

第 7 章

動画像符号化のための可変形状領域分割動き推定方式

7.1 まえがき

従来より、さまざまな方式が検討されてきた動画像符号化方式は、近年、ITU-T 勧告 H.261 や H.262(MPEG-2)[1, 2] により標準化が行われた。これらは、現在の技術レベルで実現可能なものとしてまとめられたものである。そのため、比較的簡単な要素技術の組合せとなっており、例えば、ブロックベースの符号化を基本としている。すなわち、画像を多数の方形ブロックに分割し、ブロック毎に動き補償予測を行い、ブロック DCT を行っている。

この方法では、異なる物体や背景に属し信号の性質の違う領域が混在するブロックにおいて、符号化効率が低下するという問題がある。その理由としては、そのブロックでは (1) 動き補償による予測効率が低下すること、(2) 予測誤差画像にエッジなど高周波成分が含まれ、DCT による圧縮効果が低下することなどがあげられる。

この問題を避けるために、画像を性質の似た領域毎に分割して符号化を行なう領域分割符号化の潜在力は大きいといえる。領域分割符号化は、もともと静止画を対象として、任意形状領域の直交変換符号化法と組み合わせて検討されてきた [3, 4]。これを動画像に適用した領域分割に基づく動き補償予測符号化では、各領域毎に動き推定をする必要がある。

従来より、動画像の領域分割と動き推定処理は、互いに密接に関係付けながら同時に行なうことが望ましいとされ、幾つかの方法が提案されている [5, 6, 7]。これらは、フレーム内挿やシーン記述への応用の立場から、分割処理結果、推定動ベクトルが、なるべく画像の見た目と一致することを目的としている。

一方、動き補償予測を目的とする領域分割・動き推定法として代表的な方式に、分離統合法 [8, 9, 10] がある。しかしながらこの方式は、符号化の立場から必ずしも最適とは言えない。その理由としては、7.2で述べるように、(1) 分離・統合でのパラメータ設定により性能が大きく左右されること、(2) 動ベクトル・形状・予測誤差の各情報量を同時に考慮しつつ統合を行なうことが困難であること、(3) 隣接領域毎の統合が基本となっているため、局所最適に陥る可能性が高いことがあげられる。そこで、これらの問題のない領域分割動き推定法が望まれる。

本章では、この目的のため、予測誤差情報量、動ベクトル情報量、形状情報量の全発生情報量を最小にする意味で最適な、動的計画法を用いた領域分割と動き推定の同時処理方式を提案する。以下、7.2で動き補償予測のための領域分割・動き推定法の問題点について述べ、これを解決する最適方式を 7.3で提案する。7.4で種々の一般動画像を用いて計算機シミュレーションにより提案方式の有効性を示す。

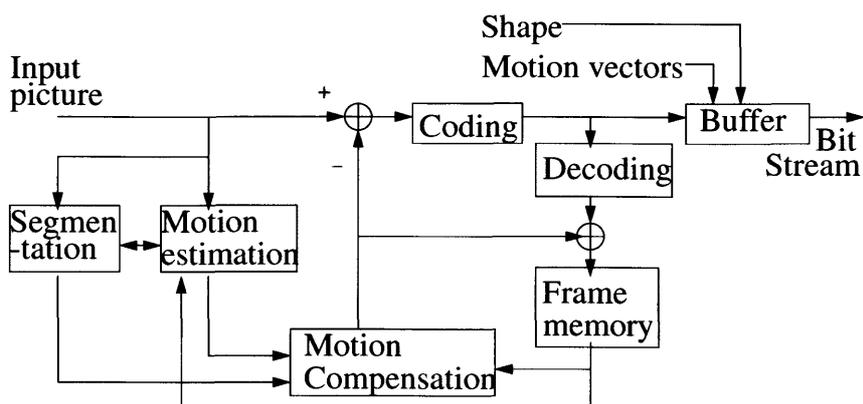


図 7.1: 領域分割動き補償予測符号化.

7.2 動き補償予測のための領域分割・動き推定法

動画像の領域分割に基づく動き補償予測符号化のブロック図を図 7.1 に示す。矩形ブロックマッチングによる動き補償予測符号化と異なり、領域分割と動き推定の両方に基づき動き補償予測が行なわれるため、一般的に予測誤差が減少するが、分割形状情報を新たに伝送する必要がある。

従来、領域分割動き補償予測のための代表的な動き推定・領域分割の方法として、分離統合法がある。これは、まず画像を何らかの方法で細かく分離（クラスタリング）し、その後統合を行なうものである [8, 9, 10]。[8] では K 平均クラスタリングを使用して分離を行ない、その結果の各領域の動き補償予測誤差を基にして統合を行なう。[9] では、エッジ検出を行ないフレーム間の相関も利用したクラスタリングの後、領域毎の動き補償予測誤差に基づく領域統合を、閾値を設定することにより行なっている。また [10] では、閾値を用いた四分木可変サイズブロック分割を行なった後、動ベクトルと予測誤差の情報量を勘案した統合を行なう。

しかしながらこれらの方法には、次のような問題点があるといえる。

- (1) 分離及び統合時の閾値などのパラメータに性能が左右されるため、画像によって特性に大きな差がでる。
- (2) 統合時に、動ベクトル情報量、形状情報量、予測誤差情報量を全て考慮に入れて処理することが困難である。
- (3) 隣接領域毎の統合が基本となっているため、局所最適に陥る可能性が高い。

動き補償予測のための領域分割と動き推定は、オブジェクト指向の符号化である場合（MPEG-4[11] など）を除いては、その目的は全発生情報量（予測誤差情報量、動ベクトル情報量、形状情報量の和）を最小にするものが最適であるといえる。その意味からは、従来の方法は必ずしも最適な領域分割・動き推定法とはなっていない。

そこで、動き補償予測のための領域分割動き推定法として、動的計画法に基づく最適化方式を提案する。

7.3 動的計画法に基づく領域分割・動き推定同時処理方式

本方式では、動き推定・領域分割を、全発生情報量が最小となるように各動ベクトルに領域を割り当てる最適化問題とみなし、次の手順で動的計画法を適用する。

- (1) 各動ベクトル, 各領域パターン毎に, その発生情報量を見積もり, テーブル化する.
- (2) 動的計画法における段階を動ベクトル, 状態を領域パターンとし, 前記テーブルを利用して各動ベクトルの占有すべき領域を得る.

情報量の見積りに関し 7.3.1, 7.3.2で述べた後, 7.3.3で動的計画法の簡単な適用例を示す.

7.3.1 発生情報量の見積もり

ある動ベクトル \mathbf{v} の占める領域が領域パターン k であった時の, その領域からの発生情報量 $E(\mathbf{v}, k)$ は, 動き補償予測誤差情報量 E_E と領域形状情報量 E_S , 動ベクトル情報 E_V の和として, 次のように表すことができる.

$$E(\mathbf{v}, k) = E_E(\mathbf{v}, k) + E_S(k) + E_V(k) \quad (7.1)$$

各情報量は, 次のように見積もることができる. ここで, 「領域パターン」とは, 後の図 7.3に示すように, 単独または複数の連結領域からなることに注意する.

(1) 予測誤差情報量

領域分割動き補償予測誤差は, 領域毎に符号化されると想定できる. 各領域からの予測誤差情報量の見積もりとしては, 使用する符号化法が確定している時は, その方式で符号化してみるのが最も正確である. しかしここでは, 特定の符号化法を前提としないという条件のもと, なるべく妥当な見積りができる方法として, 以下の方法を使用する.

動き補償予測誤差信号の分布は一般にラプラス分布で近似できる. この時, 動ベクトル \mathbf{v} による領域パターン k の動き補償予測誤差分散を σ^2 , k 内総画素数を N 個とすると, 予測誤差情報量 $E_E(\mathbf{v}, k)$ は

$$E_E = \frac{N}{2} \log_2 2e^2 \sigma^2 \quad [\text{bits}] \quad (7.2)$$

と見積もることができる.

(2) 動ベクトル情報量

動ベクトル 1 個あたりの平均のビット量 b を固定的に割り当てる. 領域パターン k が a 個の連結領域からなっているとすると, その動ベクトル情報量 $E_V(k)$ は

$$E_V = ab \quad [\text{bits}] \quad (7.3)$$

となる.

(3) 領域形状情報量

領域形状情報量は, 前提とする形状符号化法に依存する. 本方式に利用できる見積もり法には, 次のような条件が必要である.

- 各領域パターン毎に, 独立して見積もりが可能であること.
- すなわち, 当該動ベクトルの占める領域以外の領域の分割形状とは無関係に, その領域だけの形状情報量を見積もれることが, 動的計画法の適用上, 必要となる.

7.3.2 領域形状情報量の見積もり法

領域形状情報を伝送する手段としては, 輪郭線を伝送する差分チェーン符号がある. これは簡易である反面, 次のような欠点がある.

- (1) 込み入った領域では, 符号量が非常に大きくなる.
- (2) 符号量は, 輪郭線の形状よりむしろその長さに依存する [12].

これらは、我々の目標とする種々の動画像へ適用可能な領域分割符号化への応用の立場からは、特に問題となる。すなわち、(1)のように領域を細かくするにつれての形状符号量増加が急激であると、予測誤差情報量の減少分を打ち消す傾向が強くなる。また、(2)については、仮に形状が矩形である場合とそうでない場合とで見積もり符号量による差別化ができず、画像によってはこれがブロックマッチング法に比べて性能低下の原因となり得るからである。

そこで我々は、

- (1) 領域が細かくなっても、符号量の増加が比較的緩やかで、
- (2) 大きな矩形の形状であるほど符号量が少なくなり、
- (3) さらに、動的計画法への適用上、領域パターン毎に独立に形状情報量見積もりが可能な、領域形状符号量見積もり法として、グラフ理論の四色定理を応用した階層的形状符号化法を考案し利用することとした。以下に詳述する。

四色定理を応用した形状表現法

四色定理によると、任意の領域分割状態は4色で塗分けられる。すなわち、画像上の最小分割単位毎に、2ビットで表現できることとなる。この四色塗分けのアルゴリズムは、文献[13]などにある。

したがって、サイズ4x2のブロックを最小分割単位とする時、このブロック毎に2ビットを割り当てれば、完全に分割形状を表現できることとなる。しかしこれでは2ビット/8画素=0.25bppとなり、多すぎる。そこで、この2ビットの割当を階層化ブロック毎に行う。図7.2に分割形状の簡単な例を示す。この例では、ちょうどA~Dの4記号で分割領域が表現されている。最小分割単位はサイズ4x2の小ブロック、その上位階層ブロックは、サイズ8x8と16x16のそれぞれ中、大ブロックであるとする。本節では、領域パターン毎でなく、この画面全体の分割形状の符号量を計算する(次節でパターン毎の見積りに応用する)。

まず大ブロック層からの発生情報量は、次のように計算する。

- そのブロックが領域境界を含む場合：1ビット(領域境界含有フラグオン)
- そのブロックが領域境界を含まず、1色(記号)で表される場合：上記のフラグ1ビット(オフ)+色(記号)情報2ビット、計3ビット

図7.2では、

$$1\text{bit} \times 3\text{個} + 3\text{bit} \times 17\text{個} = 54\text{bits}$$

となる。領域境界含有フラグがオンの大ブロックに限り、中ブロック層に入る。中ブロック層からの発生情報量計算は、大ブロックに同じである。図7.2では、

$$1\text{bit} \times 3\text{個} + 3\text{bit} \times 9\text{個} = 30\text{bits}$$

となる。次に、領域境界含有フラグがオンの中ブロックに限り、小ブロックに入る。小ブロック層からの発生情報量は、ブロック毎に、色情報2ビットのみである。図7.2では、

$$2\text{bit} \times 24\text{個} = 48\text{bits}$$

となる。3階層で計132bits(1画素当たりの平均0.026bpp)となる。

この方式の4分木分割法との大きな違いは、4分木分割は表現可能な分割形状に大きな制約がある(上位階層ブロックをまたがった領域を表現できない)のにたいし、本方式ではそのような制約がないという点である。

次に、実際の領域分割例に対して、本方式と、従来の差分チェーン符号で表現した場合の比較を表7.1に示す。ここで用いた分割画像は、7.4の計算機シミュレーションによる領域分割結果画像

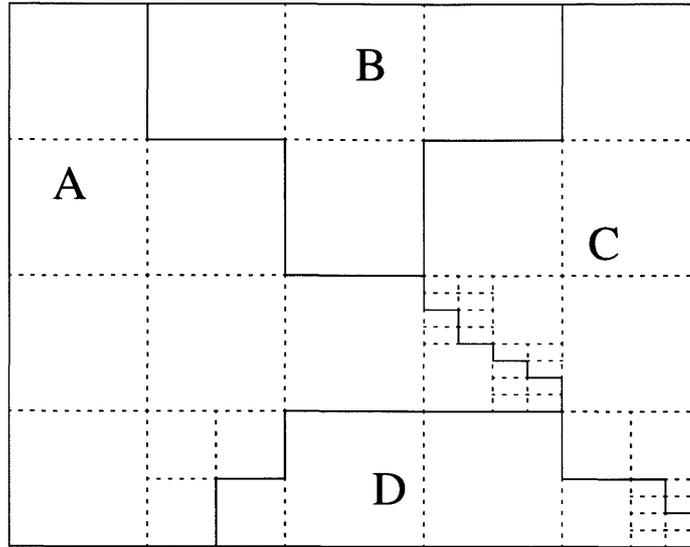


図 7.2: 分割形状の簡単な例.

表 7.1: 階層的三色塗り分け方式 (本方式) と差分チェーン符号化方式での発生形状符号量 [bpp] の比較.

| 本方式 | 差分チェーン符号 |
|------|----------|
| 0.17 | 0.65 |

(cheerleaders) である (図 7.7). 分割形状が細かいため, 差分チェーン符号では大幅に形状符号量が増大することが分かる.

形状符号量見積もりへの適用

上記の方式を 7.3.1 の動的計画法における形状符号量見積もりへ適用する. すなわち, 領域パターン毎の見積りに応用する.

ある動ベクトルの領域パターンとして, 図 7.3 の網点部のような形状の符号量見積もりを行なう例を考える. まず, 領域境界を含まない大中小ブロックからの発生形状情報量は,

$$3\text{bits} \times 8\text{個} + 3\text{bits} \times 4\text{個} + 2\text{bits} \times 10\text{個} = 56\text{bits}$$

次に, 領域境界を含む大中ブロックからの発生形状情報は, フラグ 1 ビットのみであり, これは隣接領域と共有されていると言える. したがって, 見積もりの上では, 1 領域あたり平均 $1/(\text{そのブロック中に混在する領域数})$ ビットとなる. 混在領域数はこの段階では不明である. しかしいずれにせよ, このビット数は, 前記のものと比較して充分小さいため, 無視できる. 以上より, 図 7.3 の例では, 形状符号量見積もりは 56 ビットとなる.

このように形状符号量を見積もることにより, 大きな矩形形状ほど, 見積もりビット数が小さくなるため, 分割形状として計算上優先されることとなる.

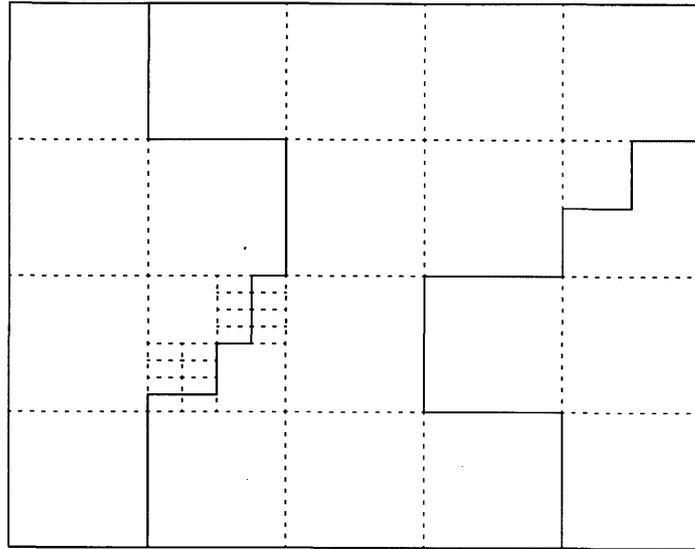


図 7.3: 領域パターンの例 (網点部) .

