

DCT領域に対するスカラー量子化—パターンベクトル量子化の ハイブリッド符号化

水野 佐年士 デニス・ムワンサ
真弓 政彦 藤村 誠
黒田 英夫

長崎大学工学部

MPEG2 標準規格による 64kbit/sec から 2Mbit/sec 程度の低ビットレートでの高能率符号化の一方式として、画像信号のフレーム間差分信号を DCT 処理して求められる DCT 係数の高周波成分をカットする方式が提案されている。この方式によると、DCT 係数の高周波成分に大きな値を有する場合、復号化した時の画品質劣化が大きくなる。この問題を解決するために、本稿では、高周波成分の+の成分の平均値、-の成分の平均値と0の3値にパターン化してベクトル符号化を行うパターンベクトル量子化を提案する。また、このパターンベクトル量子化を用いる際のコードブックを複数の画像から作成することにより、一般的な画像に対応させた。このコードブックを用いることにより、効率的で、一般的な符号化を行なうことが明らかになった。

A Hybrid Coding Scheme with Scalar/ Pattern Vector Quantization in Interframe DCT Coding

Satoshi MIZUNO Dennis MWANSA
Masahiko MAYUMI Makoto FUJIMURA
Hideo KURODA

Faculty of Engineering, Nagasaki University

In MPEG2, for low bitrate coding of between 64Kbit/sec and 2Mbit/sec, the high frequency components obtained after DCT processing are discarded. However, if the high frequency components contain a significant amount of energy, there is considerable deterioration in the reconstructed image.

In this paper, we propose a method in which the distribution pattern of the high frequency components are vector quantized. We design a codebook for the tri-valued pattern from several training images. A pattern's index and two averages; positive and negative, are transmitted. The reconstructed images are of better quality than the reconstructed images in where the high frequency components are simply discarded while they are of comparable quality to images where the high frequency components are used but with much lower bitrates.

1 まえがき

国際標準化方式 MPEG2 ではフレーム間差分を離散コサイン変換 (DCT) する符号化方式が採用されている。この DCT 係数は線形量子化されるが、従来の符号化の一方式として人間の視覚感度が低下する高周波成分を 0 にするものがある。この方式ではかなりの情報量の削減は望めるが、再生画像による雑音が認識され易くなる。特に情報発生量が多い場合、即ち高周波成分まで大きい値を有する場合には、この零化制御による画質劣化が大きくなる。

本稿では、この問題を改善する方法として、DCT 係数の高周波の成分をパターン化したものをベクトル量子化するパターンベクトル量子化 (以下、PVQ) 方式を提案する。具体的には、DCT 係数の高周波成分の + の係数の平均値、- の係数の平均値及び 0 の値のみを用いたパターンベクトル量子化を行なう。受信側には PVQ の出力インデックスとともに、ベクトル量子化される + の係数の平均値と - の係数の平均値を伝送する。以下、第 2 で DCT 符号化方について、第 3 章ではスカラー量子化について、第 4 章ではパターンベクトル量子化に用いるコードブックの設計を、第 5 章でパターンベクトル量子化の説明をする。

2 DCT 符号化方式

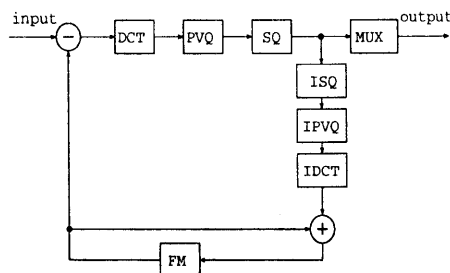
JPEG, MPEG など画像符号化技術の一つとして DCT (Discrete Cosine Transform: 離散コサイン変換) が採用されている。図 1 は、DCT を用いた符号化器の構成図である。

この図の DCT では画像の色や輝度の変化を空間周波数に変換する。以下にその変換式を示す。

与えられたデータの並びを $f(x, y)$ 、DCT の変換係数を $F(u, v)$ とすると、2 次元 DCT は次式で定義される。

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \times \cos \left\{ \frac{(2x+1)u\pi}{16} \right\} \cos \left\{ \frac{(2y+1)v\pi}{16} \right\} \quad (1)$$

$(u, v, x, y = 0, 1, \dots, 7)$



PVQ : pattern VQ (IPVQ : inverse process)
 SQ : scalar quantization
 (ISQ : inverse process)
 FM : frame memory
 MUX : multiplexer

