

## ■2 群 (画像・音・言語) -5 編 (画像符号化)

### 4 章 予測符号化

(執筆著者：宮田高道) [2013年3月 受領]

#### ■概要■

一般に静止画像や動画では、空間的あるいは時間的に近接する画素の値(画素値)は近い値をとることが多い。このような関係は空間的・時間的な画素間相関と呼ばれており、画像符号化の符号化効率を向上させるためには、この相関を可能な限り除去することが有効である。既存の画像符号化に関する検討では、この空間的・時間的画素間相関除去という問題に対して、離散コサイン変換(DCT)やウェーブレット変換などの変換符号化(Transform Coding)、及び本章で解説する予測符号化(Predictive Coding)の二つのアプローチが主にとられてきた。

変換符号化による相関除去は、画像の一部あるいは全体をベクトルとみなし、座標軸の回転などに相当する線形変換を行うことで特定の出力係数にエネルギーを集中させ、それによって符号化効率の向上を可能とするものである。これに対して予測符号化とは、着目している画素(対象画素)と空間的・時間的に近い位置にある画素(参照画素)の値から、予測関数に従って対象画素の画素値を予測し、このときの予測値と実際の画素値との誤差(予測誤差)を符号化する手法の総称である。予測関数が十分に正確であれば、予測誤差は0や0に近い値をとる可能性が高く、この誤差に対して変換、量子化及びエントロピー符号化を適用することで大幅な圧縮効率の向上が期待できる。現在では、空間的、時間的画素間相関の両方の除去について、予測符号化と変換を組み合わせた枠組みが一般的に用いられている。

予測符号化の元祖は、1950年代にCutlerによって提案されたDPCM(Difference Pulse Coded Modulation)であり<sup>1)</sup>、デジタル画像符号化への応用についても1960年代後半から続く長い歴史をもっている(例えば文献2), 3)など)。最も単純な予測符号化は、対象画素と同一ライン上の直前の既符号化画素を参照画素とする前値予測である。本章で解説する様々な予測符号化法は、この手法をベースとして、a)予測に使用する参照画素を、どこからどれだけもってくるのか、b)対象画素と参照画素との間にどのような関係を仮定するのかの2点について、それぞれ改良を施したものであるといえる。予測符号化は長い歴史をもつが、近年にいたってもH.264やH.265の標準化において非常に多くの研究がなされており、画像符号化の中でも特に活発な分野の一つであるといえる。

#### 【本章の構成】

本章では予測符号化の原理(4-1節)を説明したのち、静止画像や動画の一フレーム内で、空間的な画素間相関を用いられるフレーム内予測(4-2節)、時間的な画素間相関を除去するために動画の複数のフレームをまたいで処理を行うフレーム間予測(4-3節)、更に、フレーム間予測の予測精度を向上させるために必要不可欠な動き補償(4-4節)を解説し、最後にこれらを発展させたより高度な予測処理(4-5節)について述べる。

## ■2群 - 5編 - 4章

### 4-1 予測符号化の原理

(執筆者：宮田高道) [2013年3月 受領]

前節で述べたように、画像符号化における予測符号化とは、時間的・空間的に近接する画素の画素値に近い値をとる傾向（空間的・時間的冗長性）があることを利用し、符号化対象の画素値とその予測値との差分を符号化することで符号化効率の向上を図る技術である。予測符号化は、除去する冗長性が空間的か時間的のいずれであるかによってフレーム内予測符号化とフレーム間予測符号化の二種類に分類できるが、本節ではそれらに共通する原理の解説を行う。

#### 4-1-1 一般的な予測符号化モデル

符号化対象のデジタル画像信号（静止画像あるいは動画像）を  $\mathbf{u}$  とし、その時空間上のある位置  $n \in \mathbf{Z}$  の画素値を  $u(n) \in \mathbb{R}$  とする。いま、符号化処理が  $u(m-1)$  の画素まで行われ、これから  $u(m)$  の符号化を行うところであるとする。このとき、予測誤差  $e(m)$  は

