

スケーラブル符号化におけるドリフト誤差低減のための動き補償

岩橋 政宏[†] 神林 紀嘉[†] 貴家 仁志^{††}

A Motion Compensation of Scalable Decoder for Drift Reduction

Masahiro IWAHASHI[†], Noriyoshi KAMBAYASHI[†], and Hitoshi KIYA^{††}

あらまし デジタル動画像データの通信や蓄積において、単にデータ量を圧縮するのみならず、圧縮されたデータの一部から縮小画像を復号できるスケーラブル符号化が最近注目を集めている。スケーラブル符号化による復号画像には、一般に、量子化誤差、解像度変換誤差、ドリフト誤差などの画質劣化が含まれている。なかでもエンコーダ・デコーダ間での動き補償による予測信号の不一致により発生するドリフト誤差は、予測処理を繰り返すほど累積するため他の画質劣化に比べ顕著である。そこで本論文では、ドリフト誤差を低減できる新しい動き補償を提案する。そのためにまず、予測信号の生成過程をマルチレート信号処理に帰着させ、動き補償をフィルタ処理としてとらえる。次に、ドリフト誤差が零となるときの動き補償フィルタの理想周波数特性を明らかにし、これを近似する動き補償フィルタの設計法を与える。最後に、設計された動き補償フィルタをスケーラブル符号化に組み込み、提案法のドリフト低減効果を確認する。自然画像に対する実験の結果、従来法に比べ約2 dB以上の画質改善効果が確認された。

キーワード 符号化、動き補償、マルチレート、動画像、スケーラブル、画質劣化

1. まえがき

デジタル信号処理技術の発展とその応用分野の拡大に伴い、動画像データのデジタル化がマルチメディア通信や蓄積メディアなどのあらゆる分野で要求されている。しかし動画像データの量は文書や音声に比べ膨大であるため、これを圧縮する技術の研究と開発が必須である。なかでも離散コサイン変換(DCT)と動き補償を要素とするハイブリッド符号化[1],[2]は、ISO/IEC JTC1やITU-TによるMPEG(moving picture experts group)国際標準[3]としても採用されており、エンコーダ・デコーダのLSI化やシステムの開発が、近年、急速に進展している。

一方、画像データの解像度(画像データ1枚あたりの画素数)は各種のアプリケーションに応じて多くの種類が混在している。例えば標準的な画像データとそれよりも高い解像度をもつ高精細画像や、デスクトップコンピュータ上へ表示するための任意の解像度の画

像、あるいはCIF(common intermediate format)およびQCIF(quarter CIF)などの各種画像フォーマットである[3]。

こうした解像度の多様化に伴い、解像度変換機能をもつハイブリッド符号化すなわちスケーラブル符号化が最近盛んに研究されている[4]~[9]。これにより、エンコーダから出力されるビットストリームの一部をデコードすることで低解像度の(縮小された)画像データが再生できる。しかしこの場合、縮小画像にドリフト誤差(drift error)と呼ばれる誤差が重畳してしまう[4]~[9]。これは、動き補償における予測信号が、縮小の結果、エンコーダとデコーダの間で一致しなくなるために発生する。また、時間方向予測を繰り返すほど累積してしまう[5]。

ドリフト誤差の低減方法に関する研究は、最近、いくつか報告されている[8],[9]。R. Mokryらは、デコーダにおける4点離散コサイン逆変換(IDCT)の前に8点IDCTと8点DCTを付加し、縮小前の解像度で動き補償を行うことでドリフト誤差を除去できることを報告している[8]。しかし変換回数の増加に伴いハードウェア量が増加し、特に動画像処理のように高速化が要求される環境では望ましくない。一方、A.W. Johnsonらは、デコーダ内にて動き補償された信号に

[†] 長岡技術科学大学工学部電気系学科, 長岡市
Faculty of Engineering, Nagaoka University of Technology,
Nagaoka-shi, 940-21 Japan

^{††} 東京都立大学工学部電子情報工学科, 東京都
Faculty of Engineering, Tokyo Metropolitan University,
Hachioji-shi, 192-03 Japan

ブロックに基づくフィルタ処理を行うことでドリフト誤差を低減できることを報告している [9]。しかし、その処理は複数のブロックに及ぶため、対象となる画素とは空間的に非常に離れたほとんど相関のない画素がフィルタ処理に使われてしまうという問題がある。

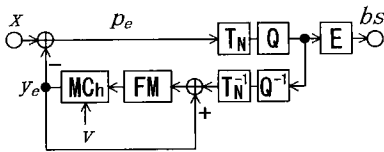
そこで本論文では、デコーダの動き補償を修正しデコーダの予測信号をエンコーダの予測信号に近づけることでドリフト誤差を低減する方法を提案し、その効果を確認する。2. ではドリフト誤差を定義し、その原因となる動き補償についてまとめる。3. ではマルチレート信号処理の表現を用いて動き補償をフィルタ処理に帰着させ、フィルタの満たすべき条件を導く。4. ではドリフト誤差を零にするフィルタの理想周波数特性を明らかにし、これを近似するフィルタの設計方法を与える。5. では提案法のドリフト低減効果を確認する。本論文の結論を 6. に述べる。

2. スケーラブル符号化

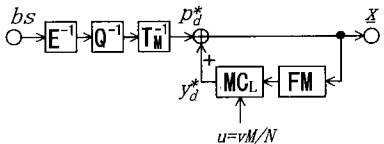
2. ではスケーラブル符号化にて発生するドリフト誤差を定義し、ドリフト誤差の発生原因となる従来の動き補償を z 領域で記述する。

2.1 ドリフト誤差 [3]~[9]

図 1 を用いてスケーラブル符号化を簡単に説明する。エンコーダは基本的に MPEG 国際標準 [3] と同一である。まず、空間位置 n 、時刻 t における入力信号を



(a) Encoder



(b) Scalable decoder
(N:M scaling)

- MC : motion compensation
- FM : frame memory
- E : entropy coding
- Q : quantization
- T_N : N point DCT

図 1 動画のスケラブル符号化

Fig.1 The scalable coding of video data.

