

変換符号化における基底係数のビット数低減法

岩橋 政宏[†] 神林 紀嘉[†] 貴家 仁志^{††}

Bit Reduction of DCT basis for Transform Coding

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

あらまし デジタル画像データを圧縮するにあたり、離散コサイン変換 (DCT) を用いた変換符号化が盛んに研究されている。一方、その応用分野においては画素数の増加による高精細化が進みつつあり、1画素当りの符号化に要する時間の短縮が強く要求されている。これに対し符号化演算における乗算器係数、すなわち DCT 基底係数の値を2進数表現する際に、できるだけ少ないビット数で表現することで処理の簡略化や演算時間の短縮が可能となる。但し、このビット数低減により再生画像の画質が劣化してはならないという制約条件を考慮する必要がある。そこで本論文では、こうした条件の下においても変換係数の量子化処理を考慮することで、基底係数のビット数を低減できることを示す。また、JPEG 等の具体的な符号化例におけるビット数低減の限界値を明らかにする。最後に、実際の画像データを変換符号化することで、従来報告されていたビット数よりも少ないビット数で基底係数値を表現しても、本論文で示した限界値以内であれば再生画像の画質がほとんど変化しないことを確認する。

キーワード 符号化, コサイン変換, 乗算器, ビット数, ハードウェア, 画像

1. ま え が き

近年、音声や文書に比べて膨大なデータ量をもつ画像データを圧縮する画像符号化が盛んに研究されている。なかでも離散コサイン変換 (DCT) [1] を用いた変換符号化は、ISO/IEC や ITU-T による国際標準 [2], [7], [8] の骨子となっており、あらゆる分野にて幅広く研究されている。一方、その応用においては画素数の増加によるハイビジョン化が進みつつあり、1画素当りの符号化に要する時間の短縮や処理の簡略化が必須となっている。そこで本論文では、変換符号化における乗算器、すなわち DCT 基底の値をできるだけ少ないビット数で表現することで符号化処理の簡略化を行う。基底係数のビット数に関する研究はこれまでにいくつか報告されている [3]~[5]。CCITT 勧告である H.261 では、デコーダ間での受信画像の不一致を回避するための DCT の演算精度を規定しており [2]、このとき発生する演算誤差については文献 [3]

にて解析された。また、異なるデコーダ間で再生画像が一致するために必要なビット数やその決定手順は、文献 [4], [5] に報告されている。これら従来の報告では、順変換の結果である変換係数を量子化しないと仮定している。しかし、実際の符号化では量子化により変換係数のビット数を減少させてデータ圧縮を行う。このため、量子化を行った場合、従来報告された基底係数のビット数は、必要となる最低ビット数に対して過剰となってしまふ。そこで本論文では、符号化処理の簡略化を目的とし、従来法における基底係数の過剰ビットを削減する方法について考察する。そのためにまず、変換係数と基底係数それぞれのビット数をパラメータとした場合の、量子化誤差のエネルギーを理論的に推定する。その結果、基底係数のビット数を従来使われていたビット数よりも少なくしても、ある限界値以上では量子化誤差のエネルギーに影響を与えないことがわかる。すなわち、基底係数の必要最小ビット数が明らかになる。この結果を用いれば、従来法と同程度の再生画像の画質を保ちながらも、従来法よりも少ないビット数で基底係数が表現でき、符号化装置のハードウェア規模の簡略化又は処理時間の短縮が可能となる。以下、2. では、変換符号化について説明する。3. では、基底係数のビット数低減方法を述べる。4. で

[†] 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

^{††} 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

は、具体的な量子化に対する基底係数の最低ビット数をシミュレーションにより調べる。最後に 5. では、本論文で得られた結論をまとめる。

2. 変換符号化 [1], [6]

初めに、変換符号化の基本処理および変換係数の量子化について説明する。

2.1 基本処理

変換符号化の基本処理を、図 1 を用いて説明する。まず、符号化する原信号を N 点ごとに区切り、いくつかのブロックに分割する。ここで、原信号は B_s ビットの 2 進数整数で表されているとする。つぎに、各ブロックにおいて、 N 点の原信号 $s(n)$ ($n = 0, 1, \dots, N-1$) を離散コサイン変換 (DCT) することで N 種類の変換係数、

$$\begin{aligned}
 t(f) &= I[T_N[s(n)]] \\
 &= I\left[\sum_{n=0}^{N-1} b(n, f)s(n)\right] \\
 &\quad (f = 0, 1, \dots, N-1)
 \end{aligned}
 \tag{1}$$

を得る。但し、DCT 結果である変換係数の値は整数化される。なお、記号 $T_N[\]$ および $b(n, f)$ はそれぞれ DCT および基底係数とする。また、記号 $I[x]$ は x をこれに最も近い整数へ丸めた結果を表す。式 (1) における基底係数 $b(n, f)$ は、

$$b(n, f) = c(f) \cos \frac{(2n+1)f\pi}{2N}
 \tag{2}$$

但し、

$$\begin{cases} c(f) = \frac{1}{\sqrt{N}} & (f = 0) \\ c(f) = \sqrt{\frac{2}{N}} & (f = 1, 2, \dots, N-1) \end{cases}$$

