロックのサイズ (w) を小さく設定すると、MC の性能が向上して予測誤差の符号量は低減するが、一方で動ベクトル数の増加に伴って付加情報の増大を招くというトレードオフの関係が生じる。MC ブロックのサイズと符号化効率の関係について調査した結果¹¹⁾では、多くの画像で $w = 16$ の時に全体の符号量が最小となることが示されている。しかし、実際には MC ブロックのサイズの最適値は、画像の局所的な性質に大きく依存すると考えられ、これを可変とすることでさらなる符号化効率の改善が期待できる。そこで本論文では、四分木分割¹²⁾に基づいた可変ブロックサイズ MC の手法を導入する。具体的には、まず 4. の (1) で述べた処理において、3 種類のサイズ $w = 8, 16, 32$ の MC ブロックについてそれぞれ動ベクトル $\mathbf{v}$ の初期値を求めておく。また、(2) の処理では $w = 16$ のサイズを前提に符号化パラメータ $m, Th_m(n), a_m(k)$ の初期値を決定する。次に (7) の処理において、 32×32 画素 ($w = 32$) の基本ブロックを $w = 8$ になるまで再帰的に 4 分割すると共に、対応するサイズの動ベクトル $\mathbf{v}$ の修正を行い、各分割状況におけるコスト J の値を記憶しておく。このときの分割状況は、図4のように $w = 32$ の基本

図 4 四分木に基づいた可変ブロックサイズ MC
Quadtree-based variable block-size motion compensation.

ブロックをルートとする四分木で表現され、分割によって枝となるノード（ブロック）には‘1’、末端の葉となるノードには‘0’のフラグが割り当てられる。ブロックサイズの符号化は、各ノードの階層と周囲のブロック分割状況（同一階層のノードにおける‘1’の数）をコンテキストとするフラグ情報の適応 2 値算術符号化により行う。この符号量も含めたコスト J の最小値を求め、対応する MC ブロックのサイズと動ベクトルの組合せを次の反復処理において採用する。これら一連の手順により、MC ブロックのサイズとそれを前提とした各種パラメータを同時に最適化することが可能となる。

ところで MPEG などの非可逆符号化方式では、MC の性能を改善するため、動ベクトルを非整数画素精度で検出することが一般的である。しかし、提案方式において動ベクトルを $1/2$ 画素精度で求めた際の符号化効率を調査した結果からは、動ベクトルの精度を向上させたことによる明確な利得は見出せなかった¹¹⁾。これは、非整数画素位置の予測値を算出する際に、近傍画素値の線形結合による内挿処理が行われるのに対し、提案方式では式 (1) の第 2 項に示す予測係数 $a_m(k)$ ($k > K_p$) を最適化することで、より自由度の高い内挿処理を等価的に実現しているためであると考えられる。したがって、予測器の種類 M および前フレームに配置した参照画素の数 K_q が充分多いならば、提案方式において非整数画素精度の動ベクトルを検出する必要はないといえる。

6. 特性評価

提案方式の性能を評価するため、C 言語を用いて符号器・復号器をソフトウェア実装し、付加情報を含めた符号化レートの調査を行った。以下の実験では、すべて CIF サイズのモノクローム動画（352×288 画素、30Hz、Y 信号）を使用し、25 フレーム分の符号化区間を一つの GOP（最初の 1 枚のみイントラフレーム）として時系列順に符号化を行った。図 5 に使用した各動画の先頭フレームを示す。

まず、付加情報の符号量に影響を与えるパラメータの適切な設定値の範囲を知るため、予測器の種類 M 、予測係数の量子化精度 h_a 、予測次数（当該フレームおよび前フレームの参照画素数 K_p, K_q ）の各設定値と符号化レートの関係について調査した。図 6 は、予測器の種類（クラス数）

図 5 テスト動画
Test sequences.

図 6 予測器の種類 (M) と符号化レートの関係
Coding rates vs. the number of predictors (M).

M を変化したときのインタフレーム（24 枚分）に関する平均符号化レートを示している。ただし、予測係数の量子化精度および予測次数に関するパラメータは、それぞれ $h_a = 1/64$, $K_p = 20$, $K_q = 25$ に固定した。提案方式では、 8×8 画素のブロック毎にどの予測器を選択したかを示すクラス選択情報 (m) およびクラス毎に $K_p + K_q$ 個の予測係数を付加情報として符号化しており、 M の値を増加させることでこれらの付加情報も増大する。図 6 において、 M の値を大きくすることで一般に符号化効率は改善されるものの、 $M \geq 20$ の範囲では符号化レートがほぼ一定となっている様子を確認できる。これは、選択可能な予測器の種類が増えたことによって、適応予測の精度は向上するものの、同時に上記の付加情報が増大することで、その効果が相殺されてしまうためであると考えられる。一方、予測器の種類を $M = 24$ に固定し、予測係数の量子化精度 h_a のみを変化させた場合について同様な調査を行った結果を

図7 予測係数の量子化精度 (h_a) と符号化レートの関係
Coding rates vs. precision of prediction coefficients (h_a).