$$e(m) := u(m) - \bar{u}(m) \quad (4 \cdot 1)$$

$$\bar{u}(m) := P_m(N_m(\tilde{\mathbf{u}})) \quad (4 \cdot 2)$$

と定義される。ここで、 $\tilde{\mathbf{u}}$  は既に符号化済みの画像信号である。関数  $N_m$  は位置  $m$  からみた（一般に複数の）参照画素を  $\tilde{\mathbf{u}}$  から抽出する働きをもつ参照関数、関数  $P_m$  は参照画素から予測画素値  $\bar{u}(m)$  を得る予測関数である。これらの関数は一般には位置  $m$  に依存する。また、量子化後の予測誤差を  $\tilde{e}(m)$  とすると、復号時の再生値  $\tilde{u}(m)$  は  $\tilde{u}(m) = \bar{u}(m) + \tilde{e}(m)$  と表せる。

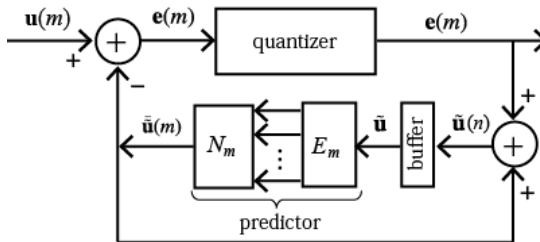


図 4・1 予測符号化の処理の流れ

以上の処理の流れを図 4・1 に示す。ここで、係数の量子化が予測のループの内部に配置されていることに注目されたい。仮に量子化を予測器の外側で行った場合には、一般にドリフトと呼ばれる量子化誤差の蓄積が生じ、復号後の再生値と入力信号間に大きな誤差が生じる。図 4・1 の構成であれば、符号化誤差  $\tilde{u}(m) - u(m) = \tilde{e}(m) - e(m)$  となり、各  $m$  の量子化誤差がそのまま  $m$  の符号化誤差となる。ロスレス符号化の場合には、図中の量子化器が存在せず、 $\tilde{\mathbf{u}} = \mathbf{u}$ 、 $\tilde{\mathbf{e}} = \mathbf{e}$ 、 $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}$  が成り立つ。

#### 4-1-2 予測誤差の分散

本節では、前節で用いた単純な予測符号化を例にとり、予測誤差  $\mathbf{e}$  の分散を求めることで、予測符号化の効果を確認する。簡単のため、信号  $u(m)$  が平均 0 となるよう前処理されているとすると、その分散は  $\sigma_u^2 := E_m[u(m)u(m)]$  と表せる。ここで、 $E_m[\cdot]$  は  $m$  についての期待値を意味する。また、信号  $u(m)$  と予測信号  $u(m-1)$  との（自己）相関係数を

$$\rho := \frac{1}{\sigma_u^2} E_m[u(m)u(m-1)]$$

と定義する。このとき、予測誤差の分散  $\sigma_e^2$  は、

$$\sigma_e^2 = E_m[\{u(m) - u(m-1)\}^2] \quad (4 \cdot 3)$$

$$\begin{aligned} &= E_m[u(m)^2] - 2E_m[u(m)u(m-1)] + E_m[u(m-1)^2] \\ &= 2\sigma_u^2 - 2E_m[u(m)u(m-1)] \\ &= 2\sigma_u^2 - 2\rho\sigma_u^2 \\ &= 2(1 - \rho)\sigma_u^2 \end{aligned} \quad (4 \cdot 4)$$

となる。

自然画像では、多くの場合  $\rho = 0.90 \sim 0.98$  であることが知られている。式(4・4)より、相関係数  $\rho$  が大きくなると、それにつれて予測誤差の分散は小さくなるのが分かる。仮に予測誤差の分布がラプラス分布であると仮定すると、分散が小さくなるにつれてエントロピーも単調減少する。すなわち、画素間の相関が十分に高ければ、予測によって平均符号長を短くできる。

## ■2 群 - 5 編 - 4 章

### 4-2 フレーム内予測

(執筆著：宮田高道) [2013年3月 受領]

#### 4-2-1 単純フレーム内予測

予測符号化を、静止画像あるいは動画の単一フレームの符号化に対して行うことをフレーム内予測符号化、あるいはイントラ符号化 (Intra-frame Predictive Coding) と呼ぶ。いま、ラスタスキャン順に 1 次元ベクトルとして並べ直した  $w \times h$  画素のデジタル画像信号  $\mathbf{u} \in \mathbb{R}^{wh}$  に対して、 $N_m: \mathbb{R}^{wh} \rightarrow \mathbb{R}: \mathbf{u} \mapsto u(m-1)$ 、 $P_m: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto x$  とおいた、フレーム内線形予測ロスレス符号化を考える。このとき、予測誤差は、 $e(m) = u(m) - u(m-1)$  となる。

