

ショートノート

ハードウェア構成の簡単なスケーラブルデコーダ

岩橋 政宏<sup>†</sup> (正員)                      神林 紀嘉<sup>†</sup> (正員)  
 貴家 仁志<sup>††</sup> (正員)

A Scalable Decoder with Reduced Hardware Complexity

Masahiro IWAHASHI<sup>†</sup>, Noriyoshi KAMBAYASHI<sup>†</sup>, and  
 Hitoshi KIYA<sup>††</sup>, Members

<sup>†</sup> 長岡技術科学大学工学部電気系, 長岡市  
 Faculty of Engineering, Nagaoka University of Technology, Nagaoka-shi,  
 940-21 Japan

<sup>††</sup> 東京都立大学工学部電子情報工学科, 八王子市  
 Faculty of Engineering, Tokyo Metropolitan University, Hachioji-shi,  
 192-03 Japan

**あらまし** 離散コサイン変換 (DCT) と動き補償 (MC) を用いた動画のスケーラブル符号化において、ドリフト誤差を低減できるハードウェア構成の簡単な復号方法を提案する。また、従来法との比較を行い提案法の有効性を確認する。

**キーワード** 離散コサイン変換, 動き補償, スケーラブル, 符号化, 画質

1. まえがき

離散コサイン変換 (DCT) と動き補償 (MC) を用いたスケーラブル符号化は、ビットストリームの一部を復号すると低解像度の画像を復号でき、全部を復号するとフル解像度の画像を再生できるという機能を持ち、近年、盛んに研究されている [1]~[7]。しかし、復号された低解像度の画像にはドリフト誤差と呼ばれる画質劣化が発生する [3]~[5]。このドリフト誤差はエンコーダ-デコーダ間での MC 予測の不一致により発生し、これを低減するデコーダがいくつか提案された [6],[7]。

A.W. Johnson らの方法 [6] は画素をサブサンプルしてから MC を行うために処理が簡単になる。しかし、MC 予測の際に既にサブサンプルされた画素を近隣画素から補間するため、ドリフト誤差の低減には限界がある。一方、R. Mokry らの方法 [7] は画素をサブサンプルする前に MC を行うために補間の必要がなく、ドリフト誤差の低減には効果的である。しかし、ポスト処理として DCT と逆離散コサイン変換 (IDCT) を必要とするため処理が複雑になる欠点がある。

本論文では、R. Mokry らの方法におけるポスト処理の簡略化を行うことで、ドリフト誤差を効果的に低減できるハードウェア構成の簡単な復号方法を提案する。また、画質およびハードウェア面で従来法と比較する。

2. スケーラブル符号化

以下、エンコーダにおける符号化方法、および、従来報告された低解像度画像の復号方法についてまとめる。但し、分離型の 2 次元信号処理を対象とするため、1 次元信号処理により説明を行う。

2.1 エンコーダ [4],[5]

スケーラブル符号化における符号化処理を図 1 に示す。同図はシングルループ型と呼ばれ、動き補償 (MC) は原画像と同じ解像度でハーフペル精度で行われる [5],[8]。直交変換 (T) には 8 点の DCT が用いられ、DCT 係数 8 点のうち、低域の 4 点と高域の 4 点は別々に符号化される。従って、ビットストリームは二つのサブセット ( $bs_1, bs_2$ ) に分離され、低域側の DCT 係数を含むサブセット ( $bs_2$ ) のみを復号することで、低解像度の画像が再生されるという特徴をもつ。しかし、再生された低解像度の画像にはドリフト誤差と呼ばれる画質劣化が発生する。本論文の目的はドリフト誤差の低減方法を検討することである。