図 1: 符号化器の構成図

ただし、 $C(u)$ 、 $C(v)$ は

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & (\text{for } u = v = 0) \\ 1 & (\text{otherwise}) \end{cases} \quad (2)$$

である。

また DCT の逆変換、IDCT は次式で定義される。

$$f(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u) C(v) F(u, v) \times \cos \left\{ \frac{(2x+1)u\pi}{16} \right\} \cos \left\{ \frac{(2y+1)v\pi}{16} \right\} \quad (3)$$

式 (1)、あるいは式 (3) より、DCT 処理する前とした後の要素の個数は変化しない。本論文では $f(x, y)$ として、現フレームと前フレームの差分を取ったものとする。こうすることによりフレーム間の時間に対する冗長を削減できる。図 4 に現フレームと前フレームの差分信号を DCT 処理した DCT 係数を示す。

図 2 において、 $u=v=0$ の DCT 係数は DC 成分であり、 $f(x, y)$ の平均の値を有する。DC 成分より右下の係数ほど高周波の成分となり、原画像が複雑に変化する部分を表している。一般的に人間の視覚特性は複雑に変化するものは見えにくいことが知られている [1]。したがって、従来の DCT を用いた符号化方式の一方式には、 $u+v \leq 8$ を満たす DCT 係数はスカラー量子化を施して、 $u+v > 8$ を満たす DCT 係数をすべて 0 にする零化制御を施すものがあ

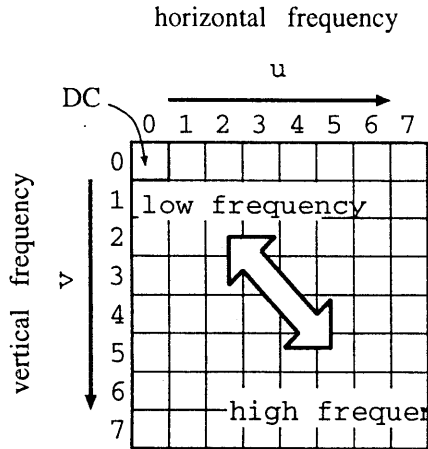


図 2: フレーム間差分の DCT 係数

る。しかし、まえがきでも述べたように原画像の複雑に変化する度合いが大きく、高周波成分まで大きい値を有する場合には画像劣化が大きくなる。この問題を解決するために、低周波成分については従来方式と同様にスカラー量子化して、高周波成分についてはパターンベクトル量子化を用いる。

3 スカラー量子化

DCT 係数の低周波の成分はスカラー量子化を用いて符号化を行なう。スカラー量子化では、DCT 係数のスカラー量子化対象要素にデッドゾーンつき直線量子化を適用する。図 3 にスカラー量子化の入出力特性を示す。スカラー量子化により入力値は、ステップ幅毎の大まかな値として出力される。H.261 勧告では、この図の閾値 T と、ステップ幅 g は、31 種類が用意されていて [2]、入力画像により選択できるようになっている。本稿でも H.261 の仕様を採用する。

4 ベクトル量子化用コードブックの設計

前述のように、人間の視覚感度が低下する DCT 係数の高周波成分を符号化する場合、これらの成分を削除せずに、おおまかに符号化する PVQ 処理を

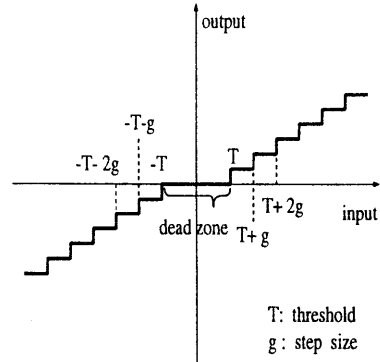


図 3: スカラー量子化の入出力特性

施す。PVQ では、処理の対象となる値の符号および 0 に着目して符号化する。ベクトル量子化では複数の入力信号をまとめてベクトルとして扱うためのコードブックが必要になる。このコードブックは、符号化を行なう以前にオフラインで作成する。コードブックの作成には基本的には LBG アルゴリズム [3] [4] を用いているが、PVQ を行なうための変更を行なったので、以下にそのアルゴリズムを示す。コードブックの要素は、正の符号を表す 1、0 を表す 0、負の符号を表す -1 の 3 値だけから構成される。