標準画像 Lena の画素値  $\mathbf{u}$  と、画像の各横ラインごとに予測行ったときの予測誤差  $\mathbf{e}$  の生起頻度の分布を図 4・2、図 4・3 に示す。これらの図から、 $\mathbf{e}$  の分布はラプラス分布に近く、生起確率が 0 付近に集中していることが分かる。よって、エントロピー符号化を施した際の平均符号長も小さくなることが期待される。実際、信号  $\mathbf{u}$  のエントロピーが 7.42 [bit] であるのに対し、予測誤差  $\mathbf{e}$  のエントロピーは 5.37 [bit] となっており、単純な差分をとるだけの処理でも、平均符号長を大幅に短くできることが分かる。

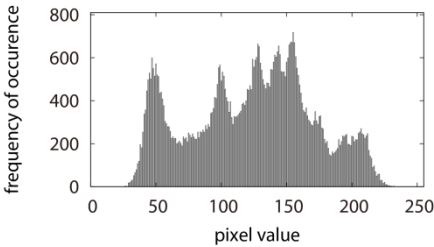


図 4・2  $\mathbf{u}$  のヒストグラム

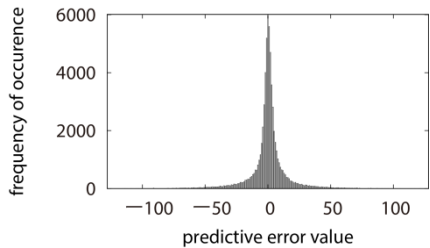


図 4・3  $\mathbf{e}$  のヒストグラム

この単純フレーム内予測と同様の手法は、ロッシェン符号化である JPEG において、ブロック DCT 係数の直流成分の符号化に用いられている<sup>5)</sup>。

#### 4-2-2 最適係数

予測誤差の分散を最小にするような予測を最適予測 (Optimal Prediction)<sup>2)</sup> と呼ぶ。いま、予測関数と参照関数の合成  $P_m(N_m(\cdot)): \mathbb{R}^{wh} \rightarrow \mathbb{R}$  を

$$P_m(N_m(\mathbf{u})) : \mathbf{u} \mapsto \sum_{i=1}^q \alpha_i u(m - p_i) \quad (4 \cdot 5)$$

と定義する。このとき、予測に使用する画素の数  $q$  を次数 (Order) と呼び、ベクトル  $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_q\} \in \mathbb{R}^q$  を予測係数ベクトルと呼ぶ。変位  $p_i$  は  $i$  番目の参照画素と  $m$  の位置のずれを表しており、便宜上  $p_0 := 0$  とおく。このとき、共分散  $R_{ij}$  を  $R_{ij} := E_m[\{u(m-p_i) - u(m-p_j)\}^2]$  と定義する。予測誤差の分散  $\sigma_e^2 = E_m[\{u(m) - \tilde{u}(m)\}^2]$  を  $\alpha_j (j = 1, \dots, q)$  で偏微分して 0 とおくことにより、 $R_{0j} = \sum_{i=1}^q \alpha_i R_{ij}$  という  $q$  本の連立方程式が得られる。これより最適な係数ベク

トル  $\alpha^* := \{\alpha_1^*, \alpha_2^*, \dots, \alpha_q^*\} \in \mathbb{R}^q$  は

$$\begin{pmatrix} \alpha_1^* \\ \alpha_2^* \\ \vdots \\ \alpha_q^* \end{pmatrix} = \begin{pmatrix} R_{11} & R_{21} & \cdots & R_{q1} \\ R_{12} & R_{22} & \cdots & R_{q2} \\ \vdots & \vdots & \ddots & \vdots \\ R_{1q} & R_{2q} & \cdots & R_{qq} \end{pmatrix}^{-1} \begin{pmatrix} R_{01} \\ R_{02} \\ \vdots \\ R_{0q} \end{pmatrix} \quad (4 \cdot 6)$$

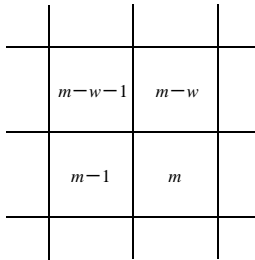
として得られる<sup>3)</sup>.