以下、エンコーダにおける処理をまとめる。但し、DCT 係数の量子化 (Q) および可変長符号化 (V) はドリフト誤差に関与しないため、これらの処理を文献 [7] と同様に割愛して議論する。また、フル解像度の画像にはドリフト誤差が発生しないため、低解像度画像のみを扱う。まず、 $t$  および  $n$  を正の整数として、時刻  $t$ 、空間位置  $n$  における原画像の画素値を  $x_{org}(n, t)$  とする。以下、 $t = 0$  ではフレーム内符号化、 $t > 0$  ではフレーム間予測符号化を行う場合を考える。すなわち、フレームメモリ (FM) には時刻  $t$  において、 $t = 0$  ならば零値、 $t > 0$  ならば画像  $x_{org}(n, t-1)$  が蓄積されているとする。このときエンコーダでは、ハーフペル精度の動き補償を記号  $MC_h[\cdot]$ 、8 点 DCT を記号  $T_8[\cdot]$  とすると、予測誤差の DCT 係数、

$$T_8[x_{dif}(n, t)] = T_8[x_{org}(n, t)] - T_8[MC_h[x_{org}(n, t-1)]] \quad (1)$$

が計算される。 $MC_h[\cdot]$  の具体的な処理方法は付録 1.

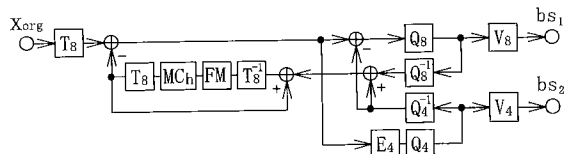


図 1 スケーラブル符号化におけるエンコーダ  
 Fig.1 Encoder of the scalable coding.

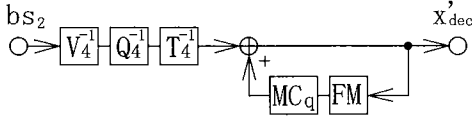


図2 従来のデコーダ1  
Fig.2 Conventional decoder 1.

に示す。次に、8点係数のうちの低域の4点のみを取り出す処理を  $E_4[\cdot]$  とすると、

$$bs_2 = E_4[T_8[x_{diff}(n, t)]] \quad (2)$$

がスケーラブルデコーダへ送られる。

## 2.2 従来のデコーダ1 [5], [6]

文献[5]のデコーダ(以下、デコーダ1)を図2に示す。この方法では、4点IDCTを記号  $T_4^{-1}[\cdot]$  と表すと、

$$x'_{dif}(n, t) = T_4^{-1}[bs_2] \alpha \quad (3)$$

により解像度が縦横ともに半分となった予測誤差が復号される。但し、 $\alpha$  は利得調整用のスカラ値であり、分離型2次元処理を行う場合、直流利得を1とするため  $\alpha = 2^{-1/2}$  が選ばれる。次に、フレームメモリ(FM)内の画像が動き補償され、式(3)の予測誤差に加算される。従って、

$$x'_{dec}(n, t) = x'_{dif}(n, t) + MC_q[x'_{dec}(n, t-1)] \quad (4)$$

により低解像度画像が復号される。但し、記号  $MC_q[\cdot]$  はクォーターペル精度の動き補償を表し、具体的な処理方法は付録1に示す。ここでは、画像の解像度が半分となるためにMCの精度をエンコーダの精度の半分としている[5]。

式(1)~(4)をまとめると、

$$\begin{aligned} x'_{dec}(n, t) &= RC[x_{org}(n, t)] \\ &\quad - RC[MC_h[x_{org}(n, t-1)]] \\ &\quad + MC_q[x'_{dec}(n, t-1)] \end{aligned} \quad (5)$$

となる。但し、

$$\begin{aligned} RC[\cdot] &= T_4^{-1}[E_4[T_8[\cdot]]] \alpha \\ &= T_4^{-1}[T_8[\cdot]] \alpha \end{aligned} \quad (6)$$

とする。更に、式(5)を再帰的に用いると、

$$x'_{dec}(n, t) = RC[x_{org}(n, t)] + d'_{err}(n, t) \quad (7)$$

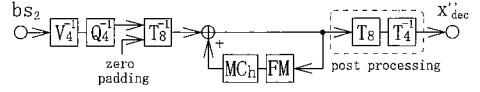


図3 従来のデコーダ2  
Fig.3 Conventional decoder 2.