$x(n, t)$ 、予測信号を $y_e(n, t)$ とすると、両者の差分である予測誤差信号、

$$p_e(n, t) = x(n, t) - y_e(n, t) \quad (1)$$

が N 点の DCT により変換される。なお、空間位置 n は画像の場合、縦方向と横方向の 2 変数で表される。しかし、本論文では分離型の信号処理を対象とするため、1 変数を用いて説明する。また、予測信号 $y_e(n, t)$ は、フレームメモリに蓄積されている信号 $x(n, t-1)$ をある大きさのブロックごとに動き補償することで得られる。詳細は 2.2 で説明する。続いて、量子化およびエントロピ符号化によりデータ量が圧縮される。このとき一般に量子化誤差が発生する。しかし、本論文において解析の対象とするドリフト誤差は量子化とは関係なく発生する。またエントロピ符号化もドリフト誤差には寄与しない。従って以下、量子化およびエントロピ符号化を除外して議論を進める。

デコーダでは、エンコーダから送られた N 点の DCT 係数のうち、低域側の M 点のみが、 M 点の IDCT により逆変換される。但し、 $M < N$ とする。また、MPEG アルゴリズムでは $N = 8$ である。その結果、 M/N 倍に解像度変換 (縮小) された予測誤差信号、

$$p_d(n, t) = x^*(n, t) - y_e^*(n, t) \quad (2)$$

が得られる。なお、本論文では記号「*」により、信号が解像度変換されていることを表す。次にこの予測誤差信号にデコーダにおいて生成された予測信号 $y_d^*(n, t)$ が加算される。従って、出力信号は、

$$\underline{x}(n, t) = p_d^*(n, t) + y_d^*(n, t) \quad (3)$$

となる。 $y_d^*(n, t)$ の生成方法は 2.3 で説明する。

以上の結果をまとめると、式 (1), (2), (3) より、

$$\underline{x}(n, t) = x^*(n, t) + d(n, t) \quad (4)$$

但し、

$$d(n, t) = -y_e^*(n, t) + y_d^*(n, t) \quad (5)$$

が得られる。上式から、 $y_e^*(n, t) \neq y_d^*(n, t)$ のとき、復号信号 $\underline{x}(n, t)$ は、縮小された入力信号 $x^*(n, t)$ のほかに不必要な信号 $d(n, t)$ を含むことがわかる。本論文では、この不必要な信号 $d(n, t)$ をドリフト誤差と定義する。そして、デコーダの予測信号 $y_d^*(n, t)$ を

解像度変換されたエンコーダの予測信号 $y_e^*(n, t)$ に近づけることで、ドリフト誤差を低減する方法について考察する。

なお、式 (4) における解像度変換された信号 $x^*(n, t)$ には、ドリフト誤差や量子化誤差とは独立な解像度変換に伴う誤差が含まれている。以下、この誤差を解像度変換誤差と呼ぶ。その発生量は M/N が 1 に近づくほど減少し、解像度変換比 M/N が一定の場合は N の値が大きいほど少なくなる [10]。特に、 $N = 8$ 、 $M = 4$ の場合、自然画像の例では視覚的にほとんど認知できないことが知られている [11]。

2.2 エンコーダにおける動き補償 [1], [3]

ここで、式 (1) の右辺における予測信号 $y_e(n, t)$ の生成方法について説明する。この信号は、フレームメモリに蓄えられている過去の信号 $x(n, t - 1)$ を動き補償することで得られる。従って、動き補償を記号 $MC_h[\]$ で表すと、予測信号 $y_e(n, t)$ は、

$$y_e(n, t) = MC_h[x(n, t - 1)] \tag{6}$$

と表される。MPEG で用いられる符号化アルゴリズム [1], [3] の場合、動き補償は 1/2 画素単位で行われる。これは、動く画素数および方向を v とすると、 v の値が、

$$v = 0, \pm 0.5, \pm 1, \pm 1.5, \pm 2, \dots \tag{7}$$

のように、符号を伴う 0.5 の整数倍として表現されることを意味する。なお、この v は動きベクトルと呼ばれ、式 (1) の左辺における予測誤差信号 $p_e(n, t)$ の領域が狭くなり、データ圧縮に寄与するよう、動き推定アルゴリズム (motion estimation algorithm) により決定される [1], [3]。ところで、式 (6) は、 $F_{h,p}(z)$ を後に式 (10) で定義するフィルタの伝達関数を用いて、

$$Y_e(z, t) = F_{h,p}(z)z^v X(z, t - 1), \tag{8}$$

$$p = (2v) \bmod 2$$

と z 領域にて表すことができる。但し、任意の信号 $x(n, t)$ の時刻 t における z 変換を、

$$X(z, t) = \sum_{n=-\infty}^{\infty} x(n, t)z^{-n} \tag{9}$$

とする。また、記号「 $a \bmod b$ 」は a を b で割ったときの余りを表す。式 (8) は v が 0.5 の偶数倍のとき $p = 0$ となり、信号 $X(z, t - 1)$ はフィルタ $F_{h,0}(z)$ に