#### 4-2-3 適応予測

前節で計算した最適係数は、画像がいたるところ定常的なモデルで表現可能な場合には予測誤差を最小にすることができる。しかしながら、実際には1枚の静止画像中にも様々な特徴をもつ領域が混在しているため、単一の予測係数ベクトル  $\alpha$  のみで予測符号化を行うことは符号化効率の低下につながる。また、画像中の領域ごとに異なる最適係数ベクトルを算出して予測に用いる場合は、それらの係数を符号化データに追加する必要があり、オーバーヘッドの増加を招く。そこで、あらかじめ定めた数種類の予測係数ベクトルを、位置  $m$  周囲の画素の状況に応じて切り替えて使用する方式が考えられる。このようなアプローチを適応予測 (Adaptive Prediction) と呼ぶ。

##### (1) Lossless-JPEG

ロスレス符号化の一種である Lossless-JPEG<sup>6)</sup> では、着目している画素の周囲の3画素から線形予測を行う。式(4.5)において  $q=3$  とし、 $p_1 := 1$ ,  $p_2 := w$ ,  $p_3 := w + 1$  とおくと、空間位置  $m$  の画素を符号化する際の Lossless-JPEG の各参照画素の位置は、**図 4.4** のようになる。また、係数ベクトル  $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$  の各値は**表 4.1** のように定義され、この7種類の予測係数  $\alpha$  (予測モードと呼ばれる) を適宜切り替えながら用いる。



**図 4.4** Lossless-JPEG の予測画素

**表 4.1** Lossless-JPEG の予測係数

係数 \ 予測モード	1	2	3	4	5	6	7
$\alpha_1$	1	0	0	1	1	0.5	0.5
$\alpha_2$	0	1	0	1	0.5	1	0.5
$\alpha_3$	0	0	1	-1	-0.5	-0.5	0

## (2) H.264 フレーム内符号化

標準化された動画像符号化方式である MPEG-4 AVC/H.264<sup>7)</sup> では、フレーム内符号化に予測符号化が採用されている。H.264 のフレーム内予測は、これまで見てきた画素ごとの予測符号化と異なり、画素をまとめたブロックごとに予測を行う。ブロックサイズは、H.264 では 16×16, 4×4 が用いられる。また、H.264 の拡張である FExt (Fidelity Range Extensions)<sup>8)</sup> では更に 8×8 が追加された。

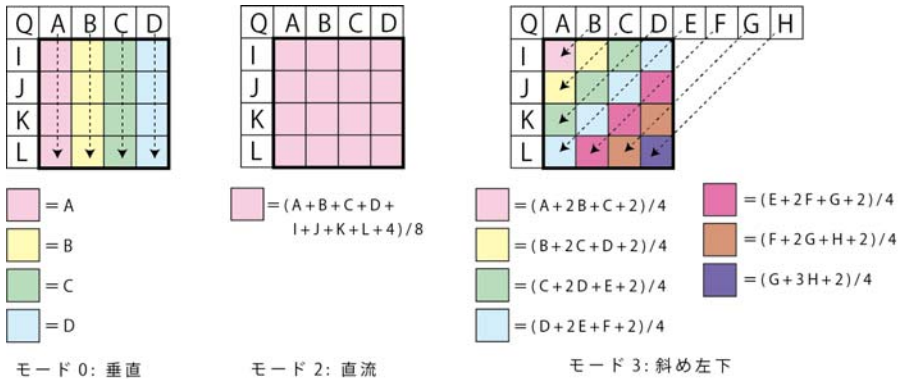


図 4・5 H.264 フレーム内予測符号化, 4×4 ブロックの予測モード (抜粋)