となる。但し、

$$d'_{err}(n, t) = \sum_{k=0}^{t-1} MC_q^k[d'_{tmp}(n, t, k)] \quad (8)$$

であり、 $MC_q^k[\cdot]$  は  $k$  回の  $MC_q[\cdot]$  を行うことを意味し、

$$\begin{aligned} d'_{tmp}(n, t, k) &= -RC[MC_h[x_{org}(n, t-1-k)]] \\ &\quad + MC_q[RC[x_{org}(n, t-1-k)]] \end{aligned} \quad (9)$$

とする。

式(7)の右辺において、第1項は解像度変換された原画像を表し、第2項はドリフト誤差を表す。ドリフト誤差は式(8)からわかるようにMCを繰り返すたびに蓄積される[3]。式(9)の右辺を零とするには、デコーダの予測画像である第2項をエンコーダの予測画像である第1項に一致させる必要がある。しかし、第2項では第1項よりも少ない画素数に対してMCが行われるため、一般に両者は一致しない。文献[6]の方法のように、 $MC_q[\cdot]$  における補間精度を高次フィルタを用いて向上させることで、ある程度ドリフト誤差を低減できるが、補間を行う以上この方法には限界がある。

## 2.3 従来のデコーダ2 [7]

デコーダ1では失われた画素値を補間して予測画像を生成するため、ドリフト誤差が発生した。これに対し文献[7]では、図3に示すようにエンコーダと同じ解像度でMCを行い、ポスト処理として解像度変換することでドリフト誤差を低減している。但し、エンコーダで生成された8点DCT係数のうち、高域の4点はデコードされないため、これによるドリフト誤差が発生する。この方法(以下、デコーダ2)では、

$$x''_{dif}(n, t) = T_8^{-1}[bs_2] \quad (10)$$

によりエンコーダと同じ解像度の予測誤差が復号される。但し、8点IDCTを行う場合、高域の4点には零値が代入される。更に、FM内の画像がエンコーダと

同じ方法で動き補償され、式(10)に加算されて、

$$x''_{tmp}(n, t) = x''_{dif}(n, t) + MC_h[x''_{tmp}(n, t-1)] \quad (11)$$

となる。最後にポスト処理として8点DCTと4点IDCTが行われ、式(6)の表現を用いると、

$$x''_{dec}(n, t) = RC[x''_{tmp}(n, t)] \quad (12)$$

により低解像度画像が復号される。

式(1), (2), (10)~(12)をまとめると、式(7)と同様に、

$$x''_{dec}(n, t) = RC[DL[x_{org}(n, t)]] + RC[d''_{err}(n, t)] \quad (13)$$

が得られる。但し、

$$d''_{err}(n, t) = \sum_{k=0}^{t-1} MC_h^k[d''_{tmp}(n, t, k)] \quad (14)$$

であり、

$$d''_{tmp}(n, t, k) = -DL[MC_h[x_{org}(n, t-1-k)]] + MC_h[DL[x_{org}(n, t-1-k)]] \quad (15)$$

とする。また、

$$DL[\cdot] = T_8^{-1}[E_4[T_8[\cdot]]] \quad (16)$$

とおく。なお、ここでは $E_4[\cdot]$ は高域の4点に零を代入する演算と解釈する。

以下、デコーダ2によりドリフト誤差が低減される理由を考察する。式(15)の右辺において、 $DL[\cdot]$ はマルチチャネルフィルタバンクによるマルチレート信号処理と等価であるため、付録2.で定義するエリアス項と非エリアス項に分離できる。デコーダ2ではエンコーダとデコーダにて同じ $MC_h[\cdot]$ が使われるため、付録3.に示すように、非エリアス項については式(15)の右辺の第1項と第2項が相殺する。これに対し、デコーダ1では相殺しない。一方、エリアス項については一般に相殺しないため、これによるドリフト誤差が発生する。しかし、式(13)の第2項における $RC[\cdot]$ によりドリフト誤差のエネルギーが半減されるため、3.3において確認するようにデコーダ1に比べて効果的にドリフト誤差が低減される。