図8 参照画素数 (K_p, K_q) と符号化レートの関係
Coding rates vs. the number of pels used for the prediction (K_p, K_q).

図7に示す。このとき、 h_a のとり得る値は2のべき乗に限定されるものとする。これにより、式(1)に示す予測値の算出が整数演算とビットシフトのみで構成できるようになり、計算速度の向上が期待できる⁹⁾。提案方式では、予測係数 $a_m(k)$ の値をステップ幅 h_a で一様量子化後に符号化しているため、 h_a を小さな値に設定することで予測値算出の精度は高くなるものの、予測係数自身の符号量は増大してしまう。このトレードオフ関係により、いずれの画像についても $h_a = 1/64$ の時には最適な性能となっている様子が図7より確認できる。また図8は、当該フレームおよび前

表1 符号化レートの比較 (bits/pel)
Comparison of coding rates (bits/pel).

Sequence	VBS (Proposed)	FBS	FBS-MMSE	Intra
Carphone	2.727	2.736 (0.3)	2.831 (3.7)	3.039 (10.3)
Container	2.345	2.349 (0.1)	2.397 (2.1)	3.536 (33.7)
Foreman	2.833	2.851 (0.6)	2.937 (3.5)	3.368 (15.9)
Mobile	3.698	3.711 (0.3)	3.867 (4.4)	4.804 (23.0)
News	1.356	1.359 (0.2)	1.447 (6.3)	2.883 (53.0)
Tempete	3.510	3.525 (0.4)	3.651 (3.9)	4.479 (21.6)
Average	2.745	2.755 (0.4)	2.855 (3.9)	3.685 (25.5)

フレームの参照画素数の組合せ K_p, K_q を変化させたときの符号化レートを表している。ただし、それぞれのフレームに配置する参照画素群は、図1に示すように符号化対象画素 p_0 または前フレーム上の中心画素 q_0 との市街地距離が、一定値以内のものをすべて含むように設定している。したがって、各参照画素数の取り得る値は $K_p = 2, 6, 12, 20, \dots$ 、 $K_q = 1, 5, 13, 25, \dots$ となる。図8より、 K_p, K_q の値を大きくするほど符号化レートも低減する傾向が認められるが、いずれか一方の値を大きくするだけでは不十分であり、当該フレームおよび前フレームにバランス良く参照画素を配置することが重要であることがわかる。また、 K_p, K_q の適切な設定値の範囲は画像により異なるものの、 $K_p \geq 20$ かつ $K_q \geq 25$ であればほぼ最適な特性が得られている。これらの調査結果を踏まえ、以降の実験ではインタフレームにおけるそれぞれのパラメータの設定値として、 $M = 24$ 、 $h_a = 1/64$ 、 $K_p = 20$ 、 $K_q = 25$ を採用する。同様な調査をイントラフレームについても行っており、GOP内の先頭フレームに対しては $M = 24$ 、 $h_a = 1/64$ 、 $K_p = 30$ 、 $K_q = 0$ と設定することとした。

次に、提案方式 (VBS) およびこれをベースに一部符号化条件を変更した比較方式について求めた符号化レート (イントラフレームを含む25フレーム分の平均) を表1に示す。ただし、括弧内の数値は次式で定義されるレートの削減率 (%) である。

$$\left(1 - \frac{\text{提案方式の符号化レート}}{\text{比較方式の符号化レート}}\right) \times 100 (\%) \quad (10)$$

ここで、比較方式の「FBS」はMCブロックのサイズを $w = 16$ に固定した方式、「FBS-MMSE」は復号手順は「FBS」と共通であるが、4で述べた符号化パラメータの最適化手順の中で式(8)で定義されるコストの代りに予測誤差の2乗和を用いる方式、「Intra」は提案方式のイントラフレームに対する符号化手順をGOP全体に適用する方式である。表1において、提案方式は「Intra」に比べて10～53%低い符号化レートを示しており、MCに基づいた3次元予測が符号化効率の改善に大きく寄与していることがわかる。また、「FBS」の符号化レートは「FBS-MMSE」のそれよりも約3.5%低く、動ベクトルの検出や予測器の設計にMMSE規範を用いることが符号化レート最小という意味で最適ではない⁹⁾ということを改めて示す結果となっている。さらに、可変ブロックサイズのMCを導入した提

(a) Foreman

(b) Mobile & Calendar

図 9 MC ブロックの分割状況 (第 2 フレーム)
Result of variable block-size partitioning (2nd frame).