H.264 の 4×4 ブロックに対する予測符号化の予測モードの一部を図 4・5 に示した。このように、予測対象のブロック内の色分けされた画素値に対し、周囲の符号化済みの画素 (A~L, Q) から、図中の予測式に従って予測を行う。実際にはここに示した 3 種類含む 9 種類の予測モードが準備されている。H.264 では、ブロックサイズと予測モードの両方を、符号化効率が高くなるように選択して使用することで高い符号化効率を実現している。一方で、適切なブロックサイズ及び予測モードの選択には膨大な計算量が必要であり、これを効率化する研究も数多く存在する<sup>13)</sup>。

## ■2群 - 5編 - 4章

### 4-3 フレーム間予測

(執筆著：宮田高道) [2013年3月 受領]

フレーム間予測符号化 (Inter Frame Predictive Coding) とは、動画像の連続するフレームの間に存在する冗長性を除去することで符号化効率を向上させる技術の総称である。最も単純なフレーム間予測符号化は、符号化対象のフレーム (対象フレーム) の画素値の予測値として、直前のフレーム (参照フレーム) の同じ位置の画素値を用いるものである。いま、時間  $t$  のフレームを対象フレームとし、その位置  $(x, y)$  における画素値を  $u(x, y, t)$  と表すと、予測されたフレーム (予測フレーム) の画素値は  $\tilde{u}(x, y, t) = \tilde{u}(x, y, t-1)$  として得られる。このときの予測誤差は  $e(x, y, t) = u(x, y, t) - \tilde{u}(x, y, t)$  である。

#### 4-3-1 動き補償

前述の単純なフレーム間予測では、対象動画像がカメラワークや物体の動きを含む場合には予測誤差が大きくなり、逆に情報量が増えることもあり得る。そこで、対象フレームと参照フレームとの間で局所的な動きを求め、これに基づいて参照画像から予測画像を作成することを考える。これを動き補償 (Motion Compensation) フレーム間予測<sup>9)</sup> と呼び、MPEG-1<sup>10)</sup>、MPEG-2<sup>11)</sup> とそれ以降の動画像符号化において圧縮効率向上の多くの部分を担っている。

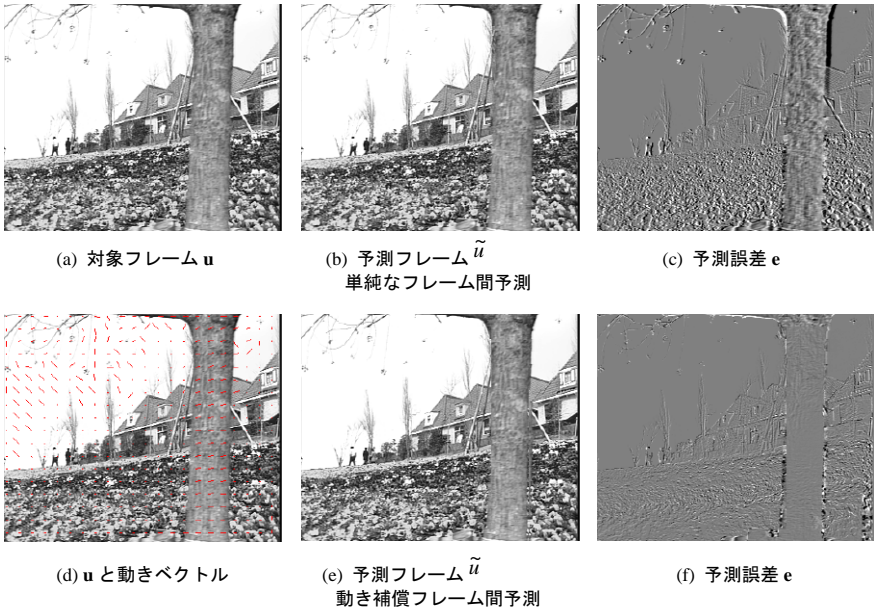


図 4・6 フレーム間予測方式の比較

いま、対象フレームを(例えば16×16画素の)ブロックに分割し、それらを対象ブロックと呼ぶこととする。また、各対象ブロックについて、参照フレーム内で最も類似したブロック(予測ブロック)を発見し、対象ブロックと参照ブロックの位置のずれを動きベクトル(Motion Vector)と呼ぶ。対象フレームの画素位置(x, y)を含む対象ブロックにおける動きベクトルを $(v_x(x, y), v_y(x, y))$ とすると、動き補償予測によって得られる予測フレームの画素値は、 $\bar{u}(x, y, t) = \bar{u}(x - v_x(x, y), y - v_y(x, y), t - 1)$ と表すことができる。前述の単純なフレーム間予測と、動き補償フレーム間予測の比較を図4・6に示す。この図より、動き予測補償を用いることにより予測誤差の分布は0付近に集中し、符号化効率の向上が期待できることが分かる。図4・6(d)中の赤い線分は、16×16画素のブロックごとの動きベクトルを表現している。これを求める具体的な方法については、4・4節で述べる。