### 3. ハードウェア構成の簡単なデコーダ

ドリフト誤差の発生するデコーダ1に対し、ドリフト誤差を低減するためのデコーダ2が提案された。しかし、この方法は式(12)のポスト処理を必要とするため、ハードウェア構成が複雑になる欠点をもつ。以下、デコーダ2において、ドリフト低減効果を損なわないようにポスト処理を簡略化する。

#### 3.1 処理手順

デコーダ2において、式(12)のポスト処理 $RC[\cdot]$ を、より簡単なハードウェアで構成される新たな処理 $DS[\cdot]$ に置き換える。すなわち、式(13)を、

$$x^*_{dec}(n, t) = DS[DL[x_{org}(n, t)]] + DS[d''_{err}(n, t)] \quad (17)$$

に置き換える。

次に、ポスト処理 $DS[\cdot]$ の具体的な方法を考える。デコーダ2では、式(13)の第1項を式(7)の第1項に一致させるために式(6)の $RC[\cdot]$ を用いた。このことは、式(6)、式(16)より、

$$RC[DL[\cdot]] = T_4^{-1}[T_8[T_8^{-1}[E_4[T_8[\cdot]]]]] \alpha = T_4^{-1}[T_8[\cdot]] \alpha = RC[\cdot] \quad (18)$$

となることから確認できる。しかし、式(7)の第1項には $RC[\cdot]$ に起因する解像度変換誤差が含まれているため、これに一致させる必要はない。従って、解像度を半分にできるならば、 $RC[\cdot]$ 以外の簡単な処理に置き換えても問題はない。そこで本論文では、遮断周波数が $0.5\pi[\text{rad}]$ である低域通過フィルタ $L(z)$ と間引き処理 $\downarrow 2$ を $DS[\cdot]$ として用いた、図4に示す復号方法を提案する。デコーダ2と同様にドリフト誤差を低減させるため、 $L(z)$ は式(17)の第2項のエネルギーを半減させる必要がある。但し、式(17)の第1項からわかるように、信号 $x_{org}(n, t)$ のエリアス成分は $DL[\cdot]$ により既に抑圧されているため、 $L(z)$ は大きな阻止域減衰量をもつ必要がなく、従って、短いタップ数で構成できる。なお、タップ数を偶数とするとデコーダ1, 2と同じ位相をもつ縮小画像が復号される。

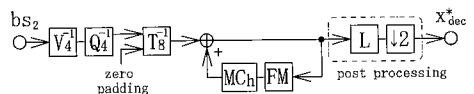
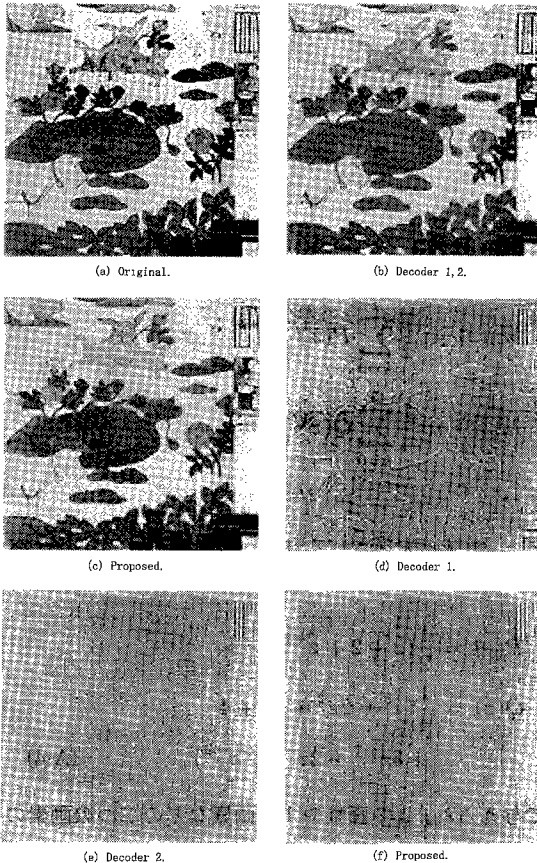


図4 提案するデコーダ  
Fig.4 Proposed decoder.