より処理された後、 v 画素だけシフトされることを意味する。一方、 v が 0.5 の奇数倍のとき $p = 1$ となり、フィルタ $F_{h,1}(z)$ が処理に用いられる。なお、MPEG で用いられる符号化アルゴリズムでは、二つのフィルタ $F_{h,p}(z)$ 、($p = 0, 1$) は、

$$\begin{bmatrix} F_{h,0}(z) \\ F_{h,1}(z)z^{-1/2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \tag{10}$$

となる [1], [3]。以上で、1/2 画素精度の動き補償が、 z 領域にて二つのフィルタを介して記述された。

2.3 デコーダにおける動き補償 [3]

次に、式 (3) における信号 $y_d^*(n, t)$ の生成方法について説明する。これは、デコーダのフレームメモリに蓄えられている信号 $\underline{x}(n, t - 1)$ により、式 (6) と同様、

$$y_d^*(n, t) = MC_L[\underline{x}(n, t - 1)] \tag{11}$$

と表される。なお、上式の予測信号 $y_d^*(n, t)$ は解像度変換された信号を動き補償したものである。一方、式 (5) における信号 $y_e^*(n, t)$ は動き補償された信号を解像度変換したものである。このように、二つの信号 $y_d^*(n, t)$ と $y_e^*(n, t)$ はそれぞれの生成過程が異なるため、必ずしも一致するとは限らない。従って、一般的にドリフト誤差が発生する。

スケーラブル符号化では、2.1 で説明したように、デコーダにおいて信号の解像度が M/N 倍に縮小されている。そのため、動き補償の精度もまたエンコーダの動き補償の精度の M/N 倍になる。例えば、MPEG アルゴリズムの場合、 $N = 8$ の DCT が使われている。従って、デコーダにて $M = 4$ の IDCT が使われるならばそのときの動き補償の精度は 1/4 画素、 $M = 2$ ならば 1/8 画素となる。但し、この精度は原画像の解像度を基準としている。動き補償の精度を $1/L$ 画素とすると、

$$L = 2N/M \tag{12}$$

なる関係がある。このとき、動きベクトルの値も M/N 倍に縮小されるため、デコーダにおける動きベクトルを u とすると、エンコーダの動きベクトル v と u には、

$$u = 2v/L \tag{13}$$

なる関係がある。式 (11) は式 (8) と同様に、

$$Y_d^*(z, t) = F_{L,p}(z)z^{2v/L} \underline{X}(z, t - 1), \tag{14}$$

$$p = (Lv) \bmod L$$

と z 領域にて表現できる。なお、従来法である文献[3]では、 $L = 4$ の場合、四つのフィルタ $F_{4,p}(z)$, ($p = 0, 1, 2, 3$) は、

$$\begin{bmatrix} F_{4,0}(z) \\ F_{4,1}(z)z^{-1/4} \\ F_{4,2}(z)z^{-1/2} \\ F_{4,3}(z)z^{-3/4} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 4 & 0 \\ 3 & 1 \\ 2 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \quad (15)$$

で与えられる。また、 $L = 8$ の場合は、

$$F_{8,p}(z)z^{-p/8} = \frac{1}{8}(8 - p + pz^{-1}), \quad p = 0, 1, \dots, 7 \quad (16)$$

が用いられている。以上で、 $1/L$ 画素精度の動き補償が、 L 個のフィルタにより z 領域で記述された。

3. 動き補償フィルタ

2. では動き補償が複数のフィルタにより表現された。3. ではそれぞれのフィルタの満たすべき条件をまとめる。次に、マルチレート信号処理を用いてこれら複数のフィルタを一つのフィルタで表現する。以下、このフィルタを動き補償フィルタと呼ぶ。本論文の目的はドリフト誤差を零とするデコーダにおける動き補償フィルタを設計することである。

3.1 フィルタの満たすべき条件

まず、エンコーダにおける二つのフィルタ $F_{h,p}(z)$, ($p = 0, 1$) の満たすべき条件を示す。第1に、 $p = 0$ となる v が0.5の偶数倍の場合、動き補償により単に v 画素だけシフトされた信号が出力されなければならない。すなわち、

$$F_{h,0}(z) = 1 \quad (17)$$

となる必要がある。一方、 $p = 1$ となる v が0.5の奇数倍の場合、伝達関数 $F_{h,p}(z)z^v$ が z の整数乗の線形結合でなくてはならない。従って、 $N1, N2$ を任意の整数として、

$$F_{h,1}(z)z^{-1/2} = \sum_{n=N1}^{N2} f_{h,1}(n)z^{-n} \quad (18)$$

となる必要がある。更に、動きベクトル v の値に依存して利得が変化してはならない。従って、

$$F_{h,1}(1) = 1 \quad (19)$$

なる条件が課せられる。

次に、デコーダにおけるフィルタ $F_{L,p}(z)$, ($p = 0, 1, \dots, L-1$) の満たすべき条件を示す。これらについても、式(17)と同様、 $p = 0$ のとき、

$$F_{L,0}(e^{j\omega}) = 1 \quad (20)$$

である必要がある。また、式(18)と同様、 $p \neq 0$ のとき、

$$F_{L,p}(z)z^{-p/L} = \sum_{n=N1}^{N2} f_{L,p}(n)z^{-n}, \quad p = 1, 2, \dots, L-1 \quad (21)$$