表 2 他の符号化方式との性能比較
Performance comparison with the other lossless coding schemes.

Sequence	VBS (Proposed)	FRExt	JPEG-LS	JPEG 2000
Carphone	2.727	3.127 (12.8)	3.493 (21.9)	3.640 (25.1)
Container	2.345	2.790 (15.9)	4.144 (43.4)	4.129 (43.2)
Foreman	2.833	3.222 (12.1)	3.881 (27.0)	4.082 (30.6)
Mobile	3.698	4.187 (11.7)	5.403 (31.6)	5.589 (33.8)
News	1.356	1.632 (16.9)	3.411 (60.3)	3.578 (62.1)
Tempete	3.510	3.940 (10.9)	4.848 (27.6)	5.076 (30.9)
Average	2.745	3.150(12.9)	4.197(34.6)	4.349(36.9)

案方式は固定サイズの「FBS」に比べて、符号化レートを平均で 0.01 bits/pel 削減しており、四分木分割に基づいた可変ブロックサイズ MC の有効性を表している。図 9 は画像 'Foreman', 'Mobile & Calendar' における MC ブロックの分割状況を示したものであり、画像のテクスチャや動きの複雑度に応じて適切なブロックサイズが選択されている様子を確認できる。

最後に、動画像および静止画像用の代表的な可逆符号化方式と提案方式との性能比較を行った結果を表 2 に示す。ここで、「FRExt」は動画像符号化の国際標準方式である H.264¹³⁾ を拡張した Fidelity Range Extensions¹⁴⁾ の可逆符号化モード (IBBBP..., CABAC, 参照フレーム 5 枚) を意味し、実験には参照ソフトウェアの JM 9.2 を使用した。また、「JPEG-LS」¹⁵⁾ および「JPEG 2000」¹⁶⁾ はそれぞれ静止画像用の国際標準方式をフレーム毎に独立に適用する方式である。表内の数値の意味は表 1 の場合と同一である。H.264 の高度なフレーム間符号化技術に基づいた「FRExt」は、フレーム内符号化方式である「JPEG-LS」

や「JPEG 2000」よりも大幅に優れた符号化特性を示していることがわかる。これに対し、提案方式は「FRExt」よりもさらに 10 % 以上低い符号化レートを達成しており、動画像の可逆符号化方式として高い水準の性能を有していることがわかる。

提案方式では、フレーム毎に各符号化パラメータを反復的に最適化しているため、必然的に符号化処理の演算量が多くなってしまふ。実際に上記の符号化実験に要した符号化時間は、汎用のパーソナルコンピュータ (Intel Xeon processor, 3.60 GHz, Linux OS) 上で 1 フレーム当り 3~4 分であった。一方、復号側では上記の最適化処理が一切不要であるため、原理的に高速な動画再生が可能である。現時点ではプログラムの最適化が充分とは言えないが、同一計算機上で毎秒 40~50 フレーム程度を復号可能であることを確認している。また、演算量の大部分が 3 次元予測に必要な計算で占められているため、用途によっては予測回数に関するパラメータ K_p , K_q を小さな値に設定することで、圧縮率は多少犠牲となるものの復号速度をさらに高速化することが可能である。

7. む す び

本論文では、動画像用の高能率な可逆符号化方式について検討した。提案方式は、先に静止画像用に開発した可逆符号化方式⁶⁾を動画像用に拡張したものであり、復号処理に必要な符号化パラメータを実際の符号量に相当するコスト関数の下で繰り返し最適化している点が特徴である。また、動画像特有のフレーム間相関を効果的に利用するため、可変ブロックサイズ MC に基づいたブロック適応 3 次元予測方式を新たに導入している。CIF サイズのモノクローム動画像を対象に符号化実験を行った結果、提案方式は最新の国際標準方式である H.264 を拡張した FRExt 方式よりも 10 % 以上優れた圧縮率を達成するなど、動画像データの可逆符号化方式として高い性能を有していることが明らかとなった。本論文では、予測器の種類 M および予測回数に関するパラメータ K_p , K_q の適切な設定値を実験的に決定したが、これらのパラメータの最適値は画像のサイズや内容に大きく依存すると考えられるため、画像毎にその設定値を自動決定する手法の開発が今後の課題の一つに挙げられる。

〔文 献〕

- 1) B. Martins and S. Forchhammer: "Lossless Compression of Video Using Motion Compensation", Proc. Data Compression Conf. (DCC '98), p.560 (Mar. 1998)
- 2) K. H. Yang and A. F. Faryar: "A Context-Based Predictive Coder for Lossless and Near-Lossless Compression of Video", Proc. IEEE Intl. Conf. on Image Processing (ICIP 2000), I, pp.144-147 (Sep. 2000)
- 3) S. Sun and S. Lei: "On Study of a Lossless Video Coding Algorithm Based on H.26L Tools", Proc. SPIE Image and Video Communications and Processing (IVCP 2003), **5022**, pp.994-1000 (Jan. 2003)
- 4) D. Brunello, G. Calvagno, G. A. Mian and R. Rinaldo: "Lossless Compression of Video Using Temporal Information", IEEE