表 1 ハードウェアの比較  
Table 1 Comparison of the hardware complexity.

	FM	MC	ポスト処理
デコーダ 1	◎: 小型	×: 特殊	◎: なし
デコーダ 2	○: 標準	○: 標準	×: 複雑
提案法	○: 標準	○: 標準	○: 簡単



(“Mobile & calendar”, interlaced,  $t = 1, v = 0.5$ )

図 5 復号された低解像度画像  
Fig.5 Reconstructed image.

### 3.2 ハードウェアの比較

従来のデコーダ 1, 2 の問題点, および提案法の特徴を表 1 にまとめる. デコーダ 1 は解像度変換後に MC を行うため, 処理対象となる画素数が縦横ともに半分となる利点をもつ. しかし, ドリフト誤差が顕著であり, これを軽減するには特殊な MC を必要とする. デコーダ 2 では, 処理すべき画素数はエンコーダと同じであるが, MC において MPEG 国際標準用に開発された標準的な LSI が利用できる上に, デコーダ 1 に比

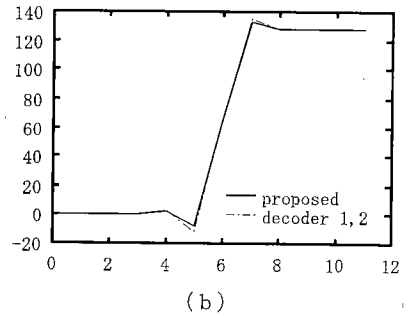
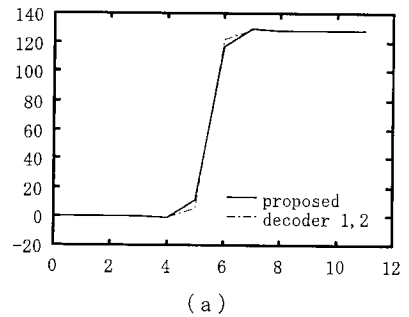


図 6 復号されたステップ入力  
Fig.6 Reconstructed image.

べてドリフト誤差を低減できる. しかし, ポスト処理として 2 回の直交変換を必要とする欠点がある. 提案法では, デコーダ 2 のポスト処理が簡略化され, 8 点 IDCT やハーフペル精度の MC といった標準的な LSI のみで構成でき, なおかつ, 以下で確認するように, デコーダ 1 と同程度以上のドリフト低減効果がある.

### 3.3 再生画像の比較

$x_{org}(n, t)$  として画像信号を用いた場合の効果を調べる. ここでは, 処理の簡単な 2 タップのフィルタ,

$$L(z) = \frac{1 + z^{-1}}{2} \quad (19)$$

を用いた. また, 付録の式 (A-3) における動きベクトルを  $v = 0.5$  とした. 実際の符号化では動き推定アルゴリズムにより入力画像に対する最適値が求められる [8]. 図 5 (a) には原画像を示す. 同図 (b) には式 (7), 式 (13) の第 1 項を, 同図 (c) には式 (17) の第 1 項をそれぞれ示す. 参考のため,  $x_{org}(n, t)$  をステップ関数とした場合を図 6 に示す. 同図の例では, デコーダ 1, 2 に比べて提案法ではエッジ部分が若干ぼけている. しかし, この違いは図 5 の画像信号の場合は顕著ではない.

図 5 (d), (e), (f) および表 2 にはドリフト誤差であ

表2 ドリフト誤差の比較  
Table 2 Comparison of the drift error.