7.3.3 動的計画法の適用

7.3.1のようにして、各候補動ベクトル毎に、それが占有し得る領域のパターンそれぞれについて、発生情報量 $E(\mathbf{v}, k)$ を見積もることができる。これをテーブル化し、それをを用いて、次のように動的計画法を適用する。候補動ベクトルの数を L 個とする。

段階 t : t 番目の候補動ベクトル \mathbf{v}_t

状態 x_t : 候補動ベクトル $\mathbf{v}_t, \mathbf{v}_{t+1}, \dots, \mathbf{v}_L$ の占める領域 R_t, R_{t+1}, \dots, R_L の全体 $\cup_{i=t}^L R_i$ (累積領域)

方策 k : 候補動ベクトル \mathbf{v}_t の取り得る領域パターン選択肢

コスト関数 $f_t(x_t)$: 状態 x_t のときの、動ベクトル $\mathbf{v}_t, \mathbf{v}_{t+1}, \dots, \mathbf{v}_L$ の占有する領域からの発生情報量の和の最小値

以上より、最適化原理式は次の式 (7.4) となる。

$$\begin{aligned}
 f_L(x_L) &= E(\mathbf{v}_L, x_L) \\
 f_t(x_t) &= \min_{\mathbf{v}k \subseteq x_t} \{E(\mathbf{v}_t, k) + f_{t+1}(x_t \ominus k)\} \\
 &\quad (2 \leq t \leq L-1) \\
 f_1(S) &= \min_{\mathbf{v}k \subseteq S} \{E(\mathbf{v}_1, k) + f_2(S \ominus k)\}
 \end{aligned} \tag{7.4}$$

ここで S は画像全体を意味し、 $x \ominus k$ は領域 x から領域 k を除いた残りの領域を意味する。また、累積領域 $x_t (2 \leq t \leq L)$ は、あらゆる領域パターンをとり得ることに注意する。

式 (7.4) を解くことにより、各動ベクトルの占有すべき領域が決定される。すなわち、全発生情報量を最小にする意味で最適な領域分割と動き推定の同時処理が実現できる。

次に、具体的な処理過程を説明する。簡単な例として、領域の最小単位 3 個だけからなる画像を考える。この場合、どの動ベクトルも、図 7.4 に示すような 8 通りの領域パターンのいずれかをとることとなる。候補動ベクトルを簡単のため 3 個とし、7.3.1 に述べた方法で、各候補動ベクトル毎、各領域パターン毎の発生情報量の見積もりが、表 7.2 のように得られたとする。これを用いて

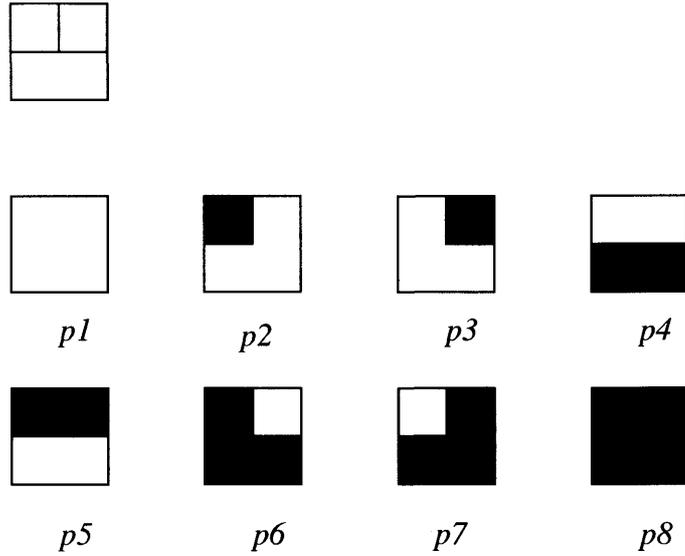


図 7.4: 分割単位 3 個からなる画像と各動ベクトルの取りうる領域パターン (黒色部)

表 7.2: 発生情報量の見積もり $E(\mathbf{v}, k)$ の例.

| 動ベ クトル | 領域パターン k | | | | | | | |
|----------------|------------|-------|-------|-------|-------|-------|-------|-------|
| | p_1 | p_2 | p_3 | p_4 | p_5 | p_6 | p_7 | p_8 |
| \mathbf{v}_1 | 0 | 10 | 20 | 40 | 30 | 60 | 60 | 90 |
| \mathbf{v}_2 | 0 | 20 | 30 | 50 | 50 | 60 | 60 | 80 |
| \mathbf{v}_3 | 0 | 20 | 20 | 10 | 40 | 30 | 20 | 40 |

式 (7.4) の漸化式に従い処理を行う際の過程は、順に表 7.3, 表 7.4, 表 7.5 のようになる. 表 7.3 の作成では, \mathbf{v}_3 のとり得る領域がそのまま累積領域 x_3 となり, 表 7.2 が用いられる. 表 7.4 では, \mathbf{v}_3 と \mathbf{v}_2 で占められ得る累積領域パターン x_2 のそれぞれについて, \mathbf{v}_2 の領域パターン選択肢に関する発生情報量の最小値を求める. この時, 表 7.2, 表 7.3 が利用される. 表 7.5 では, $\mathbf{v}_3, \mathbf{v}_2, \mathbf{v}_1$ の占める累積領域は全画面に他ならないので, x_1 は p_8 のみになる. 各選択肢について, 表 7.2 と表 7.4 から, 発生情報量が計算され, それを最小値 30 にする選択肢 p_2 が \mathbf{v}_1 のとるべき領域と決定される. この時

$$x_2 = x_1 \ominus R_1 = p_8 \ominus p_2 = p_7$$

となり, 表 7.4 より \mathbf{v}_2 の領域は p_1 となる. \mathbf{v}_3 の領域はしたがって

$$x_3 = x_2 \ominus R_2 = p_7 \ominus p_1 = p_7$$

と決定される. 以上のようにして, 動き推定と領域分割の同時処理が実行されることとなる.

次に, 簡単な例で従来の方と本方法の比較を行う. すなわち, 画像 cheerleaders の中の動きの混在する部分 (サイズ 16x16 のブロック) について, その内部をサイズ 4x4 の小ブロック 16 個に分け, これを最小単位としてこの部分の領域分割動き推定を行う. 領域形状の符号量の見積もりには 7.3.2 の方式を用いる. また動ベクトル 1 個当たりの見積もり情報量としては, あらかじめ小ブロック毎に全探索ブロックマッチングをした結果の動ベクトルエントロピーの平均値を用いる.

表 7.3: 最適化原理式の処理過程 ($f_3(x_3)$ の表).

| x_3 | $f_3(x_3)$ |
|-------|------------|
| p_1 | 0 |
| p_2 | 20 |
| p_3 | 20 |
| p_4 | 10 |
| p_5 | 40 |
| p_6 | 30 |
| p_7 | 20 |
| p_8 | 40 |

表 7.4: 最適化原理式の処理過程 (* がついたものが $f_2(x_2)$ の値).

| x_2 | v_2 の領域パターンの選択肢 k | | | | | | | |
|-------|-----------------------|-------|-------|-------|-------|-------|-------|-------|
| | p_1 | p_2 | p_3 | p_4 | p_5 | p_6 | p_7 | p_8 |
| p_1 | 0 | - | - | - | - | - | - | - |
| p_2 | 20* | 20* | - | - | - | - | - | - |
| p_3 | 20* | - | 30 | - | - | - | - | - |
| p_4 | 10* | - | - | 50 | - | - | - | - |
| p_5 | 40* | 40* | 50 | - | 50 | - | - | - |
| p_6 | 30* | 30* | - | 70 | - | 60 | - | - |
| p_7 | 20* | - | 40 | 70 | - | - | 60 | - |
| p_8 | 40* | 40* | 60 | 90 | 60 | 80 | 80 | 80 |

比較に用いた方式は、後述の 7.4.1 の方式 C を最小ブロックサイズ 4x4 にしたものである。計算機シミュレーションによる全発生情報量の比較を表 7.6 に示す。実験を行なった 2 部分とも、提案方式が従来方式より、効率的に動き推定・領域分割され、発生情報量をより小さくできている。

なお、ここでは、動きパラメータとして平行移動動ベクトルのみを扱っているが、その代わりにアフィン変換パラメータを使用することも同じ枠組内で可能である。すなわち、動的計画法における「段階」として、候補動ベクトルの代わりに、アフィンパラメータの組の候補を設定すれば良い。

7.4 計算機シミュレーション実験

ITU-R 標準テスト動画像と ITE 標準 HDTV テスト動画像 (標準 TV サイズに切り出し) 計 6 種類を使用して、上述の方式と、比較対象として全探索ブロックマッチング方式、従来の領域分割動き推定方式 2 種類の、計算機シミュレーションによる比較実験を行う。

7.4.1 計算機シミュレーションのための方式構成

比較シミュレーションの構成を、図 7.5 に示す。

表 7.5: 最適化原理式の処理過程 (* がついたものが $f_1(x_1)$ の値).

| x_1 | v_1 の領域パターンを選択肢 k | | | | | | | |
|-------|-----------------------|-------|-------|-------|-------|-------|-------|-------|
| | p_1 | p_2 | p_3 | p_4 | p_5 | p_6 | p_7 | p_8 |
| p_8 | 40 | 30* | 50 | 80 | 40 | 80 | 80 | 90 |

表 7.6: 動き混在部での提案方式と従来方式による発生情報量 [bpp] の比較.

| | 提案法 | 従来法 |
|------|------|------|
| 部分 1 | 6.41 | 6.51 |
| 部分 2 | 2.66 | 2.70 |

提案法に対するシミュレーション方式構成 (方式 A)

提案法の簡易化 7.3に述べた提案方式は、本来、画素を分割最小単位とし全画面に対して適用するのが、発生情報量最小化のためには最適である。しかし動的計画法の計算量が膨大となるため、次のような簡略化を行なう。これを、図 7.5において 2 回反復される簡易化提案法の処理の 1 回目とする。

- 領域分割の最小単位を、サイズ 4x2 の小ブロックとする。
- 動的計画法の適用を、サイズ 16x8 の大ブロック毎に行う。
- 大ブロック毎の動ベクトル情報量の見積もりにおいて、既に終了した左及び上の大ブロック内の隣接領域の動ベクトルの状況を考慮する。すなわち、7.3.1の(2)の計算において、当該連結領域の動ベクトルが、大ブロック境界を挟んで隣接する領域のベクトルと等しい時には、当該連結領域の動ベクトル情報量を 0 に設定する。これによりその隣接領域の動ベクトルが計算上優先されることとなる。

すなわち、計算量を抑えるため最小分割単位の数を 16 個にしぼる。と同時に、なるべく広範囲での最適化となるよう、最小領域ブロックサイズを 1 画素ではなく 4x2 とし、さらに大ブロック境界を挟む領域どうしの統合を図ることとする。

図 7.5の簡易化提案法の 2 回目の処理では、1 回目の結果に基づき、上・左だけでなく右・下の大ブロック境界を挟んで隣接する領域のベクトルも考慮に入れて情報量見積りを行なった後、動的計画法を適用する。

なお、動的計画法での候補動ベクトルとしては、探索範囲内の動ベクトル全てを使用するのが最も直接的な方法である。しかし、これでは段階数が非常に多くなるため、次の方法をとる。

- (1) 当該大ブロック内で、小ブロック (計 16 個) ごとに全探索ブロックマッチングを行なう。
- (2) その結果の動ベクトル (最大 16 種類) を候補動ベクトルとする。

これ以外の動ベクトルは、領域分割動き推定の結果として、当該大ブロック内で生じる可能性は非常に小さいと考えられる。したがって、この段階数削減手法による性能低下は、ほとんど無視できる。

領域統合 処理を 16x8 単位に簡略化したため、提案法の 2 回の反復が終了した段階では、画像全体では依然として過分割の状態にある。そこで、領域統合を行なう。統合法としては、従来の文献 [9] 等にある方式を、発生情報量の観点を取り込むため、次の (1)-(3) のように変えて用いる。

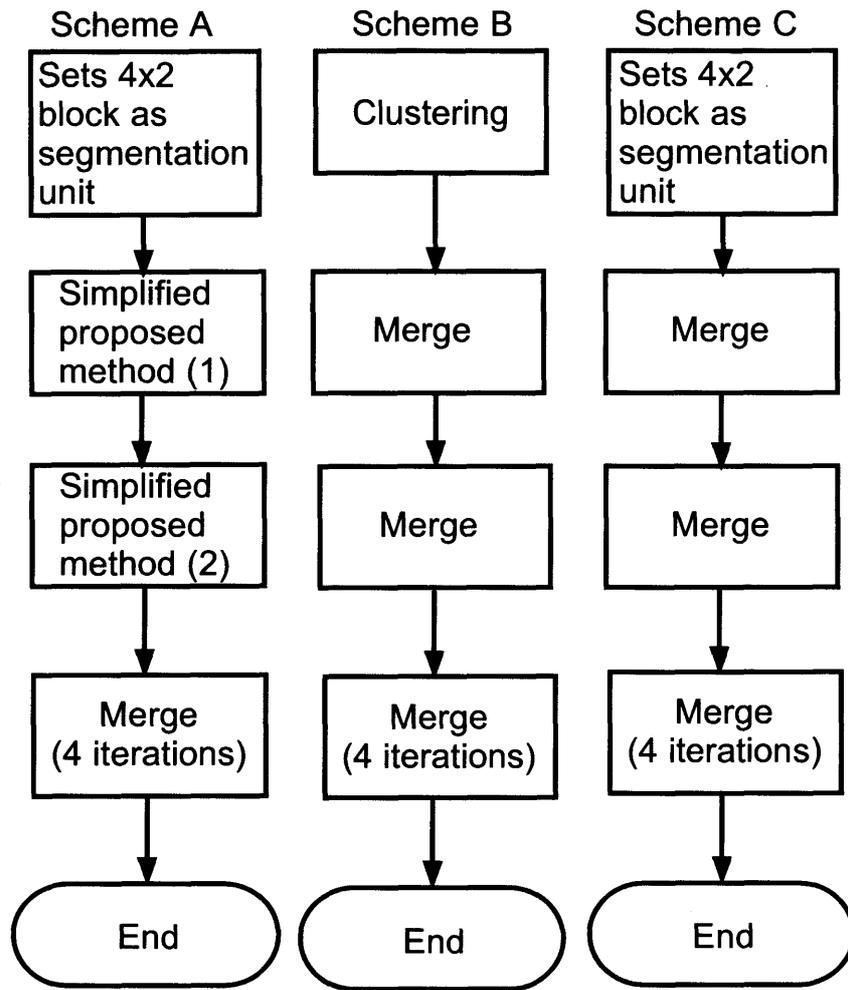


図 7.5: 領域分割動き推定方式の比較のためのシミュレーション構成.

- (1) 全領域の動ベクトルのエントロピーを計算し、動ベクトル 1 個あたりの情報量の見積もり e_v を算出しておく.
- (2) 各領域について、次の (a)-(c) の処理を行う.
 - (a) 当該領域内の画素数を N 、その動ベクトルによる動き補償予測誤差分散を σ_0^2 とする.
 - (b) 隣接領域を抽出する.
 - (c) 各隣接領域 i の動ベクトルによって当該領域を動き補償した場合の予測誤差分散を σ_i^2 とする. 次の式 (7.5) が成立するような隣接領域 i のうち、左辺の値を最も小さくするような領域 i の動ベクトルで当該領域の動ベクトルを置き換える. すなわち、当該領域と領域 i を統合する. この式の左辺は統合による当該領域の予測誤差情報量の増大分を表し、右辺は統合による動ベクトル情報量の減少分を表している.

$$\frac{N}{2} \log_2 \frac{\sigma_i^2}{\sigma_0^2} < e_v \quad (7.5)$$

- (3) (1)(2)の統合処理を、図 7.5にあるように、画像全体で 4 回反復する.