Trans. Image Processing, **12**, 2, pp.132-139 (Feb. 2003)

- 5) 松田, 森, 前田, 伊東: “可逆符号化のためのレート最小とする予測器の設計と評価”, 信学論, **J85-D-II**, 3, pp.448-456 (Mar. 2002)
- 6) 松田, 梅津, 小崎, 前田, 伊東: “画像毎に適応予測器と算術符号を最適化する可逆符号化方式”, 信学論, **J88-D-II**, 9, pp.1798-1807 (Sep. 2005)
- 7) 白井, 塩寺, 松田, 伊東: “対象画像毎に予測器と算術符号を反復最適化する可逆符号化 —予測誤差の生起確率の推定精度改善—”, 2003 情報科学技術フォーラム (FIT 2003), J-54, pp.305-306 (Sep. 2003)
- 8) 奥村: “データ圧縮の基礎から応用まで”, C Magazine, **14**, 7, pp.13-35 (Jul. 2002)
- 9) 小崎, 松田, 伊東: “対象画像毎に予測器と算術符号を反復最適化する可逆符号化 —付加情報も含んだコスト関数の導入—”, 2004 信学総合大, D-11-8 (Mar. 2004)
- 10) 加納, 松田, 伊東: “差分動ベクトルの2次元PDFモデルとその適応算術符号化への応用”, 2004 信学ソサエティ大, A-6-3 (Sep. 2004)
- 11) T. Shiodera, I. Matsuda, S. Itoh: “Lossless Video Coding Using MC Based 3D Prediction Optimized for Each Frame”, Proc. Intl. Workshop on Advanced Image Technology (IWAIT 2004), pp.159-163 (Jan. 2004)
- 12) G. J. Sullivan and R. L. Baker: “Efficient Quadtree Coding of Images and Video”, Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP '91), **4**, pp.2662-2664 (Apr. 1991)
- 13) ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC: “Advanced Video Coding for Generic Audiovisual Services” (2003)
- 14) JVT of ITU-T and ISO/IEC: “Draft Text of H.264/AVC. Fidelity Range Extensions Amendment”, JVT-L047 (Sep. 2004)
- 15) ISO/IEC, ISO/IEC 14495-1:1999: “Information Technology – Lossless and Near-Lossless Compression of Continuous-Tone Still Images: baseline” (Dec. 1999)
- 16) ISO/IEC JTC 1/SC 29/WG 1, ISO/IEC FCD 15444-1: “Information Technology – JPEG 2000 Image Coding System” (Mar. 2000)

しおでらたいちろう

塩寺 太一郎 2003 年, 東京理科大学理工学部電気工学科卒業。2005 年, 同大学大学院修士課程修了。同年, (株) 東芝入社。在学中は画像の可逆符号化に関する研究に従事。

うめづ ゆうじ

梅津 有司 2004 年, 東京理科大学理工学部電気工学科卒業。2006 年, 同大学大学院修士課程修了。同年, 富士通 LSI ソリューション (株) 入社。在学中は画像の可逆符号化に関する研究に従事。

まえだ ひろき

前田 弘樹 2005 年, 東京理科大学理工学部電気工学科卒業。現在同大学大学院修士課程在学中。画像の可逆符号化に関する研究に従事。

まつだ いちろう

松田 一郎 1991 年, 東京理科大学理工学部電気工学科卒業。1996 年, 同大学大学院博士後期課程修了。同年より, 東京理科大学理工学部電気工学科に勤務し, 現在, 同大学電気電子情報工学科講師。画像符号化等の研究に従事。博士 (工学)。電子情報通信学会, IEEE 各会員。

もりおか かずゆき

森岡 一幸 2000 年, 東京大学工学部電気工学科卒業。2005 年, 同大学大学院電気工学専攻博士課程修了。同年, 東京理科大学理工学部電気電子情報工学科助手。2006 年, 明治大学理工学部電気電子工学科講師。ロボットビジョン等の研究に従事。博士 (工学)。日本ロボット学会, 電子情報通信学会各会員。

いとう すん

伊東 晋 1976 年, 東京大学工学部電子工学科卒業。1981 年, 同大学大学院電子工学専攻博士課程修了。同年より, 東京理科大学理工学部電気工学科に勤務し, 現在, 同大学電気電子情報工学科教授。画像処理, 画像符号化等の研究に従事。工学博士。電子情報通信学会, 情報処理学会, 電気学会各会員。