### 4-3-2 予測フレームの分類とGOP構造

MPEG-1, MPEG-2などの符号化方式では、参照フレームの位置と数によって、各対象フレームをIフレーム(Intra Coded), Pフレーム(Predictive Coded), Bフレーム(Bidirectionally Predictive Coded Frame)の三種類に分類しており、Iフレームを1枚含む連続するフレームの集合をGOP(Group of Picture)と呼ぶ。その一例を図4・7に示す。

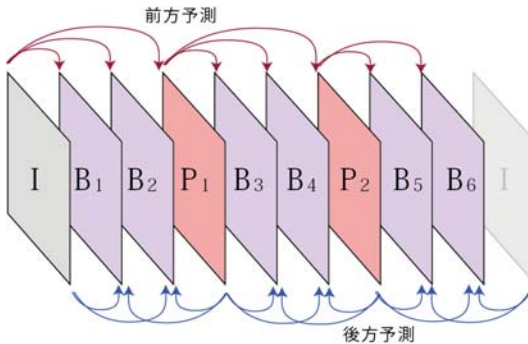


図4・7 GOP構造の例

ここで、Pフレームは、過去のIフレームから前述の動き補償を行い予測フレームを構成するフレームであり、図中のP<sub>1</sub>はIを参照フレームとしている。Bフレームは過去と未来の二つのフレームからの動き補償によって予測フレームを構成するフレームであり、双方方向からの参照を行うことによって、予測精度の向上が見込める。図ではP<sub>1</sub>とP<sub>2</sub>がB<sub>3</sub>の参照フレームとなっている。残るIフレームはフレーム間予測を行わないフレームであり、このIフレームを適当な間隔で挿入することにより、伝送誤りなどで正確な復号ができない状態が続くことを防ぐ。また、Iフレームはランダムアクセスの際の再生開始点や、編集を行う際の編集点としても使用できる。

予測フレームは符号化済みのフレームである必要がある一方で、Bフレームは時間的に未来のフレームを参照フレームとしているため、各フレームは時間順と異なる順序で符号化処



理を行う必要がある。すなわち，図4・7では符号化の順序はI, P<sub>1</sub>, B<sub>1</sub>, B<sub>2</sub>, P<sub>2</sub>, B<sub>3</sub>, …となる。そのため，符号化／復号時には複数のフレームをバッファに蓄えてから処理を行う必要があり，これは処理遅延の増大を意味する。

## ■2 群 - 5 編 - 4 章

### 4-4 動き検出

(執筆: 宮田高道) [2013年3月 受領]

前節で解説した動き補償フレーム間予測符号化においては、動きベクトルを正確に、かつ可能限り少ない計算量で求める処理が必要となる。このような処理を動き検出 (Motion Detection) といい、前節のように、画像をブロックに分割しブロックごとの動きベクトルを求める手法はブロックマッチング (Block Matching) 法と呼ばれる。いま、対象フレームを  $\mathbf{u}$ 、参照フレームを  $\mathbf{s}$  と表す。ブロックの大きさを  $w \times w$  画素とし、画素位置  $(x, y)$  をブロックの左上角にもつ対象ブロックに着目する。このとき、参照フレーム中で、対象ブロックに対応する位置から  $(v_x, v_y)$  だけ移動した位置にあるブロックの予測誤差は

$$D(v_x, v_y) := \sum_{i=0}^{w-1} \sum_{j=0}^{w-1} L(u(x+i, y+j), s(x+i+v_x, y+j+v_y)) \quad (4 \cdot 7)$$