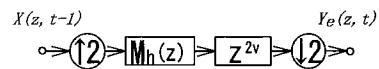
であり、また、式(19)と同様、

$$F_{L,p}(1) = 1, \quad p = 1, 2, \dots, L-1 \quad (22)$$

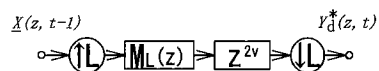
が満たされる必要がある。

3.2 マルチレート表現

これまでの議論の結果、動き補償が複数のフィルタで記述され、それぞれの満たすべき条件が明らかになった。しかし、いくつもあるフィルタのうちのいずれを用いるかは、動きベクトル v の値に依存して決定されるため、 v の値が既知である必要がある。そこで v の値があらかじめ決定されていなくても議論ができるように、まず、式(8)の関係を図2(a)に示すマルチレート信号処理の入出力関係に帰着させる。これにより、二つのフィルタ $F_{h,p}(z)$, ($p = 0, 1$) を一つの動き補償フィルタ $M_h(z)$ により表現する。次に、式(14)の入出力関係を図2(b)のマルチレート信号処理に帰着させ、 L 個のフィルタ $F_{L,p}(z)$, ($p = 0, 1, \dots, L-1$) を一つの動き補償フィルタ $M_L(z)$ により表現する。このことにより、最終的に本論文では、ドリフト誤差の低減問題を一つの動き補償フィルタ $M_L(z)$ の設計問題に帰着させる。



(a) 1/2 pel precision



(b) 1/L pel precision

図2 動き補償と等価なマルチレート信号処理

Fig.2 A multirate signal processing equivalent to the motion compensations.

図 2(a) の入出力関係は z 領域において、

$$Y_e(z, t) = \frac{1}{2} \sum_{k=0}^1 M_h(z^{1/2} e^{j\pi k}) \cdot (z^{1/2} e^{j\pi k})^{2v} X(z, t-1) \quad (23)$$

となる。従って、上式と式 (8) が等しいことから、フィルタ $F_{h,p}(z)$, ($p=0, 1$) と $M_h(z)$ の関係がわかり、

$$F_{h,p}(z) = \frac{1}{2} \sum_{k=0}^1 M_h(z^{1/2} e^{j\pi k}) (e^{j\pi k})^p, \quad p=0, 1 \quad (24)$$

または、

$$M_h(z) = \sum_{p=0}^1 F_{h,p}(z^2) \quad (25)$$

が得られる。なお、 $F_{h,p}(z)$, ($p=0, 1$) が式 (10) で与えられるとき、 $M_h(z)$ は、

$$M_h(z) = \frac{1}{2} (z + 2 + z^{-1}) \quad (26)$$

となる。

一方、式 (14) の入出力関係を、式 (23) と同様に、図 2(b) のマルチレート信号処理としてとらえると、

$$Y_d^*(z, t) = \frac{1}{L} \sum_{k=0}^{L-1} M_L(z^{1/L} e^{j2\pi k/L}) \cdot (z^{1/L} e^{j2\pi k/L})^{2v} X(z, t-1) \quad (27)$$

となる。このとき、 $F_{L,p}(z)$ と $M_L(z)$ の関係は、式 (14), (27) より、

$$F_{L,p}(z) = \frac{1}{L} \sum_{k=0}^{L-1} M_L(z^{1/L} e^{j2\pi k/L}) (e^{j2\pi k/L})^p, \quad p=0, 1, \dots, L-1 \quad (28)$$

または、

$$M_L(z) = \sum_{p=0}^{L-1} F_{L,p}(z^L) \quad (29)$$

となる。また、従来法では、式 (29) より、

$$M_4(z) = (z^3 + 2z^2 + 3z + 4 + 3z^{-1} + 2z^{-2} + z^{-3})/4 \quad (30)$$

となる。これは、式 (26) を用いると、

$$M_4(z) = M_h(z)M_h(z^2) \quad (31)$$

と表現できる。一方、 $L=8$ の場合、式 (16) より、

$$M_8(z) = M_h(z)M_h(z^2)M_h(z^4) \quad (32)$$

となる。以上で、動き補償を表す複数のフィルタが、エンコーダとデコーダそれぞれにおいて一つのフィルタで表現された。

4. 動き補償フィルタの設計

4. では、まず、ドリフト誤差を周波数領域で表現する。次に、デコーダの動き補償フィルタ $M_L(z)$ に対するドリフト誤差を零にするための理想周波数特性を導く。最後に、3.1 でまとめた条件のもとでこの理想特性を近似するフィルタ $M_L(z)$ の一設計法を述べる。

4.1 ドリフト誤差の周波数領域表現

ドリフト誤差は、式 (5) の z 変換を用いることで、周波数領域にて、

$$D(e^{j\omega}, t) = -Y_e^*(e^{j\omega}, t) + Y_d^*(e^{j\omega}, t) \quad (33)$$