で与えられる。つぎに、式 (1) で得られた変換係数 $t(f)$ を量子化することで、その値を 2 進数表現する際に必要となるビット数を削減し、

$$\hat{t}(f) = Q[t(f)] \quad (f = 0, 1, \dots, N-1)
 \tag{3}$$

を得る。量子化 $Q[\]$ については 2.3 において詳しく説明する。最後に量子化された変換係数 $\hat{t}(f)$ を逆変換し、再生信号、

$$\hat{s}(n) = I[T_N^{-1}[\hat{t}(f)]]$$

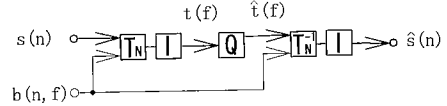


図 1 変換符号化の基本処理
Fig. 1 Basic signal processing of the transform coding.

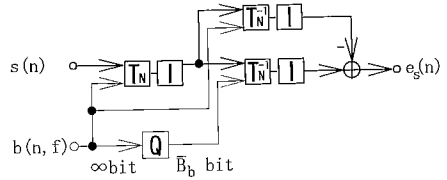


図 2 従来の評価方法
Fig. 2 Conventional method for the estimation.

$$\begin{aligned}
 &= I\left[\sum_{f=0}^{N-1} b(n, f)\hat{t}(f)\right] \\
 &\quad (n = 0, 1, \dots, N-1)
 \end{aligned}
 \tag{4}$$

を出力する。ここで、記号 $T_N^{-1}[\]$ は逆 DCT を表す。また、整数化 $I[\]$ により再生信号 $\hat{s}(n)$ の値は原信号 $s(n)$ と同じ B_s ビットの 2 進数整数で表される。

2.2 基底係数の必要最小ビット数

2.1 の基本処理において、式 (2) の基底係数値は実数で定義されており、値を 2 進数表現した場合、一般に無限のビット数を必要とする。しかし、符号化装置をハードウェア化する際に有限のビット数に語長制限されるため、DCT 演算には基底係数の有限語長化による誤差が伴う。CCITT の勧告 H.261 では、図 2 の評価方法により逆変換における基底係数に必要とされる語長を規定している。同図の評価方法では、逆変換において式 (2) の基底係数の値が無限ビットの場合と有限値の場合とで、それぞれの再生信号の差分値に対する許容範囲を規定している [2]。ここで、無限ビットとは 64 ビット浮動小数点表現を意味する。その結果、基底係数のビット数は 10 ビット以上が必要とされている [3]~[5]。但し、ここでは変換係数の量子化は行われない。

しかし実際の符号化ではデータ圧縮のために式 (3) の量子化により変換係数のビット数が削減される。これに伴い基底係数に要求される必要最小ビット数は減少し、従来報告された基底係数のビット数は、実際に必要となるビット数に対して過剰になると考えられる。本論文の目的は、変換係数の量子化により生じる基底係数の過剰ビットを削減し、乗算器係数の語長低減に

より符号化処理を簡略化することである。以下、変換係数の量子化について説明した後、3.にて基底係数の必要最小ビット数の決定法を述べる。

2.3 変換係数の量子化

式(3)の量子化 $Q[\]$ はステップサイズ $q_t(f)$ を用いて、各ブロックにおいて、

$$Q[t(f)] = I\left[\frac{t(f)}{q_t(f)}\right]q_t(f) \quad (f = 0, 1, \dots, N-1) \quad (5)$$

により行われる。上式では $t(f)$ の値が $q_t(f)$ の整数倍に量子化される。変換符号化では、このときの整数値が圧縮情報となりエンコードから出力される。一方、デコードではこの値にステップサイズが掛けられて変換係数が復号される。いま、全ブロックにおける $t(f)$ の絶対値の最大値を $R_t(f)$ とすると、第 f 帯域の変換係数 $t(f)$ を 2 進数表現するには、1 ビットの符号ビットを含め、

$$B_t(f) = \log_2 \frac{R_t(f)}{q_t(f)} + 1 \quad (f = 0, 1, \dots, N-1) \quad (6)$$

ビットの 2 値符号が必要となる。但し、式(6)の $B_t(f)$ の値は一般に実数となるので、実際にハードウェアを構成する場合は、 $[x]$ を x を超えない最大の整数とすると、 $[B_t(f)] + 1$ ビットの 2 進符号を用意することになる。一方、値域が $[-R, R]$ である信号のエネルギーは R^2 に比例するため[6]、変換係数のエネルギー、

$$E_t(f) = \{t(f)\}^2 \quad (f = 0, 1, \dots, N-1) \quad (7)$$

は、 $R_t(f)$ により、

$$E_t(f) = \alpha_1 R_t^2(f) \quad (f = 0, 1, \dots, N-1) \quad (8)$$

と表される。ここで、 α_1 は比例定数であり $t(f)$ の値の分布関数に依存する[6]。式(6)、(8)より $R_t(f)$ を消去すると、

$$q_t(f) = \sqrt{\frac{4E_t(f)}{\alpha_1}} \cdot 2^{-B_t(f)} \quad (f = 0, 1, \dots, N-1) \quad (9)$$