なお、この反復回数は、次に説明する比較方式のものも含め、種々のテスト画像を用いた予備実験により、そのすべての画像で領域統合がほぼ収束するのに十分な反復回数として採用した。

従来の領域分割動き推定方式（方式 B, C）

比較のための従来方式では、まず画像を何らかの方法で細かく分割したあと、ある規範のもとで統合するという方法を用いる（例えば文献 [9]）。

まず、比較方式（1）（方式 B とする）では、図 7.5にあるように、クラスタリングによる細分割の後、前節で説明した統合処理を 6 回反復適用する。クラスタリングには、輝度・色差画素値を用いる。また、クラスタ数が、提案方式のシミュレーション構成における最小分割単位ブロック数（標準 TV サイズで約 21000 個）になるようパラメータを調整する。

比較方式（2）（方式 C とする）では、方式 A と同じ最小分割単位（サイズ 4x2 のブロック）に分割した後、統合処理を 6 回反復適用する。方式 A とこの方式 C の違いは、反復の始め 2 回に提案方式を用いるか、形状情報量を勘案しない従来型の統合法を用いるかである。

ブロックマッチング法（BM 法）

探索方法は全探索、ブロックサイズは H.261, H.262 等の国際標準で採用されている 16x16 とする。

7.4.2 実験結果と考察

6 種類の一般動画像に対する、7.4.1 に述べたシミュレーション方式 A（簡略化した提案法を含む反復方式）による全発生情報量とその内訳を表 7.7 に示す。また、画像 cheerleaders での各方式による全発生情報量とその内訳の比較を表 7.8 に示す。さらに、その際の動き補償予測誤差画像から、注水定理によりレート歪み特性を算出した結果を図 7.6 に示す。表 7.9 には、各方式による画像毎の全発生情報量の比較を示す。

動き探索範囲は水平 -16 ~ +15, 垂直 -8 ~ +7, 整数画素精度とした。また、分割結果の形状符号量計算は、方式 A と方式 C（従来方式 2）については 7.3.2 の方法（ただし大・中・小ブロックサイズはそれぞれ 16x8, 8x4, 4x2）で行ない、方式 B（従来方式 1）については領域境界線を JBIG で符号化した（分割が画素単位であり 7.3.2 の方法では符号量が増大するため）。予測誤差情報量は、領域毎（BM 法ではブロック毎）に予測誤差分散と式 (7.2) に基づき計算し、画像全体で和をとった。領域毎に計算した理由は、予測誤差の符号化は領域毎に行なわれることが想定されるからである（領域形状が既知のため）。

テスト画像の特徴は、次のようになる。

- cheerleaders, green leaves: 異なる動きの混在
- marching in, mobile: 精細
- soccer: 速い動き
- flamingoes: ゆっくりした動き

表 7.7 を見ると、領域数の大小と動ベクトル情報量、形状情報量の大小がほぼ対応していることがわかる。また、方式 A では、動ベクトル情報量と形状情報量は、画像にもよるがほぼ同等であるといえる。この二つの情報量の和は、予測誤差情報量の 10 ~ 15 分の 1 程度となっている。

次に表 7.8 を見ると、全探索ブロックマッチング（BM 法）では、動ベクトル情報量は非常に小さいが、予測誤差情報量が非常に大きいため、全発生情報量が大きくなっている。方式 B では、領域数が多く動ベクトルと形状情報量が大きいにもかかわらず予測誤差が十分に小さくない。これ

表 7.7: 方式 A の発生情報量 [bpp] の内訳.

| 画像 | 連結 領域数 | 動ベクトル 情報量 | 形状 情報量 | 予測誤差 情報量 | 全発生 情報量 |
|--------------|-----------|--------------|-----------|-------------|------------|
| cheerleaders | 3357 | 0.17 | 0.19 | 4.47 | 4.83 |
| green leaves | 4524 | 0.23 | 0.24 | 5.94 | 6.41 |
| marching in | 3951 | 0.21 | 0.24 | 5.38 | 5.83 |
| mobile | 2467 | 0.12 | 0.20 | 5.12 | 5.44 |
| soccer | 3261 | 0.16 | 0.23 | 5.50 | 5.89 |
| flamingoes | 4417 | 0.23 | 0.24 | 4.58 | 5.04 |

表 7.8: 各方式の発生情報量 [bpp] の内訳 (cheerleaders).

| 方式 | 連結 領域数 | 動ベクトル 情報量 | 形状 情報量 | 予測誤差 情報量 | 全発生 情報量 |
|------|-----------|--------------|-----------|-------------|------------|
| 方式 A | 3357 | 0.17 | 0.19 | 4.47 | 4.83 |
| 方式 B | 9439 | 0.49 | 0.68 | 4.74 | 5.91 |
| 方式 C | 2928 | 0.15 | 0.19 | 4.65 | 5.00 |
| BM 法 | 660 | 0.02 | - | 5.32 | 5.34 |

は、クラスタリングの段階で情報量が考慮されないために、その後の領域統合の効果が上がっていないものといえる。また方式 C では、領域数が減少し、動ベクトル情報量が小さいにも係わらず、方式 A と比較して形状情報量は同等である。統合時に形状情報量が考慮されていない影響と考えられる。また、予測誤差も大きい。方式 B, C に対し方式 A の大きな特徴は、動ベクトル情報量・予測誤差情報量に対する形状情報量の割合が比較的小さい点である。B, C では統合時に形状情報量を勘案していないのに対し、A では初期の統合段階での提案法の利用により、これら 3 種の情報量を全体的に抑えることに成功していることが確認できる。

これらの性能比較は、全発生情報量によるものであり、いわば予測誤差を可逆符号化する場合に相当する。図 7.6 は、これを、通常の伝送に用いるように非可逆符号化するとどうなるかを示している。予測誤差画像の電力スペクトル密度をもとに、注水定理を適用した。ビットレート約 0.4bpp 以上の範囲で、方式 A が他方式に約 0.2dB 以上の差をつけて最も高い PSNR を実現することがわかる。約 0.35bpp 以下の範囲で方式 A, B が BM 法より低効率となるのは、予測誤差の伝送量によらず分割形状は固定のためである。また、図 7.7 の方式 A の分割結果に見られるように、細かい領域分割となっている（領域数は約 3000 個）。これも、全発生情報量を対象とした最適化であるため、形状・動ベクトル情報量の占める割合が比較的小さいからであるといえる。

伝送情報量に制限を設けた場合の領域分割動き推定の最適化法（SN 比の最大化）は、第 8 章で検討する。

画像による違いを示した表 7.9 を見ると、これら全画像で方式 A がもっとも効率の良いことが分かる。特に cheerleaders のように動きの混在の激しい画像で方式 A の効果が大きく、領域分割符号化の特長が良く表れている。BM 法との差も大きい。一方で、mobile や marching in, flamingoes のように動きのゆっくりした画像では、BM 法と方式 A の特性差は小さくなる。その理由としては、この種の画像では、矩形ブロック分割と任意形状分割との動き補償性能に大差がないこと

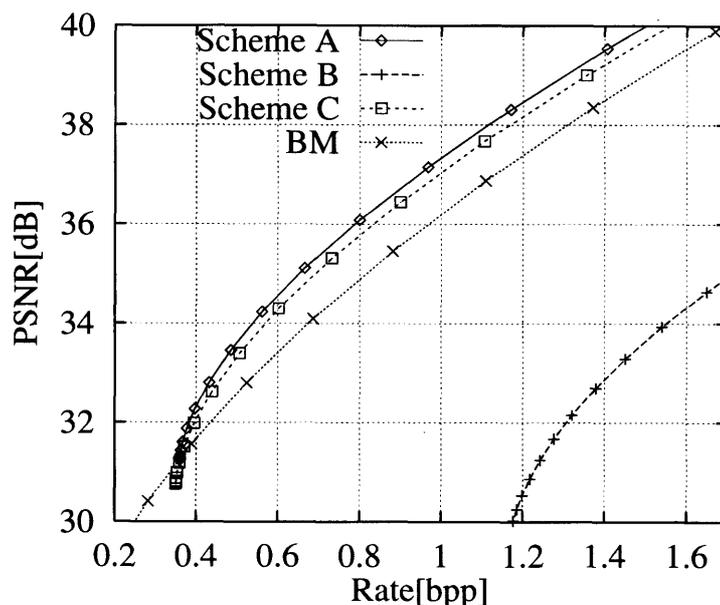


図 7.6: 動き補償予測誤差画像のレート歪み特性 (cheerleaders. 注水定理による. レートには動ベクトル情報量・形状情報量を含む.).

表 7.9: 画像毎の各方式の全発生情報量 [bpp] の比較.

| 画像 | 方式 A | 方式 B | 方式 C | BM 法 |
|--------------|------|------|------|------|
| cheerleaders | 4.8 | 5.9 | 5.0 | 5.3 |
| green leaves | 6.4 | 7.4 | 6.5 | 6.7 |
| marching in | 5.8 | 6.6 | 5.9 | 6.0 |
| mobile | 5.4 | 6.3 | 5.5 | 5.5 |
| soccer | 5.9 | 6.7 | 6.0 | 6.2 |
| flamingoes | 5.0 | 6.0 | 5.1 | 5.2 |

があげられる。全体としては、効率の高い方から順に方式 A, 方式 C, BM 法, 方式 B であることが分かる。

図 7.5 のとおり, シミュレーション方式 A, B, C の違いは反復処理の初期段階のみである。動的計画法を用いた提案方式は, シミュレーション計算量削減のため, 方式 A の反復初期段階に簡易化して組み込まれているが, 以上の実験結果より, その有効性が実証された。

なお, 本検討における動的計画法の計算においては, 性能とのトレードオフとなるような計算時間の短縮手段, すなわち状態数や選択枝数の削減は行なっていない。そのため計算時間は 1 フレーム当たり約 1 時間 (Sun SparcUltra1) と, 長くなっている。計算時間の短縮のためには, 並列処理 CPU の利用や, 性能をなるべく落さずに状態数や選択枝数の削減を行なう方法の検討が必要となる。

また, 提案法を用いた動き推定・領域分割方式を本格的に構成する場合には, 次のような方法をとることにより, 上記シミュレーション方式 A よりも性能を向上させることが可能と考えられる。すなわち, 方式 A のように提案法を反復処理の初期の大ブロック単位に適用するだけでなく, 更に階層化を行なう。具体的には, 大ブロック単位の領域分割の結果の各領域を新たな領域分割単

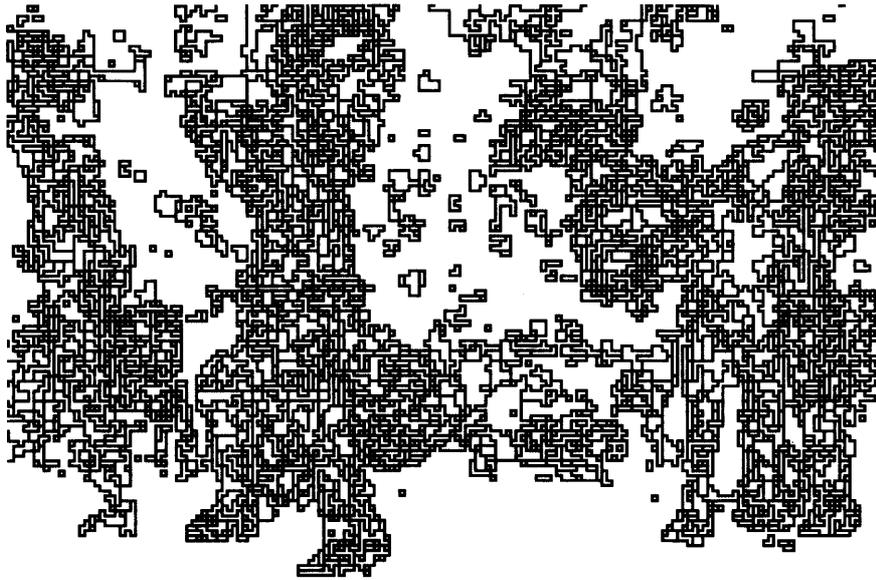


図 7.7: 分割結果の例 (方式 A, cheerleaders).

位とみなし, それらを適当な数毎にまとめて提案法を適用していく.

7.5 本章のまとめ

本章では, 動画像の領域分割動き補償の最適化を目的として, 発生情報量を最小にするべく可変形状の領域分割と動き推定を実行する方式を提案した. 本方式では, まず候補動ベクトル毎に, 取り得る各領域パターンにおける発生情報量 (予測誤差情報量, 形状情報量, 動ベクトル情報量の和) を算出する. 次に, これをもとに, 動ベクトルを段階, 領域パターンを状態とする動的計画法を適用することにより, 各動ベクトルの占めるべき領域を決定する.

本方式に用いる形状情報量の算出方法としては, 各領域パターン毎に独立して見積りが可能であることが必要となる. その一方法として, 四色定理を応用した階層的方法を考案し適用することにより, 大サイズの矩形が分割形状として優先される (見積もり情報量が小さくなる) ように考慮した. これにより, 仮に矩形分割が最適となるような画像があったとしても, ブロックマッチング動き補償に対する効率の低下が最小限に抑えられることが期待できる.

計算機シミュレーション実験により, 種々の一般動画像に対し, 本方式がブロックマッチング法や従来の領域分割動き推定法を上回る性能を持つことを示した.

今後の課題としては, 発生情報量に大きな影響を及ぼす形状符号化法の更なる検討や, 性能を保ちつつ本方式の計算量を削減する方法の検討などがあげられる. また, 発生情報量に制限を設けた場合に最大 SN 比を達成するような領域分割動き推定法の検討は, 次の第 8 章で行う.

本章の研究は, 画像の種類に性能が大きく依存するとされてきた領域分割符号化法を種々の一般動画像へ適用する方式の確立の一助となることが期待できる.

本章の参考文献

- [1] ITU-T Recommendation H.261: “Video codec for audiovisual services at p x 64 kbit/s” (1993).
- [2] ITU-T Recommendation H.262: “Information technology - Generic coding of moving pictures and associated audio information: Video” (2000).
- [3] M. Gilge, T. Engelhardt and R. Mehlan: “Coding of arbitrary shaped image segments based on a generalized orthogonal transform”, *Signal Processing: Image Communication*, 1, pp. 153–180 (1989).
- [4] 松田, 伊東, 宇都宮: “ポロノイ図を用いたカラー画像の領域分割符号化”, *信学ソ大*, No. D-118 (1995).
- [5] 小松, 斎藤: “複数の動きを推定可能なブロックマッチング法”, *テレビ誌*, **49**, 3 (1995).
- [6] 秋元, 小松, 斎藤: “Mid-level 画像符号化のための大域的動き解析法の一検討”, *信学ソ大*, No. D-113 (1995).
- [7] 高橋, 相澤, 羽鳥: “複数フレームの誤差空間に基づく動き推定法”, *信学ソ大*, No. D-112 (1995).
- [8] 小林, 伊藤, 宇都宮: “領域分割と warping 予測に基づいた動き補償方式”, *信学ソ大*, No. D-109 (1995).
- [9] L. Wu, J. Benois-Pineau, P. Delagnes and D. Barba: “Spatio-temporal segmentation of image sequences for object-oriented low bit-rate image coding”, *Signal Processing: Image Communication*, 8, pp. 513–543 (1996).
- [10] 井須, 関口, 浅井: “ブロックマッチングMCに基づくブロックベース領域分割法の評価”, *信学総大*, No. D-11-67 (1997).
- [11] ISO/IEC 14496-2: “Information technology – Coding of audio-visual objects – Part 2: Visual” (2001).
- [12] M. Eden and M. Kocher: “On the performance of a contour coding algorithm in the context of image coding Part I: Contour segment coding”, *Signal Processing*, **8**, 4, pp. 381–386 (1985).
- [13] C. Morgenstern and H. Shapiro: “Heuristics for rapidly four-coloring large planar graphs”, *Algorithmica*, 6, pp. 869–891 (1991).