と表される。また、エンコーダおよびデコーダにおける予測信号の周波数特性は、式 (23), (27) よりそれぞれ、

$$Y_e(e^{j\omega}, t) = \frac{1}{2} \sum_{k=0}^1 M_h(e^{j(\omega/2+\pi k)}) \cdot (e^{j(\omega/2+\pi k)})^{2v} X(e^{j\omega}, t-1) \quad (34)$$

および、

$$Y_d^*(e^{j\omega}, t) = \frac{1}{L} \sum_{k=0}^{L-1} M_L(e^{j(\omega+2\pi k)/L}) \cdot (e^{j(\omega+2\pi k)/L})^{2v} X(e^{j\omega}, t-1) \quad (35)$$

と表される。ところで、式 (35) における信号 $\underline{x}(n, t-1)$ は初期状態において信号 $x^*(n, t-1)$ に等しい。なぜなら、最初のフレームを符号化する際にそれ以前の信号が存在しないため、予測信号の値が零となるからである。従って、式 (4), (5) からわかるように、

$$\underline{X}(e^{j\omega}, 0) = X^*(e^{j\omega}, 0) \quad (36)$$

が成り立つ。更に、 $Y_e^*(e^{j\omega}, t)$ と $Y_e(e^{j\omega}, t)$ の関係、および、 $X(e^{j\omega}, t)$ と $X^*(e^{j\omega}, t)$ の関係がわかれば、以下に示すように、ドリフト誤差を入力信号 $X(e^{j\omega}, t)$ の関数として表すことができる。そこでこれらの関係

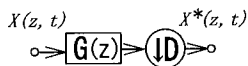


図3 解像度変換の一手法
Fig. 3 A resolution conversion method.

を調べるため、次に、解像度変換による周波数特性の変化についてまとめる。

例として、図3に示すような間引き処理により信号の解像度を $N:M$ に変換する場合を考える。この方法では解像度変換誤差を回避するため、サンプル点を間引く前に低域通過フィルタ $G(z)$ による処理が行われる。このときの入出力関係は、

$$D = N/M = L/2 \quad (37)$$

として、

$$X^*(e^{j\omega}, t) = \frac{1}{D} \sum_{k=0}^{L-1} G(e^{j(\omega+2\pi k)/D}) \cdot X(e^{j(\omega+2\pi k)/D}, t) \quad (38)$$

と表現される。なお、本論文ではサンプル値が実数である離散時間信号を対象とする。このとき、信号の周波数特性は 2π ごとの周期をもち、なおかつ $\omega = 0$ に対して対称となる。そのため以下、周波数特性が独立である $0 \leq \omega \leq \pi$ に ω の範囲を限定して議論を進める。図3の処理において、 $\omega_p = \pi/D$ として、 $0 \leq \omega < \omega_p$ の範囲においてフィルタ $G(z)$ の周波数振幅特性が一定値 α 、かつ位相特性が零、更に、 $\omega_p < \omega \leq \pi$ の範囲において振幅特性が0であれば、解像度変換誤差を含まない縮小信号が得られる。また、自然画像における解像度変換誤差は、 $L = 4, 8$ 程度ではドリフト誤差に比べて十分小さく実用上は無視できる [10], [11]。そこで本論文では、このような処理で得られる信号を解像度変換された結果とする。このとき、式 (38) より、

$$X^*(e^{j\omega}, t) = \frac{1}{D} X(e^{j\omega/D}, t) \quad (39)$$

となる。また、予測信号についても同様に、

$$Y_e^*(e^{j\omega}, t) = \frac{1}{D} Y_e(e^{j\omega/D}, t) \quad (40)$$

なる関係が得られる。但し、直流成分だけの信号に対する利得が1となるよう、フィルタ $G(z)$ の通過域利得 α を1とした。これは、画像信号の場合、直流成分のエネルギーが非常に大きいためである。

以上の結果を用いると、最初から2番目のフレームの符号化時において、式 (34), (37), (40) より、

$$Y_e^*(e^{j\omega}, 1) = \frac{1}{L} \sum_{k=0}^1 M_h(e^{j(\omega/L+\pi k)}) \cdot (e^{j(\omega/L+\pi k)})^{2v} X(e^{j2\omega/L}, 0) \quad (41)$$

となる。また、式 (35), (36), (37), (39) より、

$$Y_d^*(e^{j\omega}, 1) = \frac{2}{L^2} \sum_{k=0}^{L-1} M_L(e^{j(\omega+2\pi k)/L}) \cdot (e^{j(\omega+2\pi k)/L})^{2v} X(e^{j2\omega/L}, 0) \quad (42)$$

となる。更に、式 (41), (42) を式 (33) に代入すると、

$$D(e^{j\omega}, 1) = D_0(e^{j\omega}) X(e^{j2\omega/L}, 0) \quad (43)$$

が得られる。なお、上式において、

$$D_0(e^{j\omega}) = \frac{2}{L^2} (e^{j\omega/L})^{2v} \cdot \begin{bmatrix} 1 & (-1)^{2v} & 1 \end{bmatrix} \begin{bmatrix} D_1(e^{j\omega}) \\ D_2(e^{j\omega}) \\ D_3(e^{j\omega}) \end{bmatrix} \quad (44)$$

但し、

$$\begin{aligned} D_1(e^{j\omega}) &= M_L(e^{j\omega/L}) - \frac{L}{2} M_h(e^{j\omega/L}) \\ D_2(e^{j\omega}) &= M_L(e^{j(\omega/L+\pi)}) - \frac{L}{2} M_h(e^{j(\omega/L+\pi)}) \\ D_3(e^{j\omega}) &= \sum_{\substack{k=1 \\ k \neq L/2}}^{L-1} M_L(e^{j(\omega+2\pi k)/L}) (e^{j2\pi k/L})^{2v} \end{aligned} \quad (45)$$