が得られ、各帯域ごとのステップサイズ $q_t(f)$ とビット数 $B_t(f)$ の関係がまとめられる。以上で、各帯域における変換係数の量子化が、ステップサイズ $q_t(f)$ あるいはビット数 $B_t(f)$ をパラメータとして表現された。

2.4 変換係数のステップサイズ[6]

式(9)における帯域ごとのビット数 $B_t(f)$ は、信号の量子化誤差、

$$e_s(n) = \hat{s}(n) - s(n) \quad (10)$$

の 1 サンプル当りの平均エネルギー、

$$E_{es} = \frac{1}{N} \sum_{n=0}^{N-1} e_s^2(n) \quad (11)$$

が最小となるように決定することができる[6]。以下、式(11)で定義された E_{es} を誤差エネルギーと呼び、これを最小化する $B_t(f)$ の決定方法を簡単にまとめる。変換係数の量子化誤差、

$$e_t(f) = \hat{t}(f) - t(f) \quad (f = 0, 1, \dots, N-1) \quad (12)$$

のエネルギー、

$$E_{et}(f) = e_t^2(f) \quad (f = 0, 1, \dots, N-1) \quad (13)$$

を用いると、誤差エネルギー E_{es} は、基底係数が式(2)で与えられる場合、

$$E_{es} = \frac{1}{N} \sum_{f=0}^{N-1} E_{et}(f) \quad (14)$$

と表される。ところで式(14)の $E_{et}(f)$ はステップサイズ $q_t(f)$ により、

$$E_{et}(f) = \frac{1}{12} q_t^2(f) \quad (15)$$

と表されるため[6]、式(9)、(15)より $q_t(f)$ を消去すると、

$$E_{et}(f) = \frac{1}{3\alpha_1} E_t(f) 2^{-2B_t(f)} \quad (16)$$

となる。従って、式(14)、(16)より $E_{et}(f)$ を消すと、

$$E_{es} = \frac{1}{3N\alpha_1} \sum_{f=0}^{N-1} E_t(f) 2^{-2B_t(f)} \quad (17)$$

となり、 E_{es} が各帯域のビット数 $B_t(f)$ により表される。また、 $B_t(f)$ の全帯域での平均値、

$$\bar{B}_t = \frac{1}{N} \sum_{f=0}^{N-1} B_t(f) \quad (18)$$

はあらかじめ与えられている。従って、目的とする $B_t(f)$ の値は、式 (18) の条件下で式 (17) の誤差エネルギーを最小とする $B_t(f)$ の決定問題の解となる。この解はラグランジュの未定乗数決定法により得られ、その結果、各帯域のビット数 $B_t(f)$ が、

$$B_t(f) = \bar{B}_t + \frac{1}{2} \log_2 \frac{E_t(f)}{\left\{ \prod_{g=0}^{N-1} E_t(g) \right\}^{1/N}} \quad (19)$$

のとき E_{es} の値が最小となる [6]。そこで、式 (19) を式 (9) に代入すると、

$$q_t(f) = \alpha_t \cdot 2^{-\bar{B}_t} = q_t \quad (f = 0, 1, \dots, N-1) \quad (20)$$

但し、

$$\alpha_t = \frac{1}{\sqrt{\alpha_1}} \left\{ \prod_{g=0}^{N-1} E_t(g) \right\}^{1/2N}$$

となり、 E_{es} が最小となるときの $q_t(f)$ が \bar{B}_t により表される。この場合、式 (20) より $q_t(f)$ が全帯域で等しくなることがわかる。

以上、変換係数の量子化は式 (5) に示されるように各帯域ごとのステップサイズ $q_t(f)$ をパラメータとして記述される。また、式 (11) の誤差エネルギーが最小となるとき、 $q_t(f)$ は式 (20) のように平均ビット数 \bar{B}_t で表され、全帯域で値が等しくなる。

3. 基底係数のビット数低減方法

つぎに、2.3 で説明した変換係数の量子化を行った場合の、基底係数の有限語長化について述べる。

3.1 評価方法

本論文では、2.2 でも述べたように、変換係数を量子化した場合において、基底係数の有限語長化が式 (11) の誤差エネルギー E_{es} に影響を及ぼさない範囲での基底係数の必要最小ビット数を調べている。このため、図 2 に示した従来の評価方法に代わり、基底係数の有限語長化、

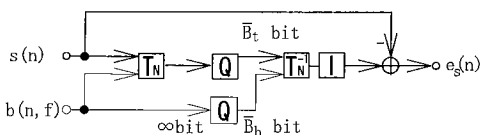


図3 提案する評価方法
Fig.3 Proposed method for the estimation.

$$\hat{b}(n, f) = Q[b(n, f)] \quad (21)$$

を導入した図 3 の評価方法を用いる。すなわち、式 (4) に代わり、語長制限された基底係数 $\hat{b}(n, f)$ を用いた、

$$\begin{aligned} \hat{s}(n) &= I[T_N^{-1}[\hat{t}(f)]] \\ &= I\left[\sum_{f=0}^{N-1} \hat{b}(n, f)\hat{t}(f)\right] \\ &\quad (n = 0, 1, \dots, N-1) \end{aligned} \quad (22)$$