第 8 章

符号化速度適応形可変形状領域分割方式

8.1 まえがき

第 7 章で述べたとおり，MPEG-2 などのブロックベースの動画像符号化では，領域の混在部で符号化効率が低下するという問題がある．その意味で，画像を性質の似た領域毎に符号化する領域分割符号化の潜在力は大きく，次世代動画像符号化方式の有力な候補といえる．この潜在力を活かすためには，領域分割の際に発生符号量を考慮することが重要となる．

従来より，領域分割の基本は，まず画像を細かく分割し，その個々の領域に対して隣接領域がある規範で逐次統合していくものである（分離統合法）[1, 2, 3, 4, 5]．しかし，これらの手法では，1 回 1 回の統合で考慮されるのは当該領域とその隣接領域のみであるため，全発生情報量の最小化という点では最適とはいえない．そのため，領域分割符号化の本来の性能が充分出ていなかったといえる．可変形状の符号化をサポートしている MPEG-4 においても，領域分割の方法自体は標準化の対象外となっている [6]．

そこで第 7 章では，動き補償予測符号化を前提とし，動的計画法の利用により，予測誤差，動ベクトル，形状の全発生情報量を最小化する領域分割動き推定方式を提案し，その有効性を示した．

本章では，この検討をさらに進め，フレーム内 / 動き補償予測いずれの符号化にも適用可能で，さらに符号化速度適応型の，動画像の可変形状最適領域分割方式を求めることを目的とする．

以下，8.2 で領域分割符号化の課題について述べた後，まず 8.3 で，フレーム内 / 動き補償予測，いずれの符号化にも適用可能な領域分割法を提案する．8.4 では，それに基づき符号化速度適応形の最適領域分割法を提案するとともに，それを用いた動画像符号化の枠組を提示する．8.5 で提案方式の有効性を示す．

8.2 領域分割符号化における課題

従来のブロックベースの符号化では，広く知られているように，ブロックに異なる領域が混在する場合に，符号化効率が低下するという問題があった．

これに対し，図 8.1 に示す領域分割に基づく符号化では，性質の一樣な領域毎に符号化されるため，このような問題を抑制することができる．

フレーム内符号化のフレームに領域分割符号化を適用する場合，原画像を性質の一樣な領域毎に領域分割し [1, 2]，それぞれを任意形状の直交変換 [7] などで符号化するのが一般的である．一方，動き補償予測符号化のフレームにおける領域分割符号化では，分割された領域毎に動き補償

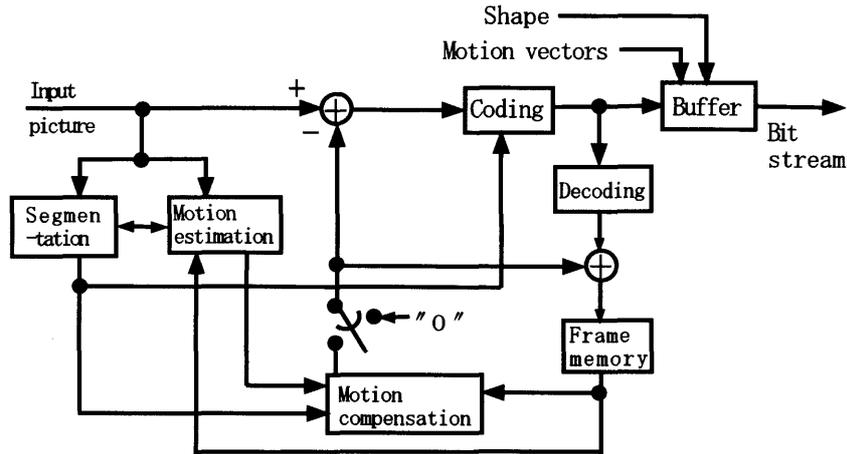


図 8.1: 領域分割符号化の構成 (decoding 部に入力される形状情報などは図から省略してある)

を行う [3, 4, 5] ことで、予測誤差の減少をはかる。この予測誤差は、フレーム内符号化と同様、分割領域毎に符号化することも、あるいは、従来のように方形ブロック毎に符号化することも可能である (予測誤差をブロック毎に符号化してもよい理由は、動き補償予測符号化の場合には、分割された各領域の予測誤差の性質が似通ってくるためである)。

いずれの場合にも、領域形状情報を新たに伝送する必要が生じるため、このトレードオフの制御法が重要となる。すなわち、フレーム内符号化 / 動き補償予測符号化それぞれについて、次のようなものが最適な領域分割といえる。

フレーム内符号化: テクスチャ情報量、形状情報量の合計を最小化

動き補償予測符号化: テクスチャ (予測誤差) 情報量、動ベクトル情報量、形状情報量の合計を最小化

従来の領域分割符号化法では、この点での検討が不十分であり、領域分割が符号化制御に充分組込まれていなかった。

そこで第 7 章で、動き補償予測符号化を前提とし、発生情報量を考慮しつつ、領域分割・動き推定を同時に行う方法を提案し、その有効性を示した。これは、動的計画法を利用して、全発生情報量を最小化するように動ベクトルと領域分割形状を決定するものである。動的計画法における「段階」を候補動ベクトルとし、これら動ベクトルに領域を割当てるという考えに基づき、最適領域分割・動き推定を実現した。

また、この方式に必要な領域毎のテクスチャ情報量の見積もりは、特定の符号化法を前提としないという条件のもと、なるべく妥当な見積りができる方法として、以下の方法を使用していた。すなわち、符号化対象領域のテクスチャ信号が、例えば、ラプラス分布で近似できるとし、その信号分散を σ^2 、領域内画素数を m 個とすると、その領域からのテクスチャ情報量 E_E を

$$E_E = \frac{m}{2} \log_2 2e^2 \sigma^2 \quad [\text{bits}] \quad (8.1)$$

と見積もっていた。

しかしながら、本方式を実際に符号化に適用するにあたっては、次のような問題が生じる。

- (1) フレーム内符号化の場合、動ベクトルがないため、上記の動的計画法 (「段階」を動ベクトルとする) が使用できず、領域分割が行えない。

(2) 符号化速度が小さくなると、形状情報の占める割合が大きくなり、ブロックベースの方式に比べ、符号化効率が低下する。すなわち、伝送可能情報量が充分大きい場合と、そうでない場合には、領域分割の形状自体も変える必要があるということが、経験的にも知られているが、式(8.1)ではそれが反映されず、厳密にはロスレスに近い高ビットレートで符号化する場合にのみ対応していた。つまり、非可逆(ロッキー)符号化の場合には、テクスチャ情報量の見積もりとして式(8.1)は不適切であり、符号化速度を勘案した見積もりが必要となる。

そこで、これらを解決するため、フレーム内/動き補償予測、いずれの符号化へも適用可能な発生情報量最小化領域分割法を8.3で提案し、さらに8.4で、この方法を使用した符号化速度適応型の最適領域分割方式を提案するとともに、それをを用いた動画像符号化の枠組を示す。

8.3 フレーム内/動き補償予測符号化へ適用可能な発生情報量最小化領域分割方式

8.3.1 前提

画像があらかじめ多数の小領域に分割されているとする。分割の例としては、クラスタリングや小矩形ブロックへの分割などがあげられる。このとき、最適領域分割とは、これらの小領域(最小分割単位)の最適な統合手法を求めることとなる。ここでの最適性とは、各領域からの発生情報量の画面全体での和を最小にすることとする。

第7章では、動き補償予測符号化を前提とすることで、小領域を各動ベクトルに割当てて最適配分問題とし、動的計画法を適用した。しかし、第7章における、動き推定を同時に行う領域分割手法、およびそれに適用した動的計画法は、動き推定が不要なフレーム内符号化の場合には、そのままでは適用できない。

そこで、次のように考える。すなわち、最小分割単位である小領域が、複数個統合された領域グループというものを考え、その領域グループの特殊な場合として、小領域をまったく含まない(すなわち、空の)仮想的な領域グループも含めて考える。最適化手法としては第7章と同様、動的計画法を利用することとし、その前提として、小領域のグループのとりうる領域パターンのそれぞれについて、発生情報量が見積もられ、テーブル化されているとする。例えば、簡単な例として、最小分割単位の小領域4個だけからなる画像を考える。この場合、どのグループも、図8.2に示すような14通りの領域パターンのいずれかをとることとなる。ここにパターン p_0 は、小

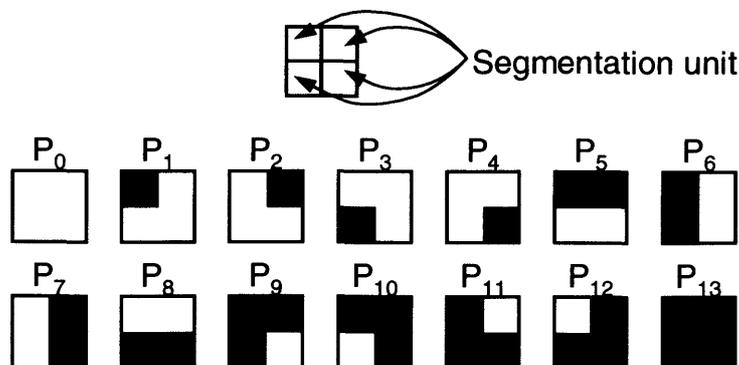


図 8.2: 分割単位 4 個からなる画像とその領域パターン (黒色部)

表 8.1: 発生情報量の見積り $E(k)$ の例

| | | | | | | | |
|--------|-------|-------|-------|----------|----------|----------|----------|
| k | p_0 | p_1 | p_2 | p_3 | p_4 | p_5 | p_6 |
| $E(k)$ | 0 | 5 | 4 | 3 | 2 | 9 | 9 |
| k | p_7 | p_8 | p_9 | p_{10} | p_{11} | p_{12} | p_{13} |
| $E(k)$ | 8 | 7 | 13 | 10 | 10 | 12 | 20 |

領域をまったく含まない仮想的な領域グループを意味している。

また、各領域パターン k 毎の発生情報量の見積り $E(k)$ として、表 8.1 のようなテーブルが得られていることを前提とする。各発生情報量の内訳は、前述の通り、フレーム内符号化の場合は、テキスト情報量と形状情報量となる。また動き補償予測符号化の場合は、予測誤差情報量、形状情報量、動ベクトル情報量となる。

ここで予測誤差情報量は、各領域パターンにそれぞれ最も適した（すなわち予測誤差情報量を最小にする）動き補償を行って得られるものとする。したがって、例えば、動ベクトル情報量として、1 領域当たり一定ビット数を想定すると、隣接する 2 領域（例：図 8.2 の p_1 と p_2 ）での最適動ベクトルが同一の場合は、2 領域を統合した領域（例：図 8.2 の p_5 ）の方が、2 個の別領域とするよりも、全発生情報量の点で有利になるといえる。

8.3.2 動的計画法の適用

画像全体が N 個の小領域（最小の領域分割単位）からなるとすると、統合後の領域グループ数は最大でも N 個（全く統合されなかった場合）である。そこで、「実際には領域を有しないグループ」も含めて、計 N 個の仮想領域グループへの統合を考える。例えば、10 個の最小領域分割単位があり、これが 4 領域へ統合される場合、 $10 - 4 = 6$ 個のグループは小領域をまったく含まない仮想的な領域グループであると解釈する。

動的計画法の適用に際しては、前記の発生情報量のテーブルを利用する。段階、状態、方策、コスト関数を次のように定義する。

段階 t : t 番目のグループの領域 R_t ($1 \leq t \leq N$)

状態 x_t : 領域 R_1, R_2, \dots, R_t の全体 $\cup_{i=1}^t R_i$ (累積領域)

方策 k : 領域 R_t の取り得る領域パターンの選択肢（ゼロ領域も取り得るとする）

コスト関数 $f_t(x_t)$: 累積領域 $x_t (= \cup_{i=1}^t R_i)$ からの発生情報量（の最小値）

以上より、最適化原理式は次の式 (8.2) となる。

$$\begin{cases} f_1(x_1) = E(x_1) \\ f_t(x_t) = \min_{\forall k \subseteq x_t} \{E(k) + f_{t-1}(x_t \ominus k)\} \\ \quad (2 \leq t \leq N-1) \\ f_N(S) = \min_{\forall k \subseteq S} \{E(k) + f_{N-1}(S \ominus k)\} \end{cases} \quad (8.2)$$

ここで S は画像全体を意味し、 $x \ominus k$ は領域 x から領域 k を除いた残りの領域を意味する。

式 (8.2) を解くことにより、各グループに属すべき小領域（領域分割単位）が決定される。具体的には、式 (8.2) を用いて、最終段階の画像全体から段階を遡って、漸的に領域パターンを求めていくことにより、領域の統合の仕方を決定する。このようにして、全発生情報量を最小にする意味で最適な領域統合が実現できる。

表 8.2: 最適化原理式の第 1 処理過程 ($f_1(x_1)$ の表)

| x_1 | $f_1(x_1)$ |
|----------|------------|
| p_1 | 5 |
| p_5 | 9 |
| p_6 | 9 |
| p_9 | 13 |
| p_{10} | 10 |
| p_{11} | 10 |
| p_{13} | 20 |

表 8.3: 最適化原理式の第 2 処理過程 (* がついたものが $f_2(x_2)$ の値)

| x_2 | R_2 の領域パターンの選択肢 k | | | | | | |
|----------|-----------------------|-------|-------|-------|-------|-------|----------|
| | p_0 | p_2 | p_3 | p_4 | p_7 | p_8 | p_{12} |
| p_5 | - | 9* | - | - | - | - | - |
| p_6 | - | - | 8* | - | - | - | - |
| p_9 | - | 13 | 12* | - | - | - | - |
| p_{10} | - | - | - | 11* | 13 | - | - |
| p_{11} | - | - | - | 11* | - | 12 | - |
| p_{13} | 20 | 14 | 13* | 15 | 17 | 16 | 17 |

8.3.3 処理過程

図 8.2, 表 8.1 の例を用いて, 具体的な処理過程を説明する. 次のようにして一般性を失わない.

- (1) 第 1 グループの領域 R_1 には少なくとも最左上の領域を含める.
- (2) 累積領域 x_t は連結領域とする.
- (3) 領域を有しないグループは, t の後の方にまとめる. すなわち, $R_t = p_0$ (ゼロ領域) であれば, $i > t$ なるすべての i についても $R_i = p_0$ であるとする. (なお, ここで R は領域であり, p は領域パターンであるが, ここでは便宜的に “=” でつなぐことにより, R がパターン p をとることを意味することとする.)

条件 (1) より, 各段階での累積領域 (状態) は, 最左上のブロックを含むパターンとなる. このことと条件 (2) より, 累積領域パターンとしてとり得るのは $p_1, p_5, p_6, p_9, p_{10}, p_{11}, p_{13}$ となる. また, 選択肢としてとりうるのは, 最左上のブロックを含まない連結領域パターン $p_0, p_2, p_3, p_4, p_7, p_8, p_{12}$ となる. そして, 条件 (3) を考慮すると, 各段階において, 累積領域が p_{13} (画像全体) である場合以外には, 選択肢として p_0 (ゼロ領域) はあり得ない.

以上に従い表 8.1 のテーブルをもとに式 (8.2) の漸化式の処理を行う際の過程は, 順に表 8.2, 表 8.3, 表 8.4, 表 8.5 のようになる. 表中, 選択肢としてあり得ない, もしくは上記条件に従い外すことのできる領域パターンは, 「-」で表現している. また, 選択肢が 1 つもない状態 x_t は, 表から外してある.

表 8.2 の作成では, R_1 のとり得る領域がそのまま累積領域 x_1 となり, 表 8.1 が用いられる. 表 8.3 では, R_1 と R_2 で占められ得る累積領域パターン x_2 のそれぞれについて, R_2 の領域パターン選択肢に関する発生情報量の最小値を求める. このとき, 表 8.1, 表 8.2 が利用される. 表 8.4 で

表 8.4: 最適化原理式の第 3 処理過程 (* がついたものが $f_3(x_3)$ の値)

| x_3 | R_3 の領域パターンを選択肢 k | | | | | | |
|----------|-----------------------|-------|-------|-------|-------|-------|----------|
| | p_0 | p_2 | p_3 | p_4 | p_7 | p_8 | p_{12} |
| p_9 | - | 12* | 12* | - | - | - | - |
| p_{10} | - | - | - | 11* | - | - | - |
| p_{11} | - | - | - | 10* | - | - | - |
| p_{13} | 13* | 15 | 14 | 14 | 16 | 16 | - |

表 8.5: 最適化原理式の第 4 処理過程 (* がついたものが $f_4(x_4)$ の値)

| x_4 | R_4 の領域パターンを選択肢 k | | | | | | |
|----------|-----------------------|-------|-------|-------|-------|-------|----------|
| | p_0 | p_2 | p_3 | p_4 | p_7 | p_8 | p_{12} |
| p_{13} | 13* | 14 | 14 | 14 | - | - | - |

は、同様に、 R_1 と R_2 , R_3 で占められ得る累積領域パターン x_3 のそれぞれについて、 R_3 の領域パターン選択肢に関する発生情報量の最小値を表 8.1, 表 8.3 により求める. 表 8.5 では、 R_1 , R_2 , R_3 , R_4 の占める累積領域は全画面に他ならないので、 x_4 は p_{13} のみになる. 各選択肢について、表 8.1 と表 8.4 から、発生情報量が計算され、それを最小値 13 にする選択肢 p_0 が R_4 のとるべき領域と決定される. このとき

$$x_3 = x_4 \ominus R_4 = p_{13} \ominus p_0 = p_{13}$$

となり、表 8.4 より R_3 も p_0 となる. 同様に

$$x_2 = x_3 \ominus R_3 = p_{13} \ominus p_0 = p_{13}$$

より、表 8.3 を参照し $R_2 = p_3$ と決定される. そして、

$$x_1 = x_2 \ominus R_2 = p_{13} \ominus p_3 = p_{10}$$

なので、表 8.2 より $R_1 = p_{10}$ となる.