(t = 1, v = 0.5, 256 × 256pxls, PSNR in [dB])		
	Mobile & calendar	Flower garden
デコーダ1	31.7 ( 0.0)	33.0 ( 0.0)
デコーダ2	36.4 (+4.7)	38.1 (+5.1)
提案法	36.8 (+5.1)	38.6 (+5.6)

る式 (7), (13), (17) それぞれの第2項とその PSNR (Peak-Signal to Noise Ratio) を示す。同図表より、デコーダ1に対する提案法のドリフト低減効果は、デコーダ2と同程度以上であることがわかる。なお、 $L(z)$  の選定については今後検討の余地がある。

#### 4. むすび

スケーラブル符号化により復号された低解像度画像に見られるドリフト誤差を低減でき、なおかつ簡易なハードウェアで実現できる復号方法を提案した。また、画像信号に適用した結果から提案法の効果を確認した。

本方法は解像度変換比が2:1に限られている。今後は、任意の有理数変換比に拡張する必要がある。

**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に深謝致します。

#### 文 献

- [1] M.R. Civanlar and A. Puri, "Scalable video coding in frequency domain," Proc. SPIE Visual Comm. and Image Processing, Boston, MA, pp.1124-1134, 1992.
- [2] 花村 剛, 富永英義, "スケーラブルビデオ符号化," テレビ技報, ICS'93-62, vol.17, no.57, pp.25-30, 1993.
- [3] M.R. Civanlar and A. Puri, "Issues in efficient frequency scalable coding," PCS '93, 12.2, 1993.
- [4] K.N. Ngan, J. Arnold, T. Sikora, T.K. Tan, and A.W. Johnson, "Frequency scalability experiments for MPEG-2 standard," Asia-Pacific Conference on Communications '93, 3D.2.1, Taejon, Aug. 1993.
- [5] "Test Model 5," Test Model Editing Committee, ISO-IEC JTCl/SC29/WG11, MPEG93/400, 1993.
- [6] A.W. Johnson and J. Princen, "Drift minimization in frequency scalable coders using block based filtering," IEEE Visual Signal Processing and Communications, 20-22, pp.231-234, 1993.
- [7] R. Mokry and D.A. Anastassiou, "Minimal error drift in frequency scalability for motion-compensated DCT coding," IEEE Trans., CAS for Video Technology, vol.4, no.4, pp.392-406, Aug. 1994.
- [8] ISO-IEC JTCl, CD11172, "Information Technology Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbits/s," 1991.
- [9] 岩橋政宏, 神林紀嘉, 貴家仁志, "低域ドリフトのないスケーラブル・デコーダ," 信学技報, DSP94-108, Jan. 1995.

#### 付 録

##### 1. $MC_h[\cdot]$ および $MC_q[\cdot]$

以下、 $x_{org}(n, t)$  の  $z$  変換を、

$$X_{org}(z, t) = \sum_{n=0}^{N-1} x_{org}(n, t) z^{-n} \quad (\text{A}\cdot\text{1})$$

と表現する。 $X_{org}(z, t)$  を  $MC_h[\cdot]$  および  $MC_q[\cdot]$  により動き補償した結果はそれぞれ、

$$\begin{aligned} MC_h[X_{org}(z, t)] &= M_h(z) X_{org}(z, t) \\ MC_q[X_{org}(z, t)] &= M_q(z) X_{org}(z, t) \end{aligned} \quad (\text{A}\cdot\text{2})$$

但し、

$$\begin{aligned} M_h(z) &= \frac{1}{2} \sum_{k=0}^1 P(z^{1/2} e^{j\pi k}) (z^{1/2} e^{j\pi k})^{2v} \\ M_q(z) &= \frac{1}{4} \sum_{k=0}^3 Q(z^{1/4} e^{j\pi k/2}) (z^{1/4} e^{j\pi k/2})^{2v} \end{aligned} \quad (\text{A}\cdot\text{3})$$

と  $z$  領域で表現される [9]。式 (A·3) における  $P(z)$ ,  $Q(z)$  は、

$$\begin{aligned} P(z) &= \frac{1}{2}(z + 2 + z^{-1}) \\ Q(z) &= \frac{1}{4}(z^3 + 2z^2 + 3z + 4 \\ &\quad + 3z^{-1} + 2z^{-2} + z^{-3}) \end{aligned} \quad (\text{A}\cdot\text{4})$$