とする。以上で、ドリフト誤差が入力信号 $X(e^{j\omega}, t)$ と関数 $D_0(e^{j\omega})$ の積として周波数領域にて表現された。

4.2 動き補償フィルタの理想周波数特性

式 (43) において、関数 $D_0(e^{j\omega})$ が零となれば入力信号に関係なくドリフト誤差が零となる。更に、式 (45) の $D_1(e^{j\omega})$, $D_2(e^{j\omega})$, $D_3(e^{j\omega})$ がすべて零であれば、式 (44) からわかるように、動きベクトルの値に依存せずにドリフト誤差が零となる。このとき式 (45) より、

$$\begin{aligned} M_L(e^{j\omega}) &= \frac{L}{2} M_h(e^{j\omega}), \quad 0 \leq \omega < \frac{\pi}{L} \\ M_L(e^{j\omega}) &= \frac{L}{2} M_h(e^{j\omega}), \quad \left(\pi - \frac{\pi}{L}\right) < \omega \leq \pi \\ M_L(e^{j\omega}) &= 0, \quad \frac{\pi}{L} < \omega < \left(\pi - \frac{\pi}{L}\right) \end{aligned} \quad (46)$$

となる。そこで本論文では、式 (46) をデコーダの動き補償フィルタ $M_L(e^{j\omega})$ に対するドリフト誤差を零とするための、理想周波数特性と考える。

4.3 動き補償フィルタの設計法

本論文のテーマであるドリフト誤差低減のためのデコーダにおける動き補償は、これまでの議論の結果、式 (20), (21), (22) の条件のもとで、式 (46) の理想周波数特性を近似する動き補償フィルタ $M_L(z)$ を設計する問題に帰着された。以下、提案する設計方法を示す。

まず、式 (46) の理想周波数特性を変形する。この周波数特性は、図 4 に示すように、理想低域通過フィルタ $H_L(z)$ の周波数特性とエンコーダの動き補償フィルタ $M_h(z)$ の周波数特性の積として表すことができる。ここで、フィルタ $H_L(z)$ は、

$$\begin{aligned} H_L(e^{j\omega}) &= \frac{L}{2}, \quad 0 \leq \omega < \frac{\pi}{L} \\ H_L(e^{j\omega}) &= \frac{L}{2}, \quad \left(\pi - \frac{\pi}{L}\right) < \omega \leq \pi \\ H_L(e^{j\omega}) &= 0, \quad \frac{\pi}{L} < \omega < \left(\pi - \frac{\pi}{L}\right) \end{aligned} \quad (47)$$

を満たすものとする。更にこのフィルタ $H_L(z)$ は、以下の特性、

$$\begin{aligned} H(e^{j\omega}) &= 2, \quad 0 \leq \omega < \frac{\pi}{2} \\ H(e^{j\omega}) &= 0, \quad \frac{\pi}{2} < \omega \leq \pi \end{aligned} \quad (48)$$

をもつフィルタ $H(z)$ により、

$$M_L(z) = \prod_{q=1}^{Q-1} H(z^{2^q}), \quad Q = \log_2 L \quad (49)$$

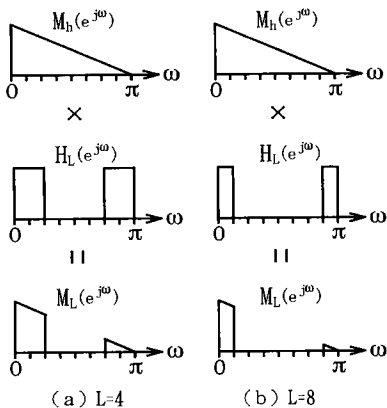


図 4 動き補償フィルタの周波数特性

Fig. 4 Frequency characteristics of the motion compensation filters.

と表すことができる。以上の結果、式 (46) の理想周波数特性は、

$$M_L(z) = M_h(z) \prod_{q=1}^{Q-1} H(z^{2^q}), \quad Q = \log_2 L \quad (50)$$

と表現できる。従って、式 (48) の周波数特性を近似するフィルタ $H(z)$ を設計して上式に代入することで、ドリフト誤差低減のための動き補償フィルタ $M_L(z)$ が設計できることになる。ところが、このフィルタは更に、条件式 (20), (21), (22) を満たさなければならない。

そこで提案法では、式 (50) におけるフィルタ $H(z)$ として、伝達関数が、

$$H(z) = 1 + F(z^2) \quad (51)$$

但し、 N を正の整数として、

$$F(z) = \sum_{n=1}^N f(n)(z^{(2n-1)/2} + z^{-(2n-1)/2}) \quad (52)$$

で表されるフィルタを用いる。但し、上式におけるフィルタ係数 $f(n)$ は、

$$|F(e^{j\omega})| = 1 \quad (53)$$

を近似するように計算する。なお、文献 [13] の方法で計算されたフィルタ係数 $f(n)$ を表 1 に示す。これにより、動き補償フィルタに課せられた三つの条件式 (20), (21), (22) がすべて満たされる。例えば、 $L=4$ の場合、式 (26), (28), (50), (51) により、フィルタ $F_{4,p}(z)$ ($p=0, 1, 2, 3$) が式 (52) のフィルタ $F(z)$ により、

$$\begin{aligned} & \begin{bmatrix} F_{4,0}(z) \\ F_{4,1}(z)z^{-1/4} \\ F_{4,2}(z)z^{-1/2} \\ F_{4,3}(z)z^{-3/4} \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \\ z^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ F(z)z^{-1/2} \end{bmatrix} \end{aligned} \quad (54)$$

と表される。上式に式 (51), (52), (53) を代入することで、三つの条件が満たされることが確認できる。

以上で、式 (20), (21), (22) の条件のもとで、式 (46) の理想特性を近似する動き補償フィルタ $M_L(z)$ の設

表1 フィルタ係数 $f(n)$
Table 1 Filter coefficients $f(n)$.