以上の最適領域統合の結果、画像全体が p_3 と p_{10} の 2 領域に分割されることになる.

本例では、最大の領域グループ数は最小分割単位の個数と同じ 4 である. これを例えば、2 に設定したいときは、表 8.3 で、状態 x_2 として $p_5 \sim p_{11}$ を除き p_{13} (全画面) のみについて計算し、表 8.4 と表 8.5 の処理を省けばよい. さらに、必ず 2 グループに統合したい場合は、選択肢からゼロ領域パターン (p_0) を外せばよい. このように、本手法を用いれば、領域数に制限を設けた最適統合も可能である.

8.3.4 提案手法の特長と実画像への適用法

以上説明したように、提案手法は、可能性のある全統合パターンの比較を動的計画法の利用により実現しており、系統的な最適領域統合手法となっている. 従来の分離統合手法では、単に隣接領域どうしの統合を閾値を用いて繰り返していくことが基本となるため、全発生情報量の最小化が達成される保証はない. また、閾値などのパラメータ設定にも性能が左右される.

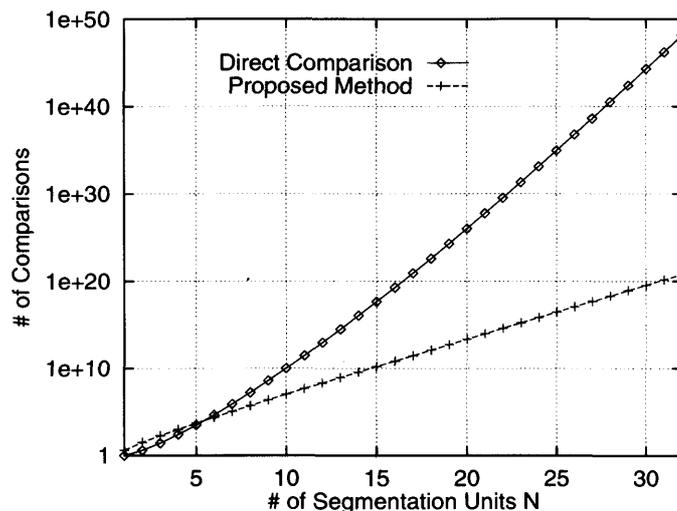


図 8.3: 全統合パターンの直接比較回数と提案手法による比較回数

全統合パターンにおける発生情報量を，各々直接比較する場合と提案手法との計算量を比較すると，次のようになる．全統合パターンの直接比較回数の目安は， N 個の最小分割単位が N 種類のグループに属する場合の数として， N^N 回となる．一方，提案手法の場合，領域パターン選択枝数，累積領域パターン数ともに最大で 2^N 個であり，段階数が N であるので，比較回数の目安は $N \cdot 2^N \cdot 2^N$ 回となる．これらを N について求めたものを図 8.3 に示す．提案手法は， N の増加に伴い急激に増大する直接比較の計算量を， $N = 15$ で約 $1/10^8$ ， $N = 30$ で約 $1/10^{24}$ と，大幅に削減することに成功している．

ただし，提案手法でも，あまり N を大きくすると計算量が膨大になって，全統合パターンの直接比較と同様，並列処理を行わない限り実現が事実上不可能となる．したがって，実画像への適用に際しては，(1) 最小分割単位の領域を十数個程度毎にまとめ，(2) 各々に対し本手法を適用し（例えば $N = 16$ の場合，Sun Ultra-1 で約 15 秒の処理時間となる），(3) 必要であればその結果の領域を新たな分割単位として (1)(2) の処理を繰り返す，などの方法が適当と考えられる．この階層的手法を用いた方式を，8.4.2 で説明する．

8.4 動画像符号化のための符号化速度適応型最適領域分割

8.4.1 理想の方式

8.2 で述べた課題の 1 つ，すなわち符号化速度に適応した領域分割を実現するには，符号化速度を勘案したテクスチャ情報量の見積もりを行う必要がある．

ある領域の符号化雑音電力を n とすると，発生テクスチャ情報量の見積りは，式 (8.1) に代え，次のようにおくことができる [8]．

$$E_E = \frac{m}{2} \log_2 \frac{\sigma^2}{n} \quad [\text{bits}] \quad (8.3)$$

ここに $n \leq \sigma^2$ である (m は領域内画素数， σ^2 は信号分散)．この式では， $n = \sigma^2$ (低ビットレートにおいて符号化ノイズが大，DCT でいえば係数の切捨て状態に相当) となると，どのような分割形状でもテクスチャ情報量は 0 となり，分割形状による差がなくなるので，なるべく形状情報の発生しない分割のほうがよいことになる．一方式 (8.1) は， n を含んでいない，すな

わち符号化雑音電力が充分小さな一定値になるまで符号量が与えられるということが仮定されている。つまり、符号化歪みの大きさ、ひいては符号化速度を勘案することができない。なお、式(8.1)にあった信号分布を表す定数 $2e^2$ は、テクスチャ情報量に対する固定的なバイアスとして働くのみである（すなわちどのような分割であっても分割結果に影響を与えない）ため、省略する。

最適化の目標は、「符号化速度の制約のもとでの符号化雑音電力の最小化」といえるので、式(8.3)を使用し、次のような動的計画法で表現できる。ここに領域グループとは、前述の通り、最小領域分割単位（例えば、サイズ 4×2 のブロック）の集合であり、空集合（領域をまったく含まない仮想領域グループ）もありえる（グループの総数は、最小領域分割単位の数 N となる）。また、「符号化速度の制約」とは、「画面全体での発生情報量の総和（これを E_S とおく）に対する制約」である。

段階 t : t 番目のグループの領域 R_t ($1 \leq t \leq N$)

状態 x_t : 領域 R_1, R_2, \dots, R_t の全体 $\cup_{i=1}^t R_i$ (累積領域) と、その累積領域からの全発生情報量の組合せ。すなわち、領域 R_i からの発生情報量を E_i と表すと、 $x_t = (r_t, e_t)$ となる。ここに $r_t = \cup_{i=1}^t R_i, e_t = \sum_{i=1}^t E_i$ である。 $t = N$ のときは $r_N = S$ (画面全体), $e_N = E_S$ となる。

方策 k : 領域 R_t の取り得る領域パターンと発生情報量の組合せの選択肢（ゼロ領域も取り得るとする）。

コスト関数 $f_t(x_t)$: 状態 x_t での符号化雑音電力（の最小値）

ここに、各領域パターン R_i の発生情報量 E_i の内訳は、テクスチャ情報量、形状情報量、動ベクトル情報量となる。このうちテクスチャ情報量は、前述のように、 R_i に最適な動き推定を行った上で、式(8.3)に基づき符号化雑音電力 n_i と関連付けられるものとする。

本方式における 8.3 の方式からの大きな相違点は、発生情報量が方策および状態に加わることである。発生情報量は、本来連続値をとるので、精度良く計算しようとする膨大な計算量増加となり、この理想方式は事実上実行不可能となる。

8.4.2 実現のための簡略化

そこで、計算量を実現に妥当なものとするため、以下の簡略化を行う。

最適化単位の設定

まず、領域分割を、画面を分割したマクロブロック内で閉じて実行することとする。このときのオーバーヘッドを考察する。

最適化単位のマクロブロックサイズを $N \times M$ とする。すべてのマクロブロック内で完全に領域統合された場合が、動ベクトルの発生情報量の最小値となる（形状情報量も、通常の形状符号化であれば、最小値をとると考えられる）。このとき、動ベクトル 1 個当たりの情報量を a [bit]、領域形状情報量を b [b/MB] とすると、換算して $(a + b)/MN$ [b/pel] が、オーバーヘッドの最大値となる。すなわち、仮にマクロブロック内で閉じずに画面全体で領域分割を行った場合には、もっと統合が進むような画像であっても、最低これだけの動ベクトル情報量と領域形状情報量が発生することとなる。例として $a = 5$ とし、領域形状の計算方法として、例えば、第 7 章にある方法

を採用すると $b = 3$ となるので、オーバーヘッド情報量は $8/MN$ [b/pel] となる。この値が、対象とする符号化速度に比べ充分小さくなるように、マクロブロックサイズ M, N を決めればよい。

例えば、 $M = N = 16$ とすると、上の値は約 0.03 [b/pel] となり、通常の符号化ビットレートとの値と比較して充分小さくなる。また、これを式 (8.3) などにより PSNR に換算すると、約 0.2 dB に相当する。これは、完全に無視できるレベルとまではいえないものの、本検討対象としては充分小さいと考え、以後の実験では、このブロックサイズを使用することとする。

MB 毎の量子化器制御

上記に加え、次のような簡略化を行う。

- 量子化器制御の単位を、そのマクロブロックごととする。すなわち、画面全体での発生情報量を所定値に合わせるため、マクロブロック単位で量子化器を制御する。

周知のように、信号分散とその分布および量子化ステップサイズが決まれば、量子化雑音は求まる（動き補償予測誤差信号はラプラス分布になることが知られている）。したがって、前節で述べたようにマクロブロック (MB) 単位で領域分割を最適化するとし、量子化器は MB 毎に固定とすると、動的計画法のコスト関数は、8.3の方式と同様、発生情報量（の最小値）とすることができ、方策および状態には、発生情報量を含む必要はなくなる。すなわち 8.3 と同じく、次のようになる。

段階 t : t 番目のグループの領域 R_t ($1 \leq t \leq N$)

状態 x_t : 領域 R_1, R_2, \dots, R_t の全体 $\cup_{i=1}^t R_i$ (累積領域)

方策 k : 領域 R_t の取り得る領域パターンの選択肢。

コスト関数 $f_t(x_t)$: 状態 x_t での発生情報量（の最小値）

各領域パターンからの発生情報量のうちの形状情報量の見積り法としては、第7章のものを使用する。またテクスチャ情報量は、量子化ステップサイズ Δ と雑音電力 n の変換テーブルと式 (8.3) に基づき求まる。この変換テーブルは、例えば、モンテカルロ法に基づき次のようにして得ればよい。分散 σ^2 でラプラス分布をなす信号をランダムに発生させ、 Δ で一様量子化し、 σ^2/Δ^2 と σ^2/n の関係をテーブル化する（これは、 σ^2, Δ に対する n の値を算出していることにも相当する）。

なお、量子化器制御法としては、MPEG-2 TM5 などの既存の方式が使用可能である。例えば、マクロブロック内での領域分割なしの状態で仮符号化してみるなどの方法が考えられる。

動的計画法の階層的適用

さらに、動的計画法を MB 内で階層的に適用することとする。その理由は、前述の通り、領域分割最小単位の数をある程度に制限しないと、動的計画法の計算量が莫大になるためである。

例えば、領域分割の最小単位をサイズ 4×2 の小ブロックとすると、サイズ 16×16 の MB 内には 32 個の最小単位が存在する。これを 1 回で動的計画法に適用すると、計算量が多いため、第 1 段階では、8 個の最小ブロックを含むサイズ 8×8 のブロック 4 個について、それぞれ動的計画法を適用する。第 2 段階以降では、前段階の統合による結果の領域を順次最大 8 領域ずつまとめ、動的計画法を繰り返し適用していく。そして、統合が起こらなくなった段階で終了とする。なお、予備実験の結果、最大 8 領域ずつまとめる際の順序が結果に与える影響は少ないことがわかった。

8.4.3 符号化の枠組

以上の方法を符号化と組合せると、次のような符号化速度適応型の領域分割動画像符号化の枠組が得られる。

- (1) マクロブロック毎に、次の処理を行う。
 - (a) 量子化器の決定. MPEG-2 TM5 などの方法が使用可能である。
 - (b) マクロブロックの符号化モードの決定および階層的動的計画法による領域分割（および動き推定）ここに動き推定は、表 8.1 のような発生情報量の見積もり表の作成のために各領域パターン毎に最適な動き補償を行う際に、マッチングにより実行する。
 - (c) 動き補償予測誤差画像の作成
 - (d) 直交変換・量子化
- (2) 符号化制御. MPEG-2 TM5 などの方法が使用可能である。これにより、画面毎の発生情報量を決定する。
特に、次のような点に注意する。

さまざまな符号化モードの実現法

次の 8.4.3, 8.4.3 で、フレーム内 / フレーム間の適応符号化や、MPEG-2/4 にあるようなさまざまな動き補償モードの組み込み方について述べる。アフィン動き補償なども、同様に組み込むことができる。なお、8.5 の計算機シミュレーション実験では、基本性能向上の要因を符号化モード選択法から切り分けて明らかとするため、あえて符号化モードは固定としている。

領域毎に適応化する場合 8.3.1 で述べたように、各領域パターン毎の発生情報量見積りテーブルを作成する際には、それぞれについて最適な動き補償のもとでの発生情報量が計算される。この動き補償の候補として、フレーム間予測なしの場合や、種々の動き補償法を適用した場合を加える。

マクロブロック毎に適応化する場合 上記のように、最適化の単位を MB 毎とすると、これに合わせ、符号化モードの切替単位も MB 毎にすることが考えられる。この場合、それぞれの符号化モードについて、その MB での最適化領域分割を行い、最も発生情報量を少なくするモードを選択すればよい。または、簡略法として、最適化の前に、ある規範により符号化モードを決定した後、そのモードのみで最適化を行う方法も考えられる。

直交変換符号化

テクスチャの符号化法としては、可変形状領域分割には本来、任意形状直交変換が適合しているといえる。しかしながら、現状の任意形状直交変換技術には、次のような問題点がある。

- (1) Gilge の任意形状直交変換法 [7] では、高次の基底の算出精度が問題となるため、高ビットレートでの適用が難しい（文献 [7] では、CIF 画像に対し約 64kb/s での符号化を行っている）。
- (2) MPEG-4 における境界ブロックのパディング法 [6] や SA-DCT [9] は、画像中のオブジェクト毎の符号化を前提としており、領域が比較的大きいことを想定している。本検討では、領域分割はあくまで符号化効率の追及であるため、後の図 8.5 に示すように、領域のサイズが小さく、また数も多くなりうる。したがって、このパディング法や SA-DCT では効率が低くなり不適切となる。

そこで、8.5の計算機シミュレーション実験では、直交変換法として、従来のブロック DCT を採用することとする。これに伴い、動き補償を行わないピクチャ (I-picture) では、領域分割も行わないことになる (ブロック DCT を使うことにするとマクロブロック内での領域分割を行う意味がなくなるため)。

提案方式では、最適領域分割の際に、発生情報量の推定を可変形状領域毎に行っている。このため、ブロック DCT の採用により、推定符号量と実符号量に mismatches を生じるといえる。したがって将来、任意形状用の優れた直交変換符号化法あるいはそれ以外の優れた符号化法が開発された場合には、それを採用することが望ましい。この場合には、I-picture に対しても領域分割符号化が適用できることとなる。

8.5 計算機シミュレーション実験

計算機を用いた符号化シミュレーションにより、提案する可変形状領域分割方式に基づく符号化と、従来のブロックベースの符号化の性能比較を行う。

従来方式は通常の MC-DCT であり、動き補償単位の MB サイズは 16x16、DCT はサイズ 8x8 のブロック単位に適用する。

提案方式でも、領域分割の最適化の単位となるマクロブロックサイズは 16x16 とする。また、最小領域分割単位の小ブロックサイズは 4x2 とし、階層的動的計画法の階層毎の領域単位数は計算時間を考慮して 8 とする。また DCT に関しては、本実験では、8.4.3 で述べたようにブロック DCT とし、そのサイズは従来方式と同じ 8x8 とする。

動ベクトル探索範囲は水平 -16 ~ +15 画素、垂直 -8 ~ +7 ライン、整数画素精度である。符号化モードと符号化制御に関しては、従来方式と提案方式とで共通化可能である。そこで、基本性能の比較を行いやすいように、ともに単純な方式を用いることとする。すなわち、ピクチャ構造は MPEG-2 におけるフィールドピクチャとし、単純フレーム間予測 (同一パリティ、すなわち 2 フィールド前からのフィールド予測)、B ピクチャなし、P ピクチャでは、画面内のすべてのマクロブロックを動き補償予測符号化する。また、量子化は固定ステップサイズで行い、これを (4,8,16,32,64) と変化させることで発生情報量と PSNR の関係を測定する。

予測誤差の発生情報量は、量子化後の DCT 係数のジグザグスキャンによるラン / レベル 2 次元エントロピーとして算出した。また形状情報量の測定は第 7 章の方法をとった。