で与えられ、 $1/2$  画素や  $1/4$  画素などの二つの画素の間に位置する分数画素を直線補間することを意味する。 $v$  はエンコードにおける動きベクトルであり、空間的なシフトの方向と画素数を表す。この場合、ハーフペル精度なので、 $0, \pm 0.5, \pm 1, \pm 1.5, \dots$  の値をとる。

##### 2. エリアス項および非エリアス項

$X_{org}(z, t)$  に  $RC[\cdot]$  および  $DL[\cdot]$  を行った結果はそれぞれ、

$$\begin{aligned} RC[X_{org}(z, t)] &= \alpha \sum_{p=0}^7 B_F(z, p) X_{org}(z^{1/2} e^{j\pi p/4}, t) \\ DL[X_{org}(z, t)] &= \sum_{p=0}^7 B_S(z, p) X_{org}(z e^{j\pi p/4}, t) \end{aligned} \quad (\text{A}\cdot\text{5})$$

但し、

$$B_F(z, p) = \frac{1}{8} \sum_{q=0}^3 H_q(z^{1/2} e^{j\pi p/4}) G_q(z)$$

$$B_S(z, p) = \frac{1}{8} \sum_{q=0}^3 H_q(z e^{j\pi p/4}) F_q(z) \quad (\text{A}\cdot 6)$$

となる。フィルタ  $H_q(z)$ ,  $G_q(z)$ ,  $F_q(z)$  は DCT の基底係数、

$$e_N(q, n) = \frac{s(q)}{\sqrt{N}} \cos \frac{(2n+1)q\pi}{2N},$$

$$s(q) = \begin{cases} 1 & (q=0) \\ \sqrt{2} & (q \neq 0) \end{cases} \quad (\text{A}\cdot 7)$$

により、それぞれ、

$$H_q(z) = \sum_{n=0}^7 e_8(q, n) z^{+n}$$

$$G_q(z) = \sum_{n=0}^3 e_4(q, n) z^{-n}$$

$$F_q(z) = \sum_{n=0}^7 e_8(q, n) z^{-n} \quad (\text{A}\cdot 8)$$

と関係づけられる [9]。本論文では式 (A.5) を、

$$RC[X_{org}(z, t)]$$

$$= \alpha B_F(z, 0) X_{org}(z^{1/2}, t)$$

$$+ \alpha \sum_{p=1}^7 B_F(z, p) X_{org}(z^{1/2} e^{j\pi p/4}, t) \quad (\text{A}\cdot 9)$$

$$DL[X_{org}(z, t)]$$

$$= B_S(z, 0) X_{org}(z, t)$$

$$+ \sum_{p=1}^7 B_S(z, p) X_{org}(z e^{j\pi p/4}, t) \quad (\text{A}\cdot 10)$$

と展開し、それぞれの第 1 項を非エリアス項、第 2 項をエリアス項と呼ぶ。

### 3. 非エリアス項の相殺

付録 1., 2. の結果を用いて式 (15) を  $z$  領域で表現すると、

$$D''_{imp}(z, t, k)$$

$$= -M_h(z) B_S(z, 0) X_{org}(z, t-1-k)$$

$$- \sum_{p=1}^7 M_h(z e^{j\pi p/4}) B_S(z, p)$$

$$\cdot X_{org}(z e^{j\pi p/4}, t-1-k)$$

$$+ M_h(z) B_S(z, 0) X_{org}(z, t-1-k)$$

$$+ M_h(z) \sum_{p=1}^7 B_S(z, p)$$

$$\cdot X_{org}(z e^{j\pi p/4}, t-1-k) \quad (\text{A}\cdot 11)$$

となる。上式より、非エリアス項である第 1 項と第 3 項が相殺することがわかる。

(平成 8 年 1 月 23 日受付)