	$N = 1$	$N = 2$	$N = 3$	$N = 4$
$f(0)$	0.5000	0.2813	0.2930	0.2991
$f(1)$		-0.0313	-0.0488	-0.0598
$f(2)$			0.0059	0.0120
$f(3)$				-0.0012

計方法が示された。また、ドリフト誤差低減の度合は、式 (53) の近似精度に比例する。すなわち、ドリフト誤差を低減するには急しゅんな特性をもつフィルタ $F(z)$ が必要である。ところが、式 (15) で示された従来法は、提案法の $N = 1$ の場合に等しく、 $F(z)$ の特性は急しゅんではない。なお、このことは $f(1) = 0.5$ を式 (52) に代入し、得られた $F(z)$ を式 (54) に代入した結果が式 (15) と一致することで確認できる。このように、本論文では動き補償フィルタに対する理想特性および満たすべき条件を明らかにした上で、従来法で使われていた低次のフィルタを急しゅんな特性をもつ高次のフィルタで置き換えることでドリフト誤差を減少させている。

5. シミュレーション

5. では提案法によるドリフト誤差の低減効果をシミュレーションにより確認する。なお、入力信号としては AR(1) モデルおよび自然画像を用いる。

5.1 AR(1) モデル

AR(1) モデルは、画像データの周波数スペクトルを近似するモデルとして広く利用されており、相関係数を ρ とすると、その周波数振幅特性は、

$$|X(e^{j\omega}, t)| = \frac{1}{\sqrt{\left(\frac{2 \sin(\omega/2)}{1-\rho}\right)^2 \rho + 1}} \quad (55)$$

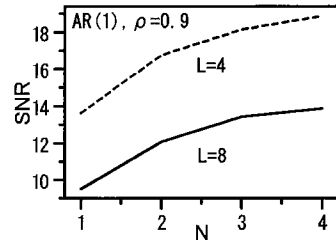
で与えられる。この周波数特性を式 (43) に代入することで、AR(1) モデルを入力信号とした場合のドリフト誤差が計算できる。ここで、ドリフト誤差の発生量を評価するため SNR (signal to noise ratio) を定義する。これは、ドリフト誤差のエネルギー密度、

$$E_N = \frac{1}{\pi} \int_0^\pi D(e^{j\omega}, 1) D(e^{-j\omega}, 1) d\omega \quad (56)$$

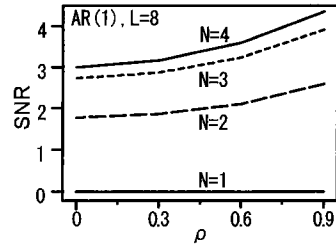
および、入力信号のエネルギー密度、

$$E_S = \frac{1}{\pi} \int_0^\pi X(e^{j\omega}, 0) X(e^{-j\omega}, 0) d\omega \quad (57)$$

により、



(a) $\rho=0.9$



(b) $L=8$

図5 ドリフト誤差の低減効果

Fig.5 Effectiveness of the drift reduction.

$$SNR = 10(\log_{10} E_S - \log_{10} E_N) \quad (58)$$

と定義する。図5に表1のフィルタ係数を用いた場合のSNRの値を示す。但し、4.1にも述べたように、解像度変換誤差は無視できるほど十分小さいと仮定している。同図(a)において、整数 N は式 (52) におけるフィルタ次数を決定するパラメータであり、整数 N の値が大きいほど式 (53) の近似精度が高い。また、 $N = 1$ のとき提案法と従来法は等しくなる。同図より、 $\rho = 0.9$ の場合、提案法による2dB以上のドリフト低減効果が確認できる。また、図5(b)には $N = 1$ を基準としたときのSNRを示す。同図より、 ρ が1に近いほど、すなわち入力信号のスペクトルが低域に偏っているほど、ドリフト誤差が小さくなることがわかる。なお、直流信号を入力した場合、ドリフト誤差は零である。これは、 $\omega = 0$ を式 (44) に代入することで確かめられる。

5.2 自然画像

次に、提案法を自然画像の縦と横、それぞれの方向に適用し、ドリフト低減効果を調べる。まず、入力信号のスペクトルをFFT (Fast Fourier Transform) で計算し、式 (56), (57) に代入する。その結果、得られたSNRを図6に示す。同図より、自然画像では、従来法の次数を若干上げるだけで2dB以上の画質改善効果が得られることが確認できる。また、同図の $L = 4$

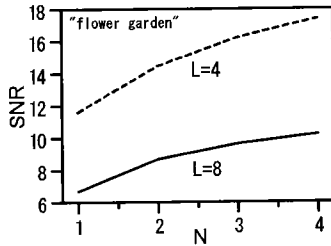


図6 ドリフト誤差の低減効果
Fig.6 Effectiveness of the drift reduction.