により逆変換を行う。

3.2 基底係数の有限語長化

本論文では式 (21) の基底係数 $b(n, f)$ の有限語長化を、式 (5) と同様、ステップサイズ $q_b(f)$ を用いて、

$$Q[b(n, f)] = C_b(f)q_b(f) \quad (23)$$

但し、

$$\begin{aligned} C_b(f) &= I\left[\frac{b(n, f)}{q_b(f)}\right] \\ &\quad (n = 0, 1, \dots, N-1, f = 0, 1, \dots, N-1) \end{aligned}$$

により行う。ここで、 $q_b(f)$ は正の整数 $Q_b(f)$ により、

$$q_b(f) = 2^{-Q_b(f)} \quad (Q_b(f) \text{ は正の整数}) \quad (24)$$

とする。式 (23) により、本来、無限語長で定義された $b(n, f)$ の値が $q_b(f)$ の整数倍に近似される。従って、式 (22) における $\hat{b}(n, f)$ と $\hat{t}(f)$ の乗算は、初めに整数 $\hat{t}(f)$ と整数 $C_b(f)$ の乗算を行った後、 $q_b(f)$ で乗算することで実行される。なお、このときの乗算は $Q_b(f)$ が整数であるため 2 進数のシフト演算となる。また、整数 $C_b(f)$ の値を 2 進数表現するには、各帯域 f におけるすべての n に対する $b(n, f)$ の絶対値の最大値を $R_b(f)$ とすると、符号ビットを含め、

$$\begin{aligned} B_b(f) &= \log_2 \frac{R_b(f)}{q_b(f)} + 1 \\ &= -\log_2 q_b(f) + \log_2 R_b(f) + 1 \\ &\quad (f = 0, 1, \dots, N-1) \end{aligned} \quad (25)$$

ビットの 2 値符号が必要となる。なお、 $R_b(f)$ の値は式 (2) の最大の絶対値から計算される。但し、式 (25) の $B_b(f)$ の値は一般に実数となるので、実際にハードウェアを構成する場合は、 $[B_b(f)] + 1$ ビットの 2 進符号を用意する。

以上、基底係数の有限語長化が $q_b(f)$ をパラメータ

として式 (23) により記述された。また、有限語長化された基底係数 $\hat{b}(n, f)$ を 2 進数表現する際に必要となるビット数 $B_b(f)$ は式 (24), (25) からわかるように、 $Q_b(f)$ に比例する。

3.3 基底係数のステップサイズ

2.4 で説明したように、誤差エネルギーを最小とする変換係数のステップサイズは全帯域で等しくなった。また、このときステップサイズは平均ビット数で表された。そこで以下、基底係数のステップサイズが全帯域で等しい、

$$q_b(f) = q_b \quad (f = 0, 1, \dots, N-1) \quad (26)$$

の場合について、ステップサイズと平均ビット数の関係を調べる。基底係数の平均ビット数 \bar{B}_b は、

$$\bar{B}_b = \frac{1}{N} \sum_{f=0}^{N-1} B_b(f) \quad (27)$$

で定義され、上式に式 (25), (26) を代入すると、

$$\bar{B}_b = -\log_2 q_b + \log_2 2 \left\{ \prod_{g=0}^{N-1} R_b(g) \right\}^{1/N} \quad (28)$$

となる。また、上式を変形すると式 (20) と同様な、

$$q_b = \alpha_b 2^{-\bar{B}_b} \quad (f = 0, 1, \dots, N-1) \quad (29)$$

但し、

$$\alpha_b = 2 \left\{ \prod_{g=0}^{N-1} R_b(g) \right\}^{1/N}$$

が得られる。以上、式 (29) により、基底係数を全帯域において同じ値のステップサイズで有限語長化した場合の、ステップサイズ q_b と平均ビット数 \bar{B}_b の関係が示された。

3.4 有限語長化の誤差エネルギーへの影響

これまでの結果、2.3, 2.4 にて変換係数の量子化が、また 3.2, 3.3 にて基底係数の有限語長化が、それぞれ平均ビット数 \bar{B}_t および \bar{B}_b をパラメータとして記述された。以下、これらの処理の誤差エネルギー E_{es} への影響を理論的に解析する。これにより、誤差エネルギーを \bar{B}_t および \bar{B}_b の関数として記述する。本論文では最終的に、与えられた \bar{B}_t の下で、 E_{es} の値が $\bar{B}_b = \infty$ の場合と同程度である最小の \bar{B}_b を求めることを目的としている。