1. 初期設定

次元数 $K=18$ 、コードブック数 $N=64$ として、コードブック $C_{64} = \{c_1, \dots, c_{64}\}$ を設計するために、1 つの画像毎にトレーニング系列 $T' = \{t'_1, \dots, t'_{2000}\}$ を用意して、4 種類の画像により $T = \{t_1, \dots, t_{8000}\}$ を作成しておく。ただし、このトレーニング系列とコードブックの要素は、画像の画素値が閾値 $Th = 0.2$ よりも大きいならば 1 に、 $-Th = -0.2$ よりも小さいならば -1、 -0.2 以上 0.2 以下ならば、0 に変換した値である。また、初期ひずみ D_{-1} としてコンピュータを扱う上で無限大と仮定した $D_{-1} = 10^{20}$ 、コードブック作成ループとして $m=0$ 、ひずみ判定用閾値 $\epsilon = 0.001$ とする。

2. ひずみ計算

コードブック作成ループ m のとき、トレーニング系列 T の中の $t_i = \{t_{i1}, \dots, t_{i18}\}$ ($1 \leq i \leq 8000$) に対するコードブック C_{64} の中の $c_j = \{c_{j1}, \dots, c_{j18}\}$ ($1 \leq j \leq 64$) において、平均ひずみが最小となるような

ひずみ最小系列 $P_{64}^{(m)} = \{P_1^{(m)}, \dots, P_{64}^{(m)}\}$ を次式により求める。

$$P_n^{(m)} = \{\vec{t}_l \mid d(\vec{t}_l, \vec{c}_n) \leq d(\vec{t}_l, \vec{c}_m)\} \quad (4)$$

$$(0 \leq l \leq 8000)$$

ただし、 $d(\vec{t}_l, \vec{c}_n)$ は一般には \vec{t}_l と \vec{c}_n の差分の 2 乗誤差が用いられるが [5]、PVQ では次式を用いる。

$$d(\vec{t}_l, \vec{c}_n) = \sum_{k=1}^{18} e_k \quad (5)$$

$$e_k = \begin{cases} 0 & (\vec{t}_l \text{ の要素 } t_{lk} \text{ と } \vec{c}_n \text{ の要素 } c_{nk} \\ & \text{が共に 0、または同符号)} \\ 1 & (t_{lk} \text{ と } c_{nk} \text{ が異符号)} \end{cases} \quad (6)$$

また、平均ひずみを次式で計算する。

$$D^{(m)} = \frac{1}{8000} \sum_{n=1}^{64} \sum_{\vec{t}_l \in P_n^{(m)}} d(\vec{t}_l, \vec{c}_n) \quad (7)$$

3. ひずみ収束判定

式 (7) により計算された $D^{(m)}$ に対して、

$$\frac{D^{(m-1)} - D^{(m)}}{D^{(m)}} < \varepsilon \quad (8)$$

を満たせば、コードブック $C^{(m)}$ が収束したとみなし、処理を終了して $C^{(m)}$ を得る。式 (4) が満たさなければ、次式によりコードブック $C^{(m)}$ を変更する。

$$\vec{c}_i = \frac{1}{\|P_i\|} \sum_{\vec{t}_l \in P_i} \vec{t}_l \quad (9)$$

ただし、 $\|P_i\|$ は、ひずみ最小系列の要素数である。この変更された $C^{(m)}$ に対し、 $m \leftarrow m+1$ として、式 (8) を満たすまで、2. ひずみ計算に戻り、この処理を繰り返す。

5 パターンベクトル量子化

4. コードブックの設計で作成したコードブックを用いてパターンベクトル量子化を行う。ここでは、最適なコードブックの検索の手順と、3 値のコード

ブックだけでは符号化する DCT 係数の符号および 0 のパターンしか復元できないので、付加情報として DCT 係数の正および負の平均値を求める手順を示す。

5.1 最適コードブックの検索

DCT 係数の高周波成分、つまり $\mathbf{X} = \{x_1, \dots, x_N\}$ のなかの $\vec{x}_i = \{x_{i1}, \dots, x_{i18}\}$ に対する、最適コードブックを、作成したコードブック $C_{64} = \{\vec{c}_1, \dots, \vec{c}_{64}\}$ から検索するために次式を最小にするような j を求めて、その数値が DCT 係数の高周波成分を代表させるコードブックのインデックスとなる。ただし、 e_k は式 (6) に示した通りである。

$$\sum_{j=1}^{64} \sum_{k=1}^{18} e_k \quad (10)$$

5.2 DCT 係数の高周波成分の平均化

復号化側でコードブックを用いて情報を復元しても、その値は 1、0、-1 だけから構成される。この 3 つの値からでは、DCT 係数の高周波成分の符号及び 0 のパターンのみしか復元したことになる。その補正として、符号化する際に DCT 係数の高周波成分の正だけと、負だけの平均値も求めて伝送する。