テスト動画像としては、表 8.6 に示すものを使用する (各 30 フレーム)。これらは、BT.601 フォーマット [10] でインタレース構造であるため、動き補償の基本性能の比較を行いやすくするため、前述のようにシミュレーションにおける動き補償を単純フレーム間予測とした。

また、本シミュレーション実験においては、前述の通り、I ピクチャでは、領域分割を行っていない。そのため、I ピクチャでの符号化性能は、提案方式と比較方式とで、全く同じになる。

表 8.6: テスト動画像

| 画像 | 特徴 |
|--------------|-------------|
| cheerleaders | 異なる動きの混在 |
| mobile | 精細、一部に動きの混在 |
| marching | 全体的に精細 |
| flamingoes | ゆっくりした動き |
| soccer | 速い動き |

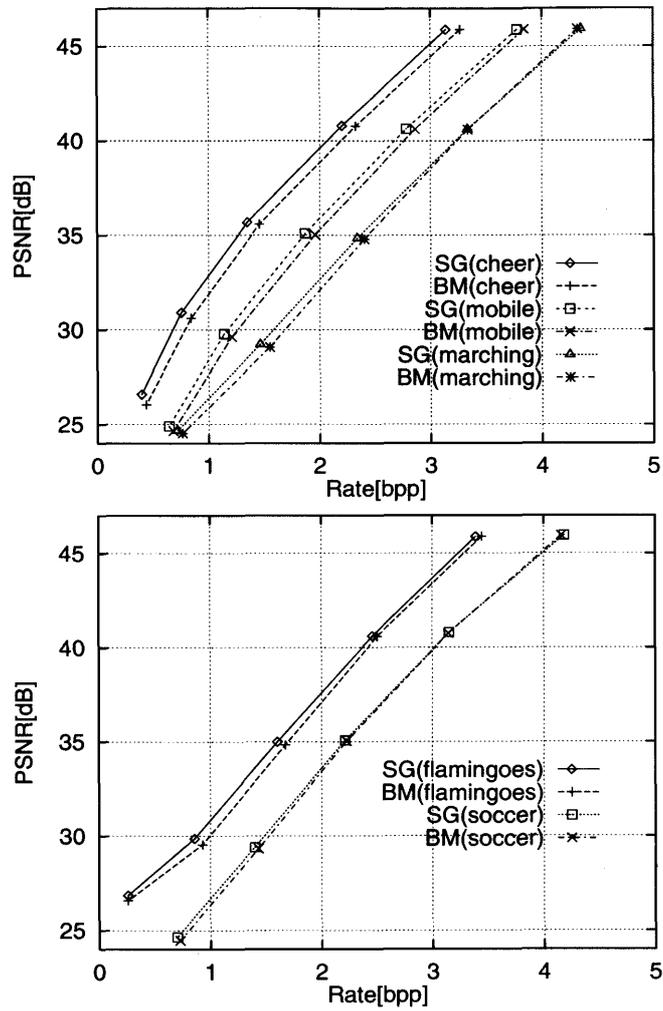


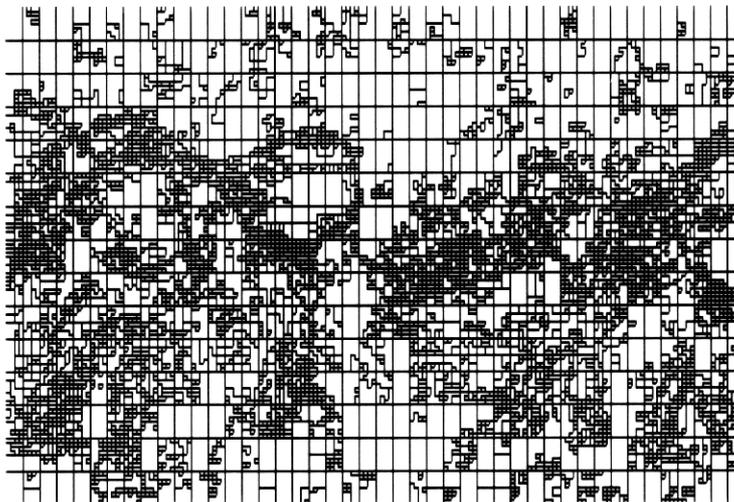
図 8.4: レート PSNR 特性の比較 (SG: 提案方式, BM: 比較方式)

表 8.7: 比較方式に対する提案方式の PSNR の増分 [dB]

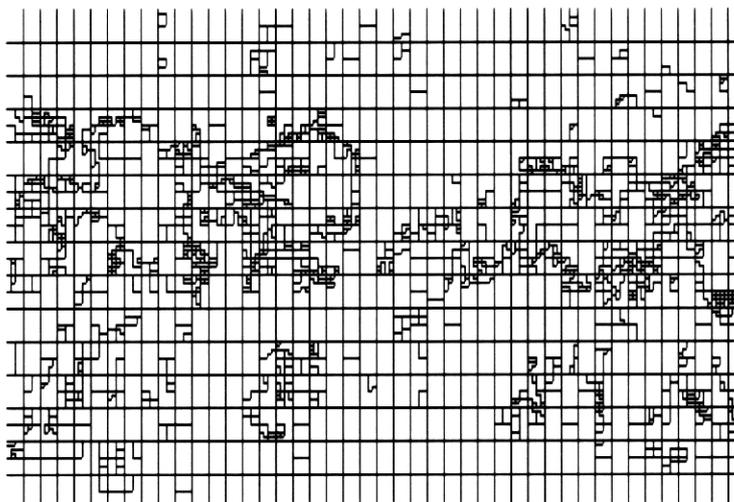
| 画像 | 1bpp | 2bpp | 3bpp |
|--------------|------|------|------|
| cheerleaders | +1.0 | +0.8 | +0.7 |
| mobile | +0.8 | +0.7 | +0.5 |



(a) 原画



(b) ステップサイズ = 4



(c) ステップサイズ = 64

図 8.5: 領域分割図 (cheerleaders)

表 8.8: 発生情報量の内訳

(a) ステップサイズ = 4

| 画像 | 全発生 情報量 [bpp] | 予測誤 差情報 量 [%] | 形状 情報量 [%] | 動ベク トル情報 量 [%] | 領域 数 [個] |
|--------------|---------------------|---------------------|------------------|----------------------|----------------|
| cheerleaders | 3.1 | 90.6 | 4.3 | 5.1 | 5367 |
| mobile | 3.8 | 94.4 | 2.8 | 2.9 | 3639 |
| marching | 4.4 | 90.3 | 4.1 | 5.6 | 8263 |
| flamingoes | 3.4 | 82.1 | 7.1 | 10.8 | 12329 |
| soccer | 4.2 | 92.7 | 3.3 | 4.0 | 5655 |

(b) ステップサイズ = 64

| 画像 | 全発生 情報量 [bpp] | 予測誤 差情報 量 [%] | 形状 情報量 [%] | 動ベク トル情報 量 [%] | 領域 数 [個] |
|--------------|---------------------|---------------------|------------------|----------------------|----------------|
| cheerleaders | 0.40 | 75.3 | 12.5 | 12.2 | 1628 |
| mobile | 0.64 | 89.1 | 5.3 | 5.6 | 1219 |
| marching | 0.72 | 91.5 | 3.9 | 4.6 | 1115 |
| flamingoes | 0.26 | 78.0 | 10.3 | 11.8 | 1030 |
| soccer | 0.70 | 92.0 | 3.7 | 4.3 | 1016 |

そこで、以後の図表における発生情報量（レート）やPSNRの数値としては、Iピクチャの結果を除外したPピクチャのみの結果を示す。

まず、図 8.4 と表 8.7 に、提案方式と比較方式のレートPSNR特性を示す。ブロックマッチング法を用いた比較方式に対し、cheerleaders で約 0.7dB 以上、mobile で約 0.5dB 以上提案方式が上回っており、その他の画像でも提案方式の方が上回っているか、または同等である。提案方式は、混在する動きが多く、領域分割符号化が有効な画像では、大きな効果があり、そうでない画像でも、従来方式に劣らずに符号化できることがわかる。

次に、高レート（ステップサイズ 4）と低レート（ステップサイズ 64）での cheerleaders の領域分割図を図 8.5 に示す。高レートでは細かく分割され、低レートではより統合が進んでいることが確認できる。また、一見動きの様な領域（例えば、画像中の地面の部分など）でも、領域分割がされていることがわかる。この傾向は高レートのものほど顕著である。これは、見た目の動きよりも、情報量の最小化（予測誤差の減少）が優先された結果であるといえる。すなわち、このような部分では、動ベクトルの微妙な違いに対して動き補償の予測誤差の変化が大きいため、一見動きが一様であっても領域分割がなされるといえる。

表 8.8 には、ステップサイズが 4 と 64 の各場合における画像毎の発生情報量の内訳を示す。これを見ると、領域数や全発生情報量は、ステップサイズや画像の種類によって大きく変化している。これに対し、形状情報量と動ベクトル情報量の全発生情報量に占める割合は、数% から十数% で推移し、比較的安定している。これらのことも、提案方式によって、符号化速度に適応した領域分割、および各情報量の制御が行われていることを示しているといえる。

以上により提案方式の有効性が確認できた。

8.6 本章のまとめ

本章では、次世代の高能率な領域分割画像符号化の実現のため、フレーム内 / 動き補償予測いずれの符号化にも適用可能な、符号化速度適応形の最適領域分割方式を提案し、それをを用いた可変形状の領域分割動画像符号化の枠組を提示した。

提案領域分割法においては、最適化手法として動的計画法を使用している。第7章の検討では、動き補償予測符号化を前提としていたのに対し、本方式ではフレーム内 / 動き補償予測いずれの符号化にも適用可能とした。また、領域分割の際、量子化ステップサイズ適応形の発生情報量見積りを行うことにより、符号化速度に応じた最適領域分割を実現する。さらに、動的計画法を階層的に適用することにより、計算量を削減し、実行可能としている。

その結果、提案領域分割法に基づく符号化では、一般動画像に対し、低レートから高レートに渡って、従来のブロックベースの符号化を上回る性能が得られる。

さらに、この符号化の枠組には、MPEG-2/4にあるようなさまざまな動き補償モードを組込むことができるなど、さらなる機能拡張が可能である。

第5章で述べたような高信頼のための伝送の冗長化は、2重化のみならず3重化が要求されることもあり、そのためには伝送容量の観点から画質を保ったまま更なる高圧縮を実現する必要がある。本提案方式は、そのための基礎とすることができる。

本検討により、次世代動画像符号化方式の確立の一助としたい。

本章の参考文献

- [1] P. Willemin, T.R.Reed and M.Kunt: “Image sequence coding by split and merge”, IEEE Trans., **COM-39**, 12, pp. 1845–1855 (1991).
- [2] 松田, 伊東, 宇都宮: “ポロノイ図を用いたカラー画像の領域分割符号化”, 信学ソ大, No. D-118 (1995).
- [3] 井須, 関口, 浅井: “ブロックマッチングMCに基づくブロックベース領域分割法の評価”, 信学総大, No. D-11-67 (1997).
- [4] 小林, 伊藤, 宇都宮: “領域分割と warping 予測に基づいた動き補償方式”, 信学ソ大, No. D-109 (1995).
- [5] L. Wu, J. Benois-Pineau, P. Delagnes and D. Barba: “Spatio-temporal segmentation of image sequences for object-oriented low bit-rate image coding”, Signal Processing: Image Communication, 8, pp. 513–543 (1996).
- [6] ISO/IEC 14496-2: “Information technology – Coding of audio-visual objects – Part 2: Visual” (2001).
- [7] M. Gilge, T. Engelhardt and R. Mehlan: “Coding of arbitrary shaped image segments based on a generalized orthogonal transform”, Signal Processing: Image Communication, 1, pp. 153–180 (1989).
- [8] 似鳥: “線形変換符号化と予測符号化 — 離散的パラメータを持つ連続情報源のための二つのデジタル符号化法”, 信学論, **53-A**, 2, pp. 97–104 (1970).
- [9] T. Sikora and B. Makai: “Shape-adaptive DCT for generic coding of video”, IEEE Trans., **CSVT-5**, 1, pp. 59–62 (1995).
- [10] ITU-R Recommendation BT.601-5: “Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios” (1995).

第 9 章

結論

本論文では、テレビ中継の高画質化や高信頼化のために筆者らが考案した、テレビ方式変換の改善や伝送の高エネルギー無瞬断化、画質遠隔監視や次世代領域分割符号化の方式について述べた。

まず、テレビ方式変換の画質改善については、画像の動き推定処理において、次の要素技術を新たに組み合わせて利用することとした。(1)DOG フィルタによる動き推定前処理(2)高精度反復勾配法(3)動ベクトル後処理絶対値順序フィルタ。これらにより、特に次のような改善が得られた。(1)カメラフラッシュなど明るさの急激な変化の際の動き推定の破綻の防止(2)格子模様など規則的パターンの絵柄の動き推定における、誤ベクトル発生的大幅な軽減(3)異なる動きの混在する部分において、より小さい動きを優先的に捉えることによる、変換画質の見た目の改善。この結果として、特に従来問題となっていた、スポーツ中継においてカメラが追尾する小飛球などまで、高画質で変換できるようになった。すなわち、放送事業者の要求する高度な画質要求を満足できるようになった。本装置は、シドニーオリンピックをはじめとする各種イベント伝送の他、日々の外国ニュースの伝送などで使用され、好評を博している。今後の課題としては、更なる動き推定の改善、特に上記の規則的模様に対する対策や動き混在部対策を更に強化することである。方式変換(フレーム内挿)とは、もともと絵の無いところに絵を作ることであり、そこには本質的な困難さが常に付きまとう。最初の動き補正方式変換が世に出てから約20年、ようやく、実用レベルに達してきたという言い方も出来る。今後も、更なる改良を目指して研究開発が続くであろう。また、理想の方式変換が安価に可能となって初めて、テレビ方式、特にフレーム周波数の世界統一化の可能性が出てくると言えるのではないだろうか。これは皮肉なことではあるが、これまでの番組資産を問題なく活用できるという条件が無ければ、フレーム周波数の統一はありえないと考えられるからである。

次に、完全冗長2重化テレビ回線の高エネルギー化と無瞬断化については、デジタル圧縮2重化系の次の特徴を利用することにした。(1)2系統の受信復号画像では、重畳する符号化ノイズが通常異なっている。(2)アナログ信号と異なり、2系統の画像を画素単位ですれなく正確に比較可能である。そこで、次の方法を考案した。(1)2系統の受信復号画像を平均化し符号化ノイズを平滑化することによる高エネルギー化の実現(2)(a)2系統の画面内ブロックごとの特徴量を比較することによる、障害発生の検出(b)画面内の障害発生領域と非発生領域との間の特徴量差分を2系統で比較することによる障害発生系統の判定(3)さらに、前記符号化ノイズの差異を利用して、受信画像のSNも推定可能であることも見出した。以上のような要素技術を取り入れ、2重化テレビ回線の高エネルギー無瞬断映像伝送装置を世界で初めて実用化した。本装置は、既存の他の伝送装置構成を変える必要が無く、低コストでの導入が可能である。すでに数十台が各テレビ中継回線に導入され、テレビ中継の高信頼化に大きく貢献している。特に本装置の特長としては、ベースバンド映像の切替のため、圧縮エンコーダ/デコーダの障害にも対応可能という点や、衛星系とケ

ーブル系など異なる伝送システムで構成された2重化系でも使用可能という点があり、実運用に非常に便利なシステムとなっている。無瞬断化の単位としては他にATMセルやIPパケット、MPEG-2 TSパケット単位での方法が考えられるが、これらは上記のような特長を兼ね備えることが困難であり、本装置の優位性が明らかである。ただし、本装置では、テレビ回線の映像のみを監視して切替を決定しているため、音声のみが障害を発生したとしても切り替えることが出来ない。そこで、音声まで含めた無瞬断化が今後の課題である。

次に、デジタル伝送映像の遠隔監視方式については、監視用回線を利用して送信映像と受信映像を比較する枠組みを採用した。そして、なるべく少ない監視用情報量でなるべく高精度に受信映像のSN比を測定可能とする方式を検討の対象とした結果、画期的な方式を考案することが出来た。すなわち、送受信映像をスペクトル拡散した上で直交変換、特定の係数だけを抽出し監視回線で伝送する。送受信側から送られた係数の差分自乗平均により受信映像の誤差自乗平均を推定するものである。スペクトル拡散することにより、少数の係数を抽出するだけで高精度な推定が可能となった(2桁の精度向上)。本方式は、現在、プロトタイプとして装置化済みであり、今後、映像ネットワークの多地点同時監視の実現へ向けて洗練化が図られていく予定である。これにより、デジタル映像伝送の低コスト、高信頼化に大きく貢献することになる。