3.4.1 量子化誤差の成分

初めに、式 (10) で定義された量子化誤差 $e_s(n)$ に含まれる成分を調べる。表現の簡略化のため $b(n, f)$ を用いた順変換 $T_N[x]$ を $b(n, f) * x$ と表すと、変換係数は、

$$t(f) = b(n, f) * s(n) \quad (30)$$

と表される。同様に、有限語長化された $\hat{b}(n, f)$ を用いた逆変換 $T_N^{-1}[x]$ を $\hat{b}(n, f) * x$ と表すと、再生信号は、

$$\hat{s}(n) = \hat{b}(n, f) * \hat{t}(f) + e_i \quad (31)$$

但し、

$$e_i = I[\hat{b}(n, f) * \hat{t}(f)] - \hat{b}(n, f) * \hat{t}(f)$$

となる。ここで、 e_i は整数化 $I[\]$ により発生する誤差である。更に、有限語長化された基底係数 $\hat{b}(n, f)$ を真値と誤差値に分けて、

$$\hat{b}(n, f) = b(n, f) + e_b(n, f) \quad (32)$$

と表す。以上をまとめると、式 (10), (12), (30)~(32) より、

$$\begin{aligned} e_s(n) &= e_b(n, f) * b(n, f) * s(n) \\ &\quad + b(n, f) * e_t(f) \\ &\quad + e_b(n, f) * e_t(f) \\ &\quad + e_i \end{aligned} \quad (33)$$

が得られ、量子化誤差 $e_s(n)$ が四つの項で表された。

3.4.2 成分のエネルギー

つぎに、これら四つの項のエネルギーを調べる。ここでは、ステップサイズ $q_t(f)$ および $q_b(f)$ が全帯域で等しい場合を考える。なお、帯域ごとに異なる場合については 3.6 において考察する。式 (33) の各項に含まれる信号の値は、表 1 にまとめられる分布範囲をもつ。例えば、信号 $e_t(f)$ はステップサイズ 2^{Q_t} による量子化誤差なので、その最大の絶対値はステップサイズの 0.5 倍となる。従って、信号 $e_t(f)$ の値の分布範囲は $\pm 2^{Q_t-1}$ となる。但し、表中の Q_t および Q_b を式 (20), (26) の q_t および q_b により、

$$Q_t = \log_2 q_t, \quad f = 0, 1, \dots, N-1 \quad (34)$$

$$Q_b = -\log_2 q_b, \quad f = 0, 1, \dots, N-1 \quad (35)$$

と定義する。ここで、 Q_b は式 (24) でも述べたように、

表1 成分の振幅値の分布範囲
Table 1 Range of the signals.

single	range
e_t	$\pm 2^{Q_t-1}$
s	$\pm 2^{B_s-1}$
e_i	$\pm 2^{-1}$
e_b	$\pm 2^{-Q_b-1}$

表2 成分の振幅値の分布範囲
Table 2 Range of the signals.

single	range
$e_b * b * s$	$\pm 2^{-Q_b+B_s-2}$
$b * e_t$	$\pm 2^{Q_t-1}$
$e_b * e_t$	$\pm 2^{Q_t-Q_b-2}$
e_i	$\pm 2^{-1}$

正の整数である。一方、 q_t には任意の正の整数が使われるため [7], [8], Q_t は正の実数となる。以下、これら対数ステップと呼ぶ。式 (34), (35) に式 (20), (29) を代入すると、

$$Q_b = \bar{B}_b - \log_2 \alpha_b \quad (36)$$

$$Q_t = -\bar{B}_t + \log_2 \alpha_t \quad (37)$$

となり、対数ステップは平均ビット数に比例する。表1を用いると式 (33) の四つの項の振幅値の分布範囲がわかり、得られた結果を表2にまとめる。但し、表2では入力信号 $s(n)$ のビット数を B_s とした。また、基底係数 $b(n, f)$ による順変換は信号の分布範囲を変えないことを利用した。これは、順変換を伝達関数が、

$$B(z, f) = \sum_{n=0}^{N-1} b(n, f) z^{-n} \quad (f = 0, 1, \dots, N-1) \quad (38)$$

である FIR (Finite Impulse Response) フィルタによる処理として捉えたとき、フィルタの最大利得が1以下となることを意味する。

一方、値が $\pm 2^{-1}R$ の範囲内に分布する信号のエネルギーは R^2 に比例するため [6], 式 (33) の誤差信号 $e_s(n)$ のエネルギー E_{es} は、

$$E_{es} = \beta_1 \cdot 2^{2(-Q_b+B_s-1)} + \beta_2 \cdot 2^{2Q_t} + \beta_3 \cdot 2^{2(Q_t-Q_b-1)} + \beta_4 \quad (39)$$

となる。ここで、係数 $\beta_1, \beta_2, \beta_3, \beta_4$ は信号値の分布関数に依存する定数であり、一様分布ならば 12^{-1} となることが知られている [6].

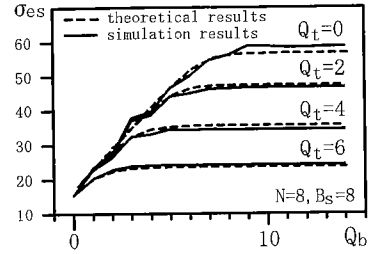


図4 誤差エネルギーと対数ステップの関係
Fig.4 Relation between energy of the error and the log-step.