(a) conventional method (N=1)



(b) proposed method (N=3)

図7 再生されたインタレース動画像
Fig.7 Decoded interlace video image. (L = 4)

の場合において、 $N = 1$ および $N = 3$ のときの再生画像を図7に示す。同図より、従来法では全体としてぼやけた画質となっていること、および、この画質劣化が提案法により改善されていることが確認できる。

6. むすび

本論文では、スケーラブル符号化において再生画像

に重畳するドリフト誤差を低減するための、デコーダにおける新しい動き補償を提案した。まず、ドリフト誤差の発生原因を考察し、動き補償フィルタの満たすべき条件をまとめた。次に、ドリフト誤差を零とするための理想周波数特性を明らかにした。更に、デコーダの動き補償フィルタを本論文で提案した方法で設計し直すことで、ドリフト誤差の低減が可能となることを示した。最後に、従来法におけるフィルタのタップ数を若干上げるだけで、自然画像に対して2dB以上の画質改善効果が確認された。

今後はフィルタ係数の語長を低減し、ハードウェアの簡略化を検討する。

謝辞 本研究を進めるにあたり御協力頂いた、グラフィックス・コミュニケーション・ラボラトリーズ(株)の大山公一主幹研究員、亀山渉主任研究員の両氏に深謝致します。

文 献

- [1] D.L. Gall, "MPEG: A video compression standard for multimedia applications," Commun. ACM, vol.34, no.4, April 1991.
- [2] K.R. Rao and P. Yip, "Discrete Cosine Transform Algorithms, Advantages, Applications," Academic, New York, 1990. (安田 浩, 藤原洋訳, "画像符号化技術—DCTとその国際標準," オーム社, 1992).
- [3] "Test Model 4/draft," Test Model Editing Committee, ISO-IEC/JTC1/SC29/WG11, MPEG93/N225b, March 1993.
- [4] 花村 剛, 富永英義, "スケーラブルビデオ符号化—MPEGでの動向と将来像," テレビ技法, vol.17-57, pp.25-30, Oct. 1993.
- [5] M.R. Civanlar and A. Puri, "Issues in Efficient Frequency Scalable Coding," Picture Coding Symposium (PCS), Lausanne, 12.2, March 1993.
- [6] M. Iwahashi, "Motion Compensation Technique for 2:1 Scaled-Down Moving Pictures," Picture Coding Symposium (PCS), 18.14, Lausanne, March 1993.
- [7] J. Ngan, T. Arnord, T.K. Sikora, A.W. Tan, and Johnson, "Frequency scalability experiments for MPEG-2 standard," Asia-Pacific Conference on Commun., pp.298-301, Aug. 1993.
- [8] R. Mokry and D. Anastassiou, "Insight into the Problem of Error Drift in Frequency Scalability," ISO-IEC/JTC1/SC29/WG11, MPEG 92/466, Sept. 1992.
- [9] A.W. Johnson and J. Princen, "Drift minimum in frequency scalable coders using block based filtering," IEEE Visual Signal Processing and Communications, Melbourne, pp.231-pp.234, Sept. 1993.
- [10] M. Iwahashi, K. Ohyama, and N. Kambayashi, "Estimation of aliasing error in layered coding system," IEEE Proc. ICASSP 49.7, April 1994.
- [11] P. Delogne, O. Poncin, B. Gaille, and L. Vandendrope,

"Compatible coding of digital interlaced HDTV," IEEE Proc. ISCAS '92, pp.196-199, 1992.

[12] M. Iwahashi, W. Kameyama, K. Ohyama, and N. Kambayashi, "Design of motion compensation filters of frequency scalable coding-drift reduction," IEEE Proc.ISCAS'94, pp.3.277, May 1994.

[13] O. Herrman, "On the approximation problem in nonrecursive digital filter design," IEEE Trans. Circuit Theory, vol.CT-18, pp.411-413, May 1971.

(平成7年5月11日受付, 7月24日再受付)



岩橋 政宏 (正員)

昭63都立大・工・電気卒。平2同大学院修士課程了。同年新日本製鐵(株)エレクトロニクス研究所入所。平3ジー・シー・テクノロジー(株)へ派遣。この間、デジタル動画像符号化の研究に従事。平5長岡技科大・工・電気系助手、現在に至る。

デジタル信号処理、特に画像の高効率符号化の研究に従事。画像電子学会、日本音響学会、IEEE各会員。



神林 紀嘉 (正員)

昭38信州大・工・通信卒。昭42東工大・工・電子物理助手。昭53長岡技科大・工・電気系助教授。現在、同教授。工博。電子回路、デジタル信号処理とその応用に関する研究に従事。著書「フィルタの理論と設計」、「基礎回路工学」(共著)など。



貴家 仁志 (正員)

昭55長岡技科大・工・電気電子システム卒。昭57同大学院修士課程了。同年都立大・工・電気工学科助手。現在、同大電子情報工学科助教授。工博。デジタル信号処理、特にマルチレート信号処理とその応用に関する研究に従事。著書「高速フーリエ変換とその応用」、「デジタル信号処理技術入門」など。画像電子学会、IEEE各会員。

エ変換とその応用」,「デジタル信号処理技術入門」など。画像電子学会、IEEE各会員。