のように計算できる。ここで、画素値間の距離  $L(a, b)$  の定義としては、2乗誤差  $L(a, b) = (a - b)^2$  や絶対値誤差  $L(a, b) = |a - b|$  が用いられる。これらの記号を用いると、ブロックマッチングによる動き検出は、 $(v_x^*, v_y^*) := \arg \min_{(v_x, v_y) \in S} D(v_x, v_y)$  と表すことができる。なお、集合  $S$  は動きベクトルの探索範囲を表す。

#### 4-4-1 全探索

全探索 (Exhaustive Search) は、予測フレーム内で、対象ブロックに対応する位置の周辺をくまなく探索して予測ブロックを発見する最も単純な動き検出手法である。全探索に必要な計算量は、探索範囲  $S$  を近傍に絞った場合でも莫大であり、現実的とはいえない。そこで、計算量削減の様々な手法が提案されている<sup>16)</sup>。以下にその一部を解説する。

#### 4-4-2 木探索

探索範囲  $S$  内の画素位置を木構造の内部ノードとみなし、根から葉にむけて木を辿るようにマッチングを行うことで計算量を抑制するアプローチを木探索という。探索範囲を水平、垂直ともに  $+s$  画素とした場合、最初のステップでは中心位置  $(0, 0)$  を含む  $(\pm 4, \pm 4)$  の 9 点について誤差が最小となる点を求め、次にその最小点を中心に  $(\pm 2, \pm 2)$  のやはり 9 点についてマッチングを行う。これを繰り返すことで、動きベクトルを探索する。この手法は  $S$  内における誤差の分布が比較的なめらかであるという仮定に基づいており、その単純性と高い性能から多くの実装で用いられている。

#### 4-4-3 階層探索

階層探索<sup>15)</sup> は、対象フレームと参照フレームとを段階的に低解像度にした複数の画像 (多重解像度ピラミッドなどと呼ばれる) を生成し、最低解像度のフレーム間でブロックマッチングを行い、その結果を一つ上の解像度のフレームでの探索の初期位置とするものである。探索範囲が広い場合に効果的であるが、最低解像度の画像では画像の精細なテクスチャが失われているため、探索の早い段階で局所最小解にとらわれる可能性もある。

## ■2群 - 5編 - 4章

### 4-5 高度な予測処理

(執筆者：宮田高道) [2013年3月 受領]

本節では、MPEG-4 part 2 から H.264/AVC までに検討、採用されてきた、種々の高度な予測処理について概説する。

#### 4-5-1 ループ内デブロッキングフィルタ

ブロック単位の符号化処理では、ブロック境界にブロックひずみと呼ばれるノイズが発生する。フレーム間予測符号化の参照フレームにおけるブロックひずみは、予測誤差を増加させる要因になる。そこで H.264/AVC<sup>7)</sup> では、予測のループ内 (図 4-1 のバッファと予測器の間) で、ブロックの境界線に隣接する画素値にフィルタ処理を施す、ループ内デブロッキングフィルタ (In-the-loop Deblocking Filtering) が採用されている。

#### 4-5-2 オーバーラップブロック動き補償

参照フレームのブロックひずみを取り除いた場合でも、予測フレームを重ねりのない参照ブロックを並べて構成する以上は、得られる予測フレーム内にブロックひずみが発生する可能性がある。そこで、予測フレームの生成時に、対象ブロックの動きベクトルで補償した画素値と、隣接する上下左右のブロックの動きベクトルで補償した画素値とを重み付け平均することで、このようなブロックひずみの発生を抑える技術がオーバーラップブロック動き補償 (Overlapped Block Motion Compensation)<sup>19)</sup> であり、MPEG-4 part 2<sup>12)</sup> ではオプションとして採用された。

#### 4-5-3 グローバル動き補償

動画像中にパン、チルト、ズームなどのカメラワークによる大域的な動きが存在する場合には、前述のブロックマッチングによる動き補償のように、予測画像を大量の局所的な動きベクトルから作るアプローチは効率的とはいえない。このような場合に、カメラワークを表現するごく少数のパラメータを用いて予測画像の生成に必要な符号量を抑える手法は古くから研究されており<sup>17)</sup>、これを発展させたものがグローバル動き補償 (Global Motion Compensation) として MPEG-4 part 2<sup>12)</sup> に採用されたものの、H.264/AVC では採用を見送られている。