復号化する時には、伝送されたコードブック番号と、コードブックを用いて符号及び 0 のパターンを求めて、1 のところを伝送された正の平均値で置き換え、-1 のところを伝送された負の平均値で置き換える。

6 PVQ の符号化構成

PVQ の符号化手順を以下に示す。符号化する入力画像信号のフレーム間差分をとり、さらに離散コサイン変換を施す。図 4 に本方式におけるフレーム間差分の DCT 係数の量子化処理の割当を示す。8×8 の DCT 係数の内、網目内の 18 個の係数に対して PVQ 処理を施す。この 18 個より低周波の成分は従来通りのスカラー量子化を行ない、18 個よりさらに高周波の成分は零にする。

図 5 に DCT 係数に対する符号化、復号化のブロック図を示す。ここで、SEP は 8×8 の DCT 係数 \vec{X}_i の内の図 4 で示した 18 個の係数、即ち 18 次元のベクトル \vec{H}_i の分離回路、MEAN は + の係数値の平均、- の係数値の平均を計算する平均値演算回路、SQ は

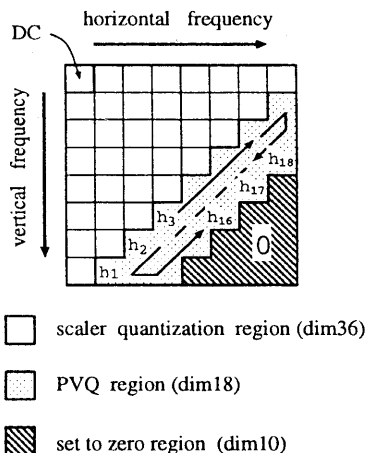
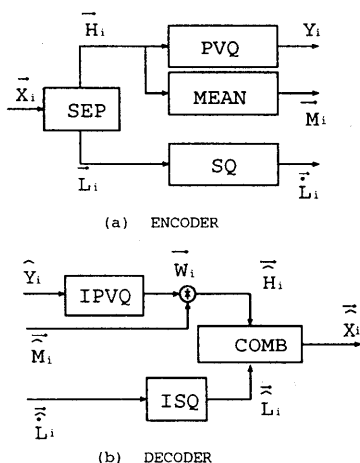


図 4: 本方式におけるフレーム間差分の DCT 係数の処理の割当



- SEP: separation of high and low frequency components
 PVQ: pattern vector quantization (IPVQ: inverse process)
 SQ: scalar quantization (ISQ: inverse process)
 MEAN: mean value computation
 COMB: insertion of high frequency components mean value according to sign pattern

図 5: DCT 係数に対する符号化、復号化のブロック図

スカラー量子化器、PVQ は \bar{H}_i に対する PVQ であり、PVQ の符号化操作は次のように表される。ただし、量子化雑音などは \wedge で表した。

$$\bar{H}_i = \text{SEP}(\bar{X}_i) \quad (11)$$

$$Y_i = \text{PVQ}(\bar{H}_i) \quad (12)$$

$$\bar{M}_i = \text{MEAN}(\bar{H}_i) \quad (13)$$

$$\text{SQ}(\bar{L}_i) = L_i \quad (14)$$

図 5(b) の COMB は高周波成分結合器、IPVQ は逆 PVQ、 \oplus は \bar{W}_i の + の符号と \bar{M}_i の + 係数の平均値と、 \bar{W}_i の - の符号と \bar{M}_i の - の係数の平均値との乗算器であり、PVQ の復号化操作は次のように表される。

$$\bar{W}_i = \text{IPVQ}(Y_i) \quad (15)$$

$$\bar{L}_i = \text{ISQ}(L_i) \quad (16)$$

$$\bar{H}_i = \bar{W}_i \oplus \bar{M}_i \quad (17)$$

$$\bar{X}_i = \text{COMB}(\bar{H}_i, \bar{L}_i) \quad (18)$$

7 シミュレーション実験

シミュレーション実験に用いた画像は標準化作業で用いられている CIF 画像 (352 画素 \times 288 ライン) での輝度信号 (256 階調、黒=0 ~ 白=255) のみを対象とし、画像は "clair", "sales man", "Miss. America", "Blue" の 4 種類を用いて、これらの画像よりコードブックを作成した。