3.4.3 誤差エネルギーと \bar{B}_t および \bar{B}_b の関係

図4の破線は $B_s = 8$ の場合の式 (39) の関係を示す。但し、誤差エネルギー E_{es} を次式、

$$\begin{aligned} \sigma_{es} &= 10 \log_{10} \frac{(2^{B_s} - 1)^2}{E_{es}} \\ &= -10 \log_{10} E_{es} + 20 \log_{10} (2^{B_s} - 1) \quad (40) \end{aligned}$$

で定義される PSNR (Peak Signal-to-Noise Ratio) で示した。同図より、 Q_b がある値以下では σ_{es} は Q_b に比例するが、ある値以上では σ_{es} は Q_b ではなく Q_t に依存することがわかる。これは、式 (39) 全体に対し、第3項および第4項は無視できるほど小さく、第1項又は第2項のいずれかが支配的になることを表している。そこで、第1項および第2項の指数部分の大小関係に着目すると、式 (39) は、

$$\begin{aligned} \text{(i)} \quad & -Q_b + B_s - 1 > Q_t \text{ のとき,} \\ E_{es} & \approx \beta_1 \cdot 2^{2(-Q_b+B_s-1)} \quad (41) \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad & -Q_b + B_s - 1 < Q_t \text{ のとき,} \\ E_{es} & \approx \beta_2 \cdot 2^{2(Q_t)} \quad (42) \end{aligned}$$

と表すことができる。ここで、式 (41), (42) に式 (36), (37) を代入すると、

$$\begin{aligned} \text{(i)} \quad & \bar{B}_b < \bar{B}_t + \log_2 \frac{\alpha_b}{\alpha_t} + B_s - 1 \text{ のとき,} \\ E_{es} & \approx (\beta_1 \alpha_b^2 \cdot 2^{2(B_s-1)}) \cdot 2^{-2\bar{B}_b} \quad (43) \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad & \bar{B}_b > \bar{B}_t + \log_2 \frac{\alpha_b}{\alpha_t} + B_s - 1 \text{ のとき,} \\ E_{es} & \approx (\beta_2 \alpha_t^2) \cdot 2^{-2\bar{B}_t} \quad (44) \end{aligned}$$

が得られる。以上で式 (11) の誤差エネルギー E_{es} が平均ビット数 \bar{B}_t および \bar{B}_b の関数として表された。

3.5 基底係数の最小ビット数の決定法

これまでの結果、与えられた \bar{B}_t のもとでの \bar{B}_b の

最小値は、式 (39) の第 1 項が第 2 項に対して無視できるほど小さいときのときの \bar{B}_b であることがわかる。そこで非常に小さい数 ε を用い、このときの両者の比を、

$$\varepsilon = \frac{\beta_2 2^{2(-Q_{b,\min} + B_s - 1)}}{\beta_1 2^{2Q_t}} \quad (45)$$

とすると、これを変形することで、

$$Q_{b,\min} = -Q_t + B_s - 1 + \frac{1}{2} \log_2 \frac{\beta_2}{\beta_1 \varepsilon} \quad (46)$$

となる。更に式 (36), (37) を代入すると、

$$\bar{B}_{b,\min} = \bar{B}_t + B_s - 1 + \frac{1}{2} \log_2 \frac{\alpha_b^2 \beta_2}{\alpha_t^2 \beta_1 \varepsilon} \quad (47)$$

となり、基底係数の最小ビット数の推定式が得られ、 $\bar{B}_{b,\min}$ は \bar{B}_t および B_s に比例することがわかる。以上で、変換係数の量子化を考慮した場合の、基底係数の最小ビット数が理論的に明らかになった。

これまで基底係数に要求されていたビット数は、変換係数の量子化を行わないという仮定の下で算出されていた [2]~[5]。しかし、実際の符号化では変換係数を量子化することでデータ圧縮が行われる。また、この場合に必要となる最小ビット数は本論文により式 (47) として明らかになった。従って、従来報告されたビット数は式 (47) の必要最小ビット数に対して過剰となり、この不要なビット数を削減することで符号化処理の簡略化が可能となる。なお、具体的な符号化例における最小ビット数は 4. にて算出する。

3.6 帯域ごとにステップサイズが異なる場合

誤差エネルギーを最小化した結果、式 (20) に示されるように、変換係数のステップサイズ $q_t(f)$ は全帯域で等しくなった。しかし、JPEG や MPEG 等の符号化方式では、帯域ごとに異なるステップサイズが用いられる [7], [8]。これは人間の視覚特性を考慮した結果である。このような場合、まず、 N 個の帯域のうちの第 f 帯域のみの変換係数を与えられたステップサイズ $q_t(f)$ により量子化する。このときの基底係数の最小対数ステップは、式 (46) と同様、

$$Q_{b,\min}(f) = -Q_t(f) + B_s - 1 + \frac{1}{2} \log_2 \frac{\beta_2(f)}{\beta_1(f) \varepsilon} \quad (48)$$

となる。また、対数ステップの定義より、

$$Q_b(f) = -\log_2 q_b(f) \quad (49)$$

の関係がある。従って、式 (48), (49) より $Q_b(f)$ を消去すると、このときのステップサイズ $q_b(f)$ が得られ、これを式 (25) に代入すると、各帯域ごとに基底係数の最小ビット数が求められる。

3.7 2次元信号処理への応用

これまでの議論は 1次元信号処理を対象としていた。本方法を 2次元信号処理に応用するには、式 (1), (22) の $T_N[\]$ および $T_N^{-1}[\]$ に代わり次式、

$$T_N[s(n_1, n_2)] = \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} b(n_1, n_2, f_1, f_2) s(n_1, n_2) \quad (f_1 = 0, 1, \dots, N-1, f_2 = 0, 1, \dots, N-1) \quad (50)$$