更に、次世代の高効率画像圧縮伝送方式については、その枠組みとして、領域分割符号化を対象として、検討を行った。従来、形状が複雑な絵柄に対して効率低下という問題のあった領域分割符号化に対して、形状情報量と動ベクトル情報、テクスチャ情報量をコスト要素として動的計画法を適用する方法を考案、レート歪みの観点から最適領域分割を実現することに成功した。これにより、絵柄によらず領域分割符号化の適用が可能となり、ブロックベース符号化特有のブロック歪みの無い画像伝送が実現できる。動的計画法適用のため計算量が多いが、次世代符号化方式の基盤として重要な方式を確立できた。現在、MPEG-4やH.264、Motion JPEG 2000などが新しい符号化方式として活発に使用され始めているが、本提案方式は、その次に来る超高効率符号化の基礎として重要となろう。上述の伝送の高信頼冗長化の観点からも、2重化のみならず現在でも要望のある3重化伝送のためには、画質を保ったままの更なる高圧縮が必須であり、そのための技術としても有効活用が期待できる。

今後の展望 今後、地上波デジタル放送が開始するなど、テレビ放送のデジタル化はますます進んでいく。また放送局内も、旧来の番組ソフトの活用のために依然として一部アナログ機材が残っているものの、それらもMPEG-2によるデジタルアーカイブ化が進み、トータルデジタル処理の流れは止まることは無いであろう。更に伝送面を見ると、現在はMPEG-2が主流であるが、処理遅延による掛け合い(中継先とスタジオとの会話)の際の不自然さの問題や、時報、CM差し替え問題(親局から地方局へ番組配信が行われる際、CMだけは地方局で差し替える場合がある。その際に、MPEG-2の処理遅延などのため、差し替えタイミングが指定時刻からずれてしまう現象)などが存在する。このため、今後、Motion JPEG2000など低遅延でも高画質な方式が広がる可能性がある。更には、光伝送技術の進展により、既存のファイバー回線の大容量化が進めば、特に国内においてテレビ中継伝送の非圧縮デジタル化が行われるようになるのもそう遠い話ではないであろう。

このようにしてデジタル中継回線が増大すると、伝送監視技術がますます重要となる。多地点、多回線を低コストで同時監視できることが監視システムの必須要件となる(あるリンクだけを正確に監視できるだけでは、全体から見るとそれほど便利とはいえない)。特に障害発生時の障害発生個所の切り分けや、無瞬断切替などが必須の技術として導入されるようになるであろう。

一方、オールデジタル化の流れの中でも、ヨーロッパと日米のテレビ方式、特にフレーム周波数の違い(50Hzと60Hz)は、既存の番組資産の問題などから、今後も残ることはまず間違いないであろう。そこで、これを変換する方式変換は今後も必須となり、その画質向上は、伝送自体での劣化が少なくなっていく中で、ますます重要な課題となっていくはずである。中継伝送において方式変換を責務の一つとしている通信会社のなお一層の努力に期待がかかるところである。

映像情報メディア、なかでもテレビジョンは、インターネットが普及した現在においても、世界の人々を結ぶ強力な力を持っている。それはむしろ衰えるどころか、通信ネットワークの充実により、回線数、中継時間数の急速な拡大により、ますます強力となっている。近くて遠かったアジア各国間の距離も、ファッションの共通化など、近くて近いものになりつつある。これには各国間でのテレビ番組の相互交換も大きな役割を果たしているといえよう。

本研究成果が、今後の国際テレビ中継の更なる普遍化と、それによる世界の人々の相互理解の更なる進展への一助となれば幸いである。

本論文に直接関係する文献

- [1] R.Kawada, T.Hamada and S.Matsumoto: “A new motion estimation with vision filter suitable for field interpolation of TV signals”, Proc. IEEE Workshop on Visual Signal Processing and Communications, Melbourne, Australia, pp. 203–206 (1993).
- [2] 川田, 浜田, 松本: “動き補正テレビ方式変換の改善”, 映情メ誌, **51**, 9, pp. 1577–1586 (1997).
- [3] R.Kawada and S.Matsumoto: “Flat multi-scalable coding for failure-free video transmission”, Proc. IEEE International Conf. on Image Processing, Vol. 3, Kobe, Japan, pp. 110–114 (1999).
- [4] 川田, 松本: “フラットマルチスケラブル高能率高信頼度映像伝送方式”, 信学論, **J84-B**, 3, pp. 621–631 (2001).
- [5] R.Kawada and S.Matsumoto: “A flat, multi-scalable, highly efficient and reliable video transmission scheme”, Electronics and Communications in Japan, Part I, **86**, 7, pp. 42–53 (2003).
- [6] 川田, 小池, 松本: “高能率無瞬断 TV 回線切換システムの開発”, 信学論, **J87-B**, 8, pp. 1106–1110 (2004).
- [7] R. Kawada, A. Koike, M. Wada and S. Matsumoto: “An efficient parallel video transmission system including codecs with functions of failure detection and coding noise evaluation”, IEEE Trans., **CSVT-xx**, xx, pp. xx–xx (in review) (2004).
- [8] R.Kawada, O.Sugimoto, M.Wada and S.Matsumoto: “Precise monitoring of remote picture PSNR by extracting reference data”, Proc. IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Kaohsiung, Taiwan, pp. 32–36 (2002).
- [9] 川田, 杉本, 小池, 和田, 松本: “スペクトル拡散と直交変換係数抽出による高精度遠隔映像 PSNR 推定方式”, 信学論, **J87-B**, 6, pp. 893–904 (2004).
- [10] 川田, 小池, 松本: “可変形状の領域分割動き推定方式”, 信学論, **J82-D-II**, 11, pp. 1972–1981 (1999).
- [11] R.Kawada, A.Koike and S.Matsumoto: “An integrated scheme of arbitrarily shaped segmentation and motion estimation”, Systems and Computers in Japan, **31**, 13, pp. 19–30 (2000).

- [12] 川田, 小池, 松本: “領域分割符号化のための領域統合最適化手法の提案”, 映情メ誌, **53**, 4, pp. 622–624 (1999).
- [13] 川田, 小池, 松本: “動画像符号化のための符号化速度適応形可変形状領域分割方式”, 映情メ誌, **55**, 5, pp. 751–759 (2001).

謝辞

終りに当たり、筆者が本研究をまとめるに際してご指導、ご鞭撻を賜った東京大学工学部相澤清晴教授に謹んで深謝の意を表します。さらに種々のご指導、ご助言を下された東京大学工学部原島博教授、相田仁教授、安田浩教授、瀬崎薫助教授、苗村健助教授に謹んで御礼申し上げます。

東京大学名誉教授、芝浦工業大学工学部高木幹雄教授には、大学院在学中並びに修了後を通じて、画像研究の基礎からはじめ、終始有益なご指導を賜った。ここに深甚の謝意を表します。

東京大学名誉教授、芝浦工業大学工学部水町守志教授には、筆者が研究を志すきっかけを与えていただくとともに暖かくご指導いただいた。謹んで感謝いたします。

また KDDI 株式会社執行役員技術開発本部長 村上仁己博士には、たえず暖かい励ましとともに卓越した見通しを持って御指導、御教示を賜った。深く感謝いたします。また (株) KDDI 研究所代表取締役 浅見徹所長、東京工業大学教授 元 KDDI 研究所副会長 羽鳥好律博士には、本研究をまとめるに当たり多大の御助言と御激励を賜った。謹んで感謝の意を表します。

KDDI 研究所取締役 松本修一博士、KDDI 株式会社開発推進部長 和田正裕博士、KDDI 研究所映像通信グループリーダ 小池淳博士、株式会社 KDDI メディアウィル社長 浜田高宏博士には、本研究遂行の上で直属の上司として、細部にわたる御指導と暖かい励ましを賜った。無量の感慨を込めてここに謝意を表します。

筆者が研究を進めるに当たり、装置開発において御協力いただいた沖電気工業株式会社、株式会社シバソク、エレテックス株式会社の関係者に、心から感謝いたします。

本研究の一部は、通信・放送機構からの委託研究「ISDB 技術に関する研究開発」に基づき行われた。ここに感謝の意を表します。

最後に、本研究を進めるに当たり、色々と御協力いただいた KDDI 研究所映像通信グループ及びマルチメディア通信グループの皆様に感謝いたします。

研究業績

(I) 学会誌論文

- [1] 川田, 浜田, 松本: “テレビジョン標準方式変換に適した動画像の動き推定方法 — 視覚フィルタの応用 —”, 国際通信の研究, 151, pp. 15–26 (1994).
- [2] 川田, 浜田, 松本: “高画質 TV 方式変換装置 VISION-TSC の開発”, 国際通信の研究, 157, pp. 41–52 (1997).
- [3] 川田, 浜田, 松本: “動き補正テレビ方式変換の改善”, 映情メ誌, 51, 9, pp. 1577–1586 (1997).
- [4] 内藤, 川田, 松本: “各種動き補償方式の実画像による比較検証”, 映情メ誌, 51, 9, pp. 1560–1567 (1997).
- [5] 浜田, 川田, 松本, 村上: “第二世代テレビ多重信号符号化システム MUCCSII の開発 — 動き補償直接符号化技術の導入 —”, 映情メ誌, 51, 4, pp. 583–591 (1997).
- [6] 浜田, 川田, 松本: “動き補償 EDTV-II 直接符号化システム”, 信学論, J81-B-I, 6, pp. 401–407 (1998).
- [7] 川田, 小池, 松本: “可変形状の領域分割動き推定方式”, 信学論, J82-D-II, 11, pp. 1972–1981 (1999).
- [8] 川田, 小池, 松本: “領域分割符号化のための領域統合最適化手法の提案”, 映情メ誌, 53, 4, pp. 622–624 (1999).
- [9] R.Kawada, A.Koike and S.Matsumoto: “An integrated scheme of arbitrarily shaped segmentation and motion estimation”, Systems and Computers in Japan, 31, 13, pp. 19–30 (2000).
- [10] 川田, 松本: “フラットマルチスケラブル高能率高信頼度映像伝送方式”, 信学論, J84-B, 3, pp. 621–631 (2001).
- [11] 川田, 小池, 松本: “動画像符号化のための符号化速度適応形可変形状領域分割方式”, 映情メ誌, 55, 5, pp. 751–759 (2001).
- [12] 杉本, 川田, 和田, 松本: “フレーム間差分信号への不可視マーカ埋め込みによる符号化画像自動客観画質評価方式の高精度化”, 情報科学技術レターズ, 1, 1, pp. 183–184 (2002).
- [13] R.Kawada and S.Matsumoto: “A flat, multi-scalable, highly efficient and reliable video transmission scheme”, Electronics and Communications in Japan, Part I, 86, 7, pp. 42–53 (2003).
- [14] 杉本, 川田, 小池, 和田: “低速監視回線に適用可能な特徴量抽出型画質監視方式”, 情報科学技術レターズ, 2, 2, pp. 235–237 (2003).
- [15] 川田, 杉本, 小池, 和田, 松本: “スペクトル拡散と直交変換係数抽出による高精度遠隔映像 PSNR 推定方式”, 信学論, J87-B, 6, pp. 893–904 (2004).
- [16] 松村, 内藤, 川田, 小池, 松本: “動的に更新される多層背景バッファを用いた高精度な自由視点映像生成方式”, 映情メ誌, 58, 6, pp. 850–853 (2004).

- [17] 川田, 小池, 松本: “高能率無瞬断 TV 回線切換システムの開発”, 信学論, **J87-B**, 8, pp. 1106–1110 (2004).
- [18] R. Kawada, A. Koike, M. Wada and S. Matsumoto: “An efficient parallel video transmission system including codecs with functions of failure detection and coding noise evaluation”, IEEE Trans., **CSVT-xx**, xx, pp. xx–xx (in review) (2004).

(II) 国際学会

- [1] R.Kawada, T.Hamada and S.Matsumoto: “A new motion estimation with vision filter suitable for field interpolation of TV signals”, Proc. IEEE Workshop on Visual Signal Processing and Communications, Melbourne, Australia, pp. 203–206 (1993).
- [2] R.Kawada and S.Matsumoto: “Flat multi-scalable coding for failure-free video transmission”, Proc. IEEE International Conf. on Image Processing, Vol. 3, Kobe, Japan, pp. 110–114 (1999).
- [3] O.Sugimoto, R.Kawada, M.Wada and S.Matsumoto: “A method for estimating psnr of coded pictures using embedded invisible markers without reference pictures”, Proc. IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Nashville, Tennessee, USA, pp. 42–46 (2001).
- [4] O.Sugimoto, R.Kawada, M.Wada and S.Matsumoto: “Automatic objective picture quality measurement method using invisible marker signals as reduced references”, Proc. World Multi-Conference on Systemics, Cybernetics and Informatics, Orlando, Florida, USA, pp. 311–316 (2001).
- [5] R.Kawada, A.Koike, M.Wada and S.Matsumoto: “A study on priority-based scalable coding”, Proc. International Workshop on Advanced Image Technology, Taejon, Korea, pp. 49–59 (2001).
- [6] R.Kawada, M.Wada and S.Matsumoto: “A method for automatic visual surveillance of wave overtopping”, Proc. International Workshop on ITS Telecommunications, Yokosuka, Japan, pp. 59–64 (2001).
- [7] R.Kawada, O.Sugimoto, M.Wada and S.Matsumoto: “Precise monitoring of remote picture PSNR by extracting reference data”, Proc. IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Kaohsiung, Taiwan, pp. 32–36 (2002).
- [8] R.Kawada, A.Koike, M.Wada and Y.Hatori: “Flat-scalable video communication for wireless transmission error tolerance”, Proc. SPIE Visual Communications and Image Processing, Lugano, Switzerland, pp. 1402–1409 (2003).
- [9] R.Kawada, A.Koike, M.Wada and Y.Hatori: “A method for precise detection of local impairment of coded video by use of reduced reference”, Proc. SPIE Visual Communications and Image Processing, Lugano, Switzerland, pp. 641–648 (2003).

(III) 研究会等

- [1] 川田, 松本, 村上: “ブロックマッチング動き検出における視覚フィルタのパラメータ最適化”, 画像符号化シンポジウム, No. 8-3, pp. 271-274 (1992).
- [2] 川田, 松本, 村上: “視覚モデルを用いたブロックマッチング動き検出”, 信学技報, No. IE92-37, pp. 25-31 (1992).
- [3] 川田, 齊藤, 松本: “A study on motion estimation using visual perception model”, 信学春大, No. D-299, pp. 7-41 (1992).
- [4] 川田, 松本: “Block matching motion estimation using a vision model”, 信学秋大, No. D-130, pp. 6-131 (1992).
- [5] 川田, 松本: “視覚モデルを用いた動き推定に関する検討”, テレビ年大, No. 11-10, pp. 221-222 (1992).
- [6] 浜田, 酒澤, 川田, 松本: “アダマール変換を用いたコンポジット動き補償方式”, テレビ年大, No. 12-3, p. 215 (1993).
- [7] 川田, 浜田, 松本: “動きの局所的变化に対応可能な動き補正 TV 方式変換”, テレビ技報, No. ICS94-44, pp. 13-18 (1994).
- [8] 川田, 浜田, 松本: “EDTV-II の高能率符号化に関する一検討”, 信学技報, No. IE94-86, pp. 47-52 (1994).
- [9] 川田, 浜田, 松本: “フィールド内挿のための視覚フィルタを用いた動き推定の検討”, 信学春大, No. D-352, pp. 7-85 (1994).
- [10] 川田, 浜田, 松本, 深沢, 大内: “前処理・動ベクトル整形付き階層型動き補正 TV 方式変換”, テレビ年大, No. 21-2, pp. 335-336 (1994).
- [11] 川田, 浜田, 松本: “コンポジット TV 信号符号化への MPEG-2 適用可能性について”, 信学ソ大, No. D-88, p. 90 (1995).
- [12] 川田, 浜田, 松本: “動き補正 TV 方式変換における高精度反復勾配動き推定の効果”, テレビ年大, No. 11-4, pp. 183-184 (1995).
- [13] 川田, 浜田, 松本: “EDTV-II 信号の分離符号化劣化に関する一検討”, 信学総大, No. D-350, p. 76 (1995).
- [14] 川田, 浜田, 松本: “EDTV-II 信号の分離符号化劣化に関する一検討”, 信学春大, No. D-350, p. 86 (1995).
- [15] 松本, 浜田, 川田: “45Mbps HDTV 符号化システム”, 信学技報, No. IE95-97, pp. 91-96 (1995).
- [16] 川田, 宮地, 浜田, 松本, 後藤, 長坂: “光海底ケーブルを介した日米間 HDTV / SDTV デジタル伝送実験”, テレビ年大, No. 23-2, pp. 342-343 (1996).