#### 4-5-4 輝度の重み付き予測

フェードイン／アウトのように、動画像中のフレーム全体の輝度値が同時に変化する場合には、この輝度変化を少数の係数で表現し、これに基づいて予測画像を作ることで予測誤差を減少させることが可能である。この技術は輝度の重み付き予測 (Weighted Prediction)<sup>18)</sup> と呼ばれ、H.264<sup>7)</sup> で採用されている。

#### 4-5-5 マルチ参照フレーム動き補償

MPEG-2 の P フレームや B フレームでは、その参照フレームは過去の 1 枚あるいは過去と

未来から 1 枚ずつ、とそれぞれ定められていたが、H.264/AVC<sup>7)</sup>ではこの制限を取り除き、既に復号済みでバッファに蓄えられている複数のフレームから自由に予測ブロックを選ぶことが可能となっている。この技術をマルチ参照フレーム動き補償 (Multiple Reference Frame Motion Compensation) と呼ぶ。

#### ■参考文献

- 1) C. C. Cutler, "Differential Quantization of Communication Signals," U.S. patent 2605361, 1952.
- 2) J. B. O'Neal, "Predictive Quantizing System (Differential Pulse Code Modulation) for the Transmission of Television Signals," BSTJ, vol.45, no.5, pp.689-721, 1966.
- 3) A. Habibi, "Comparison of nth-Order DPCM Encoder With Linear Transformations and Block Quantization Techniques," IEEE Trans. on Communication Technology, vol.19, no.6, 1971.
- 4) M. Rabbani and P. W. Jones, "Digital Image Compression Techniques," SPIE Press, 1991.
- 5) ISO/IEC 10918-1, "Digital Compression and Coding of Continuous-tone Still Image," ITU-T Rec. T.81, 1994.
- 6) ISO/IEC 14495, "Lossless and near-lossless compression of continuous-tone still images," ITU-T Rec. T.87, 1999.
- 7) ISO/IEC 14496-10, "Advanced video coding," ITU-T Rec. H.264, 2003.
- 8) G. Sullivan et al., "Draft Text of H.264/AVC Fidelity Range Extensions Amendment," Doc. JVT-L047, 2004.
- 9) A. N. Netravali and J. D. Robbins, "Motion-compensated television coding: Part I," The Bell System Technical Journal, vol.58, no.3, pp. 631-670, 1979.
- 10) ISO/IEC 11172, "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s," 1993.
- 11) ISO/IEC 13818, "Generic coding of moving pictures and associated audio information," ITU-T Rec. H.262, 1996.
- 12) ISO/IEC 14496-2, "Coding of audio-visual objects," ITU-T Rec. H.263, 2000.
- 13) K. Lee, T. S. Kim, H.-J. Lee, "A survey of fast mode decision algorithms for inter-prediction and their applications to high efficiency video coding," IEEE Trans. on Consumer Electronics, vol.58, no.4, pp.1375-1383, 2012.
- 14) T. Koga, K. Iinuma, A. Hirano, Y. Iijima and T.Ishiguro, "Motion-compensated interframe coding for video conferencing," Proceedings of NTC 1981, pp.G5.3.1-G5.3.5, 1981.
- 15) M. Bierling, "Displacement estimation by hierarchical block matching," Proc. of Visual Communications and Image Processing, vol.1001, no.2, pp.942-951, 1988.
- 16) Y.-W. Huand, C.-Y. Chen, C.-H. Tsai, C.-F. Shen, L.-G. Chen, "Survey on Block Matching Motion Estimation Algorithms and Architectures with New Results," Journal of VLSI Signal Processing Systems, vol.42, no.3, pp.297-320, 2006.
- 17) M. Hoetter, "Differential estimation of the global motion parameters zoom and pan," Signal Processing, vol.16, no.3, pp.249-265, 1989.
- 18) K. Kamikura, H. Watanabe, H. Jozawa, H. Kotera, and S. Ichinose, "Global Brightness-Variation Compensation for Video Coding," IEEE Trans. on Circuits and Systems for Video Coding, vol.8, no.8, pp.988-1000, 1998.
- 19) T. Orchard and G. J. Sullivan, "Overlapped Block Motion Compensation: An Estimation-Theoretic Approach," IEEE Image Processing, vol.3, no.5, pp.693-699, 1994.