および、

$$T_N^{-1}[\hat{t}(n_1, n_2)] = \sum_{f_1=0}^{N-1} \sum_{f_2=0}^{N-1} b(n_1, n_2, f_1, f_2) \hat{t}(f_1, f_2) \quad (n_1 = 0, 1, \dots, N-1, n_2 = 0, 1, \dots, N-1) \quad (51)$$

を用いる。ここで、基底係数 $b(n_1, n_2, f_1, f_2)$ には式 (2) に代わり、

$$b(n_1, n_2, f_1, f_2) = c(f_1, f_2) \cos \frac{(2n_1+1)f_1}{2N} \pi \cos \frac{(2n_2+1)f_2}{2N} \pi \quad (52)$$

但し、

$$c(f_1, f_2) = c(f_1)c(f_2)$$

を用いる。

4. シミュレーション

以下、3. で述べた方法の有効性を計算機シミュレーションにより確認する。

4.1 信号モデル

初めに、画像データの特性を近似する信号モデルを用いて実験を行う。ここでは乱数を伝達関数、

$$H(z) = \frac{1}{1 - \rho z^{-1}} \quad (53)$$

に入力したときの出力信号を用いる。但し、出力信号の値は B_s ビットの整数とする。また、 ρ は信号の相関係数であり、値が 1 に近いほどエネルギーが低周波数領域に偏在する。

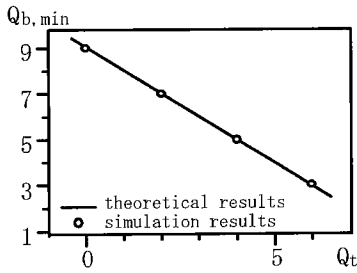


図5 基底係数の最小対数ステップ

Fig.5 Relation between energy of the error and the log-step.

図4の実線は、対数ステップをパラメータとして誤差エネルギー σ_{es} をプロットした結果を示す。但し、 $B_s = 8$, $\rho = 0.9$ とした。また、同図の破線は式(39)の理論値を示す。但し、一様分布を仮定した。なお、誤差エネルギーの理論式は式(39)に示したように信号の値域により決定され、相関係数 ρ には依存しない。同図より、シミュレーション結果は理論値とほぼ等しくなることが確認できる。また、 $Q_b = 14$ のときの σ_{es} に対する差が δ dB 以内となる最小の $Q_{b,min}$ を図4から読み取ると、図5の丸印で示される結果となる。同図の実線は式(46)の理論値であり、実測値とほぼ一致することが確認できる。但し、式(46)では $\beta_1 = \beta_2$ とした。また、非常に小さい数 ε の値は 2^{-4} とした。このとき、

$$\delta = 10 \log_{10}(1 + \varepsilon) \tag{54}$$

より $\delta = 0.3$ dB となる。主観評価上、この程度の違いは検出できない。また、図6には基底係数の最小ビット数 $\bar{B}_{b,min}$ と PSNR σ_{es} の関係を示す。最小ビット数 $\bar{B}_{b,min}$ は式(36)を用いて図5の最小対数ステップ $Q_{b,min}$ から計算した。例えば誤差エネルギーが 32 dB の場合、従来 10 ビット以上必要とされていた基底係数のビット数を 3 ビットとしても、再生信号に含まれる量子化誤差のエネルギーには影響を与えないことがわかる。

4.2 画像データ

つぎに、入力信号に標準画像を用いた場合の実験結果を示す。図7はこの場合の誤差エネルギーと対数ステップの関係を示す。画像データにはインターレース動画画像“flower garden”の1フレームを用いた。但し、フィールド分離はしていない。また、変換係数の量子化には文献[7]に示される帯域ごとに異なるステップ

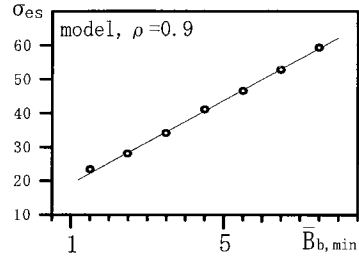


図6 誤差エネルギーと最小ビット数の関係

Fig.6 Relation between energy of the error and the minimum number of bit.

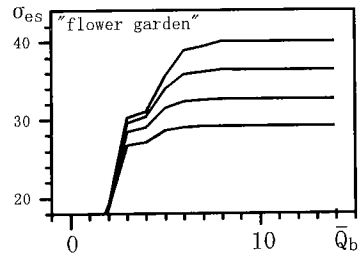


図7 誤差エネルギーと対数ステップの関係

Fig.7 Relation between energy of the error and the log-step.

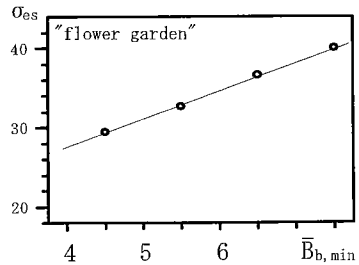


図8 誤差エネルギーと最小ビット数の関係

Fig.8 Relation between energy of the error and the minimum number of bit.