7.1 画像の種類による比較

ベクトル量子化は、符号化を行う画像から作成したコードブックを用いて量子化を行うと SN 比などは良好となるが、他の画像をそのコードブックを用いて量子化を行うと復号化したときの画像劣化が大きくなる [5] [6]。その画像劣化を少なくするために、本論文ではコードブックを 4 種類の画像から作成することにより、コードブックに一般性を持たせた。図 6 が、そのコードブックを用いて、画像の種類別に PVQ を行い、SN 比を比較した図である。コードブック数は、64 に固定した。この図 6 によると、画像の種類を変えて PVQ 処理を施しても、SN 比が改善されているのがわかる。

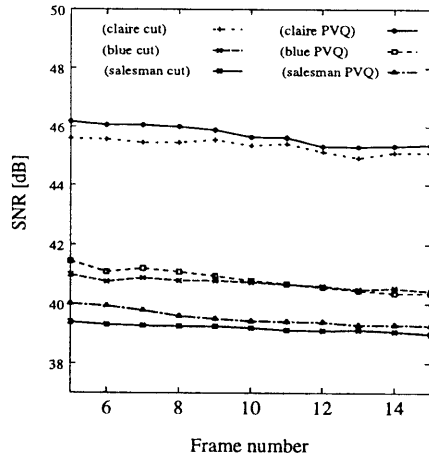


図 6: 画像の種類別の SN 比

7.2 コードブック数による比較

ベクトル量子化では、符号化特性がコードブック数に依存するので、コードブック数は 64 の他にも 128、256、512 のコードブックを作成した。

図 7 がコードブック数を変化させた時の SN 比を比較した図である。この図で、(normal)、(cut)、(cb) はそれぞれ、高周波成分をカットしない画像、高周波成分をカットした画像、高周波成分に PVQ 処理を施した時のコードブック数を変化させた時の画像の SN 比である。

この図 7 によると、コードブック数を増やすことにより、入力値に対する、より最適なコードブックが増えて、SN 比が改善されているのがわかる。

8 むすび

本稿ではフレーム間 DCT 符号化における高周波成分を粗量子化し、かつ情報発生量を少なくする方式として、DCT 係数の各要素の符号に着目したパターンベクトル量子化を提案した。この PVQ により高周波成分の効率的な符号化が行なえることが明らかになった。

シミュレーション実験では 4 種類の画像から作成したコードブックを用いて画像別に PVQ 処理を施した SN 比の比較を行ない、どの画像に対してもほぼ同等の改善が得られた。また、コードブック数を

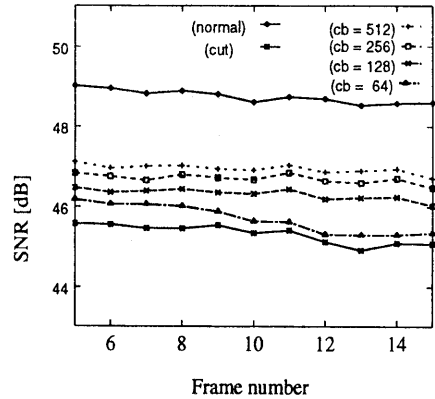


図 7: コードブック数を変化させた時の SN 比の比較

変化させた時の SN 比の比較も行ない、コードブック数を増やすことにより、SN 比を改善できた。今後は、情報発生量の比較に関する検討を行なう予定である。

参考文献

- [1] 松本, 中島, "映像信号の高能率符号化", テレビジョン学会誌, vol.48, No.3, pp274-282(1994-3)
- [2] 安田, "MPEG / マルチメディア符号化の国際標準", 丸善株式会社, (1994-9)
- [3] Y.Linde, A.Buzo, R.M.Gray, "An Algorithm for Vector Quantizer Design", IEEE Trans. Comm. COM-28, 1, pp.84-95(1980-1)
- [4] 田崎, 山田 "ベクトル量子化", 電学誌, vol.67, No.5, pp532-536, (1984-5)
- [5] 斎藤, "画像符号化アルゴリズム (III)", テレビジョン学会誌, vol.43, No.11, pp1276-1284(1989-11)
- [6] 伊藤, 内藤, 佐藤, 宇都宮, "画像の子測誤差信号ベクトル量子化法", 信学論, vol.J70-B, No.3, pp337-345(1987-3)