- [17] 小池, 川田, 松本: “領域分割に基づく動画像の高能率符号化方式”, 信学技報, No. IE96-108, pp. 67-72 (1996).
- [18] 川田, 浜田, 松本: “圧縮符号化の立場から見た YC 分離合成の検討”, 信学総大, No. D-213, p. 1 (1996).
- [19] 川田, 浜田, 松本: “MPEG-2 符号化制御情報を利用した TV 方式のビットストリーム変換”, 信学ソ大, No. D-209, p. 211 (1996).
- [20] 内藤, 川田, 松本: “MPEG-2 における各種動き補償方式の比較検証”, 信学総大, No. D-11-62, p. 62 (1997).
- [21] 内藤, 川田, 松本: “MPEG-2 における各種動き補償方式の比較検証”, 信学技報, No. IE96-133, pp. 77-82 (1997).
- [22] 小池, 川田, 松本: “任意形状単位の変換符号化に関する基礎検討”, 信学ソ大, No. D-11-43, p. 135 (1997).
- [23] 小池, 川田, 松本: “画像符号化のための多重オプティカルフローに基づく画像の領域分割”, 信学総大, No. D-11-64, p. 64 (1997).
- [24] 小池, 川田, 松本: “任意形状の直交変換を用いた画像符号化に関する基本的な検討”, 信学技報, No. IE96-144, pp. 21-26 (1997).
- [25] 川田, 小池, 松本: “動的計画法を用いた領域分割・動き推定の最適統合方式”, 映情メ年大, No. 20-10, pp. 252-253 (1997).
- [26] 川田, 小池, 松本: “ブロックマッチングを包含する領域分割・動き推定の統合化方式の検討”, 信学ソ大, No. D-11-42, p. 134 (1997).
- [27] 川田, 小池, 松本: “動画像符号化のための領域分割・動き推定の統合化方式”, 信学総大, No. D-11-63, p. 63 (1997).
- [28] 川田, 小池, 松本: “動き補償予測符号化に適した動画像の領域分割法”, 信学技報, No. IE96-130, pp. 53-60 (1997).
- [29] 川田, 小池, 松本: “領域分割符号化における動ベクトル統合化の検討”, 信学技報, No. IE97-12, pp. 23-28 (1997).
- [30] 浜田, 川田, 松本: “動き補償 DCT 符号化における逆 MAX 型非線形量子化の適用”, 映情メ年大, No. 21-2, pp. 258-259 (1997).
- [31] 川田, 小池, 松本: “動き補償予測符号化のための動き推定と領域分割の統合処理方式”, 信学技報, No. IE98-73, pp. 35-42 (1998).
- [32] 川田, 小池, 松本: “領域分割符号化のための領域統合最適化方式”, 映情メ冬大, No. 12-3, p. 101 (1998).
- [33] 鈴木, 小池, 川田, 松本: “四分木に基づく可変領域統合化法の演算高速化”, 信学総大, No. D-11-61, p. 61 (1998).

- [34] 川田, 松本: “フラットマルチスケラブル高能率耐障害映像伝送方式”, 映情メ技報, **23**, 50, pp. 7-12 (1999).
- [35] 川田, 松本: “2 系統の独立圧縮符号化における画質評価方式”, 映情メ冬大, No. 8-6, p. 105 (1999).
- [36] 川田, 松本: “2 重化圧縮符号化伝送映像の客観品質評価方式”, 信学ソ大, No. D-11-10, p. 96 (1999).
- [37] 川田, 浜田, 松本: “2 重化無瞬断映像伝送のための一検討 — 相互参照による画質評価方式 —”, 映情メ年大, No. 24-3, pp. 332-333 (1999).
- [38] 川田, 浜田, 松本: “無瞬断映像伝送のための映像障害検出方式”, 情処研報, No. AVM25-6, pp. 31-35 (1999).
- [39] 川田, 浜田, 松本: “フラットマルチスケラブル映像伝送方式”, 信学総大, No. D-11-7, p. 7 (1999).
- [40] 杉本, 川田, 松本: “電子透かしを用いた絶対画質客観評価方式”, 信学技報, No. HCS99-44, pp. 23-30 (1999).
- [41] 杉本, 川田, 松本: “不可視信号を用いた絶対画質客観評価方式における情報埋め込み方法に関する検討”, 映情メ冬大, No. 2-5, p. 48 (1999).
- [42] 浜田, 川田, 松本: “可変レート符号化方式の特性限界”, 情処研報, No. AVM24-4, p. 17 (1999).
- [43] 浜田, 川田, 松本: “デジタル画像信号処理プロセスの性能評価ツールの開発”, 情処研報, No. AVM25-10, pp. 55-62 (1999).
- [44] 川田, 和田, 松本: “映像監視による越波検出方式の検討”, 信学技報, No. ITS2000-36, pp. 7-12 (2000).
- [45] 川田, 和田, 松本: “映像による越波自動監視方式の検討”, 映情メ冬大, No. 10-9, p. 128 (2000).
- [46] 川田, 小池, 松本: “オブジェクトスケラブル符号化方式の検討”, 映情メ年大, No. 26-5, pp. 381-382 (2000).
- [47] 川田, 小池, 松本: “動的計画法を用いた領域分割動き補償予測符号化方式”, 情処研報, No. AVM28-10, pp. 55-60 (2000).
- [48] 川田, 松本: “映像監視による越波検出のための道路領域抽出方式”, 信学総大, No. A-17-21, p. 376 (2000).
- [49] 杉本, 川田, 和田, 松本: “リファレンス画像を用いない自動画質評価方式に関する一検討”, 画像符号化シンポジウム, No. P-02.01, pp. 77-78 (2000).
- [50] 杉本, 川田, 和田, 松本: “不可視信号埋め込みによるリファレンス画像なしの画質評価方式”, 映情メ年大, No. 5-6, pp. 70-71 (2000).

- [51] 杉本, 川田, 和田, 松本: “MPEG 符号化による劣化を検出するリファレンス画像なしの画質客観評価手法”, 信学技報, No. IE2000-122, pp. 13-18 (2000).
- [52] 杉本, 川田, 和田, 松本: “リファレンス画像を用いない高精度の客観画質評価方式の検討”, 映情メ冬大, No. 6-11, p. 97 (2000).
- [53] 杉本, 川田, 松本: “不可視信号埋め込みによる絶対画質客観評価方式”, 映情メ技報, 25(BCS2000-14), pp. 7-12 (2000).
- [54] 杉本, 川田, 松本: “絶対画質客観評価方式の性能比較”, 信学総大, No. D-11-2, p. 2 (2000).
- [55] R. Kawada, A. Koike, M. Wada and S. Matsumoto: “A study on priority-based scalable coding”, 信学技報, **24**, IE2000-172, pp. 17-22 (2001).
- [56] 川田, 和田, 松本, 山内, 山本, 石田: “小型高画質動き補正 TV 方式変換装置の開発”, 信学総大, No. D-11-85, p. 85 (2001).
- [57] 川田, 和田, 松本: “スペクトル拡散と WHT 係数抽出に基づく高精度伝送画質監視方式”, 信学ソ大, No. D-11-7, p. 92 (2001).
- [58] 川田, 杉本, 和田, 松本: “スペクトル拡散と直交変換係数抽出による伝送障害映像の PSNR 推定”, 映情メ冬大, No. 4-6, p. 76 (2001).
- [59] 川田, 杉本, 和田, 松本: “特徴量抽出による遠隔映像品質監視方式”, 映情メ年大, No. 18-5, pp. 268-239 (2001).
- [60] 川田, 杉本, 和田, 松本: “スペクトル拡散と WHT 係数抽出に基づく高精度遠隔画質監視方式”, 信学技報, No. IE2001-79, pp. 35-40 (2001).
- [61] 杉本, 川田, 和田, 松本: “拡散スペクトルの劣化推定に基づく符号化画像の自動客観評価方式”, 信学ソ大, No. D-11-28, p. 113 (2001).
- [62] 杉本, 川田, 和田, 松本: “不可視マーカ信号を用いた伝送映像の障害検出に関する検討”, 映情メ冬大, No. 4-7, p. 77 (2001).
- [63] 杉本, 川田, 和田, 松本: “WHT 係数操作による不可視マーカ埋め込みを用いた動画像の PSNR 推定方式”, 映情メ年大, No. 18-4, pp. 266-267 (2001).
- [64] 杉本, 川田, 和田, 松本: “WHT を用いた不可視マーカ埋め込みによる画質評価方式”, 情処研報, No. AVM33-1, pp. 1-6 (2001).
- [65] 杉本, 川田, 和田, 松本: “不可視マーカ信号埋め込みによる自動画質評価方式の理論検証”, 信学総大, No. D-11-63, p. 63 (2001).
- [66] 高木, 川田, 和田: “2 重化伝送の音声の無瞬断切替のための障害検出法”, 映情メ冬大, No. 11-1, p. 124 (2001).
- [67] 竹沢, 長谷山, 北島, 川田: “GA を用いた IFS パラメータの決定法における探索性能の向上”, 平成 14 年度電気関係学会北海道支部連合会, No. 220, p. 272 (2002).

- [68] 川田, 和田, 松本: “動的パラメータ制御によるシーン適応型の高精度画像動き推定方式”, 映情メ年大, No. 9-3 (2002).
- [69] 川田, 杉本, 和田, 松本: “局所的画質劣化を考慮した PSNR 高精度遠隔推定方式”, 信学総大, No. D-11-3, p. 3 (2002).
- [70] 杉本, 川田, 和田, 松本: “不可視マーカ信号に基づく伝送映像の障害検出に関する検討”, 信学総大, No. D-11-6, p. 6 (2002).
- [71] 杉本, 川田, 和田, 松本: “不可視マーカ信号の特徴量抽出と統計的検定に基づく映像障害検出方式の検討”, 信学技報, No. CMN02-21, pp. 19-23 (2002).
- [72] 杉本, 川田, 和田, 松本: “フレーム間差分信号への不可視マーカ埋め込みによる符号化画像自動客観画質評価方式の高精度化”, 情報科学技術フォーラム, No. LJ-10, pp. 183-184 (2002).
- [73] 杉本, 川田, 和田, 松本: “高精度映像障害検出のための不可視マーカ信号の特徴量抽出方式”, 映情メ年大, No. 9-2 (2002).
- [74] 川田, 杉本, 和田, 松本, 羽鳥: “局所的画質劣化を考慮した特徴量抽出型遠隔 PSNR 高精度監視方式”, 信学技報, 101, 722(IE2001-234), pp. 17-22 (2002).
- [75] 川田, 山下, 内藤, 小池, 和田: “多視点映像伝送のための高精度反復勾配法による実時間ステレオマッチング方式”, 映情メ冬大, No. 7-5, p. 1 (2002).
- [76] 川田, 杉本, 和田: “伝送エラー耐性のためのフラットスケラブル画像通信方式”, 情報科学技術フォーラム, No. J-72, pp. 345-346 (2002).
- [77] 杉本, 川田, 小池, 和田: “画像特徴量の送受信間比較による高精度映像伝送障害検出方式”, 映情メ冬大, No. 5-6 (2002).
- [78] 高木, 川田, 和田: “2重化された圧縮音声信号の障害検出法の検討”, 信学総大, No. D-10-3, p. 145 (2002).
- [79] 杉本, 川田, 小池, 和田: “検出閾値の適応制御に基づく映像の局所劣化検出方式の高精度化”, 映情メ年大, No. 16-1 (2003).
- [80] 杉本, 川田, 小池, 和田: “低速監視回線に適用可能な特徴量抽出型画質監視方式”, 情報科学技術フォーラム, No. LJ-010, pp. 235-237 (2003).
- [81] 川田, 山下, 内藤, 小池: “3D 映像伝送のための階層型反復勾配法による高精度ステレオマッチング方式”, 情報科学技術フォーラム, No. J-075, pp. 347-349 (2003).
- [82] 松村, 内藤, 川田, 小池, 松本: “任意視点動画像の高圧縮伝送を目的とした隠蔽領域補完方式の提案”, 信学技報, IE2003-66(OIS2003-41), pp. 63-68 (2003).
- [83] 松村, 内藤, 川田, 小池, 松本: “背景バッファの適応処理を用いた高精度な自由視点映像の補完”, 映情メ冬大, No. 8-8, p. 88 (2003).

- [84] 杉本, 川田, 小池, 松本: “画像特徴量の多値表現に基づく遠隔自動画質監視方式の高精度化の検討”, 映情メ冬大, No. 9-4, p. 94 (2003).
- [85] 松村, 内藤, 川田, 小池, 松本: “自由視点映像に用いる動的背景バッファの適応処理に関する検討”, 信学技報, IE2003-242(ITS2003-107), pp. 107-110 (2004).
- [86] 川田, 小池, 松本: “高信頼素材伝送方式—高能率無瞬断 TV 回線切替システムの開発—”, 情処全大, No. 2TTA-4, pp. 5.113-116 (2004).
- [87] 杉本, 川田, 小池, 松本: “不可視マーカ方式による IP ベース映像配信品質自動監視システム”, 情処全大, No. 2TTB-5, pp. 5.213-216 (2004).
- [88] 杉本, 川田, 小池, 松本: “不可視マーカ埋め込み強度の適応制御による自動客観画質評価方式の高精度化”, 信学総大, No. D-11-30, p. 30 (2004).
- [89] 川田, 杉本, 小池: “方式変換映像の遠隔監視方式”, 信学総大, No. D-11-64, p. 64 (2004).
- [90] 松村, 内藤, 川田, 小池, 松本: “背景バッファの適応処理を用いた高精度な自由視点映像の補完”, 信学総大, No. D-11-123, p. 123 (2004).
- [91] 川田, 杉本, 小池, 松本: “方式変換映像の遠隔画質監視方式”, 信学技報, IE2004-15, pp. 1-6 (2004).

(IV) 解説

- [1] 川田：“マルチメディアの基本デバイス—可変長符号化・復号 LSI “VLCODEC” ”, KDD テクニカルジャーナル, 17, pp. 24–25 (1994).
- [2] 松本, 川田：“客観評価法の種類と特徴”, 映情メ誌, 53, 9, pp. 1199–1205 (1999).
- [3] 川田：“符号化効率向上のための画像の領域分割方式”, 画像ラボ, 12, 9, pp. 11–16 (2001).
- [4] 川田, 小池：“映像品質の客観評価技術の標準化動向”, 信学技報, No. CQ2003-31, pp. 41–44 (2003).
- [5] 川田, 内藤, 松本：“テレビジョン伝送と CATV に関する ITU-T SG9 の最近の標準化動向 (その 1) ・ MPEG-2 関連技術と客観品質評価技術”, 映情メ誌, 58, 5, pp. 655–661 (2004).
- [6] 川田：“デジタル・ビデオ画像の評価の実際”, トランジスタ技術, 7, pp. 193–197 (2004).

(V) 受賞

| 表彰年月 | 表彰業績名称 | 表彰主体 | 受賞者 |
|----------|-----------|----------------|--------|
| 1995年6月 | 放送文化基金賞 | (財) 放送文化基金 | 他5名と共同 |
| 1998年3月 | 学術奨励賞 | (社) 電子情報通信学会 | 本人 |
| 1998年4月 | 社長表彰 | 国際電信電話(株) | 他2名と共同 |
| 2000年12月 | 最優秀プロダクト賞 | (株) KDDI 研究所 | 本人 |
| 2000年12月 | 最多特許発明賞 | (株) KDDI 研究所 | 本人 |
| 2001年5月 | 技術振興賞開発賞 | (社) 映像情報メディア学会 | 本人他2名 |
| 2001年8月 | 鈴木記念賞 | (社) 映像情報メディア学会 | 本人 |
| 2001年12月 | 研究奨励賞 | (社) 映像情報メディア学会 | 本人 |
| 2004年3月 | 優秀プロダクト賞 | (株) KDDI 研究所 | 本人 |
| 2004年5月 | 日本ITU協会賞 | (財) 日本ITU協会 | 本人 |