サイズをスカラ倍して用いた。 \bar{Q}_b は $Q_b(f)$ の平均値である。同図より、図4と同様に、 σ_{es} は \bar{Q}_b にほぼ比例するが、 \bar{Q}_b がある値以上では σ_{es} が変化しなくなることがわかる。図8にはこの境界における \bar{Q}_b 、すなわち $\bar{Q}_{b,min}$ と σ_{es} をプロットした。同図より、例えば $\sigma_{es} = 31$ dB の場合、 $\bar{B}_{b,min} = 5$ ビットである。このことは、再生画像の画質が 31 dB である場合は、基底係数を平均 5 ビットで表現しても、それ以上のビット数を割り当てた場合に比べて再生画像の画質が変わらないことを意味する。図9には $\sigma_{es} = 31$ dB

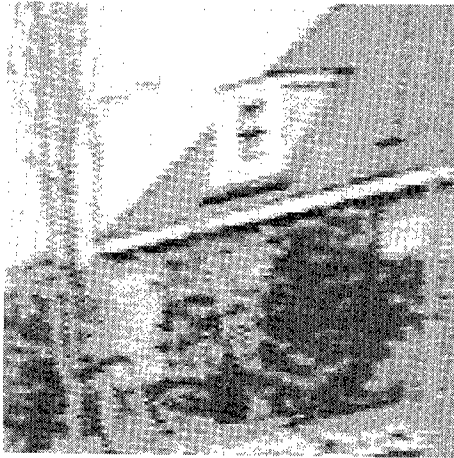
(a) $\bar{B}_b = 5 \text{ bit}$ (b) $\bar{B}_b = 14 \text{ bit}$

図9 再生画像 (インターレース画像)
Fig.9 Decoded image data (interlaced).

のときの再生画像を、 $\bar{B}_b = 5$ および $\bar{B}_b = 14$ それぞれについて示した。両者を比較すると画質の面ではほとんど遜色がないことが確認できる。

基底係数のビット数は従来 10 bit 以上必要とされてきた [7], [8]。しかし、図 8 をもとに再生画像の画質に応じた基底係数の最小ビット数を求めた結果、 σ_{es} が例えば 31 dB のときは、 \bar{B}_b を 5 ビットに低減しても画質には影響を与えず、 σ_{es} が 30 から 40 dB 程度の実用的な場合、基底係数のビット数は従来報告されたビット数以下で十分であることが示された。

5. ま と め

本論文では、画像符号化処理の簡略化を目的として、DCT を用いた変換符号化における基底係数値をできるだけ少ないビット数で表現する方法を提案した。初めに、基底係数および変換係数それぞれのビット数と量子化誤差のエネルギーとの関係を理論的に解析した。これにより、変換係数を量子化した場合の、基底係数のビット数による再生画像の画質劣化の理論的な推定が可能となった。つぎに、得られた結果をもとに変換係数を量子化した場合の、基底係数に要求される必要最小ビット数を理論的に導出した。更に、理論値の妥当性をシミュレーションにより確認した。本論文による考察の結果、従来報告されたビット数は、変換係数を量子化した場合、必要最小ビット数に比べて過剰であり、この過剰分を取り除いても再生画像の画質にはほとんど影響を与えないことが確認された。

今回の評価では、DCT 演算に標準的なアルゴリズムを用いている。今後は、DCT 演算に各種の高速アルゴリズムを用いた場合について考察する予定である。

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謝意を表します。

文 献

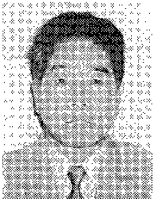
- [1] K.R. Rao and P. Yip, "Discrete Cosine Transform—Algorithms, Advantages, Applications," New York: Academic 1990. (安田 浩, 藤原洋訳, "画像符号化技術—DCT とその国際標準," オーム社, 1992).
- [2] CCITT 勧告 H. 261, "Video CODEC for Audiovisual devices at px64 kbit/s," Dec. 1990.
- [3] 望月孝志, "逆 DCT 有限語長演算の演算誤差解析," 信学論 (A), vol.J77-A, no.3, pp.357-368, March 1994.
- [4] 菊池義池, 赤嶺政巳, "DCT を用いた画像符号化における完全再生可能な DCT 演算精度," 信学論 (A), vol.J74-A, no.7, pp.1116-1120, July 1991.
- [5] 砂川智子, 尾知 博, 金城繁徳, "可変ブロックサイズ DCT において完全再生可能な演算精度の決定法," 信学技報, vol.CAS94-85, Jan. 1995.
- [6] A.N. Akansu and R.A. Haddad, "Multiresolution Signal Decomposition-Transforms, Subbands, Wavelets," New York: Academic, 1992.
- [7] JPEG CD10918-1, "Digital compression coding of continuous-tone still images," JPEG-9-R6, Jan. 1991.
- [8] MPEG CD 11172, "Coding of moving picture and associated audio for digital storage media at up to about 1.5 M bps," MPEG VIDEO COMMITTEE, Dec. 1990.

(平成 7 年 7 月 28 日受付, 12 月 28 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。IEEE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)等。



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」、「マルチレート信号処理」等。画像電子学会、IEEE各会員。