



HOKKAIDO UNIVERSITY

Title	エッジ部分の再現性を考慮したDCT符号化の検討
Author(s)	岩永, 智博; 北島, 秀夫; 白川, 智昭 他
Citation	北海道大学工学部研究報告, 164, 63-67
Issue Date	1993-05-28
Doc URL	https://hdl.handle.net/2115/42364
Type	departmental bulletin paper
File Information	164_63-68.pdf



エッジ部分の再現性を考慮した DCT 符号化の検討

岩 永 智 博 北 島 秀 夫
白 川 智 昭 小 川 吉 彦

(平成 4 年 12 月 25 日受理)

DCT Coding of Images with Improved Edge Reconstruction

Tomohiro IWANAGA, Hideo KITAJIMA, Tomoaki SHIRAKAWA, and Yoshihiko OGAWA

(Received December 25, 1992)

Abstract

This paper proposes new discrete cosine transform coding of images with edge information retained more faithfully.

Edges are often blurred in conventional transform coding.

The proposed method uses a differential filter, the outputs of which are used to determine quantization schemes.

1. ま え が き

静止画像符号化として、離散コサイン変換(Discrete Cosine Transform, 以下 DCT と略す)を基本とした手法が国際標準カラー静止画符号化方式 JPEG(Joint Photographic Expert Group)で用いられている。DCT を用いた符号化では、信号エネルギーが空間周波数の低周波部分に集中することを利用し、高周波成分のデータを削減して高能率な符号化を実現している。しかし、高周波成分に含まれるエッジ部分の情報がなくなるため、エッジがぼやけてしまうという問題がある。そこで本稿ではエッジ部分を検出し、その形状に応じた符号化をおこなってエッジの再現性を向上させる手法について提案する。また、参考文献³⁾に紹介されている、各ブロックの直流成分画像をなめらかに補間拡大する方法を用いて、低ビットレートで符号化したときに発生しやすいブロックノイズの低減を試みた。

2. 離散コサイン変換

画像信号の冗長性を除去する方法として、直交変換がよく用いられる。直交変換の代表的なものとして DCT がある。DCT もっとも理想的な直交変換である KL 変換に近い特性をもっていることが知られており、画像符号化ではよく用いられている。変換式は以下のとおりである。

$$F(u, v) = \frac{4}{N^2} C(u) C(v) \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j, k) \cos\left[\frac{(2j+1)u\pi}{2N}\right] \cos\left[\frac{(2k+1)v\pi}{2N}\right] \quad (1)$$

逆変換は

$$f(j, k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u, v) \cos\left[\frac{(2j+1)u\pi}{2N}\right] \cos\left[\frac{(2k+1)v\pi}{2N}\right] \quad (2)$$

ただし

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

3. エッジ検出

エッジは多くの高周波成分を持っている。しかし、変換符号化では圧縮率を上げると高周波成分が削られるので、エッジのぼけや歪みが目立つ。特に線画を含むような画像ではエッジの劣化は視覚的に目立ちやすい。再生画像の画質を向上させるためにはエッジの形状を検出し、その形状に応じた量子化を行う必要がある。

3.1 差分フィルタによるエッジ検出

エッジ検出の方法として差分フィルタを用いる。原画像と補間拡大画像との差分画像を 8×8 の小ブロックに分割し、各ブロックに対して 3×3 のSobelの5レベル単純マスクをかけて差分値が最大の値をとるものをその画素の差分値とする。輪郭や線画を含んだブロック内の差分値は大きいので、そのブロック内の和も大きくなる。

ブロック内の差分値の和を S 、しきい値を $Th_1, Th_2 (Th_1 < Th_2)$ と決めて、次のように3つの領域に分ける。

$S < Th_1$: 平坦な部分

$Th_1 < S < Th_2$: 緩やかな部分

$S > Th_2$: 輪郭・線分

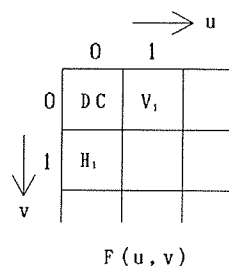
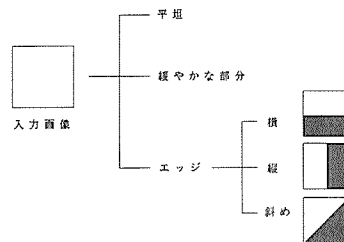


図1 DCT 係数

3.2 エッジ形状の分類

次に、輪郭部分として検出されたブロックに対してエッジパターンの検出を行う。画像を小ブロックに分割してDCTをおこなったとき、そのブロック内のエッジの状態によってDCT係数の分布が変わってくる。水平方向、垂直方向のエッジに対しては図1のDCT係数(H_1, V_1)に特徴があらわれる²⁾。そこでエッジパターンの検出は、低周波領域のDCT係数(H_1, V_1)を用いて3種類のパターンに分類する。

$|H_1| > 0 \quad |V_1| = 0$ のとき : 横エッジ
 $|H_1| = 0 \quad |V_1| > 0$ のとき : 縦エッジ
 $|H_1| > 0 \quad |V_1| > 0$ のとき : 斜めエッジ



以上の方法によって、ブロックを5つのクラスに分ける。

図2 エッジのクラス分け

4. 最適補間拡大

低ビットレートでDCT係数を符号化したとき、交流成分がカットされ直流成分のみ符号化される傾向があるので、ブロック内の画素値はほぼブロック内平均値となり、隣接ブロックとの輝度差がめだちブロックノイズとなる。そこで各ブロックの平均値よりつくられた縮小画像をなめ

らかに拡大すれば、低ビットレート部分のブロックノイズを減らすことができる。そこで文献³⁾の方法を用いて補間拡大する。N×N画素のブロックに分け、拡大対象ブロックとその上下左右のブロックの平均値をS, A, B, C, D (図3) とすると拡大ブロックの補間値を次式で表す。

$$y_{ij} = a_{ij}A + b_{ij}B + c_{ij}C + d_{ij}D + s_{ij}S \quad (3)$$

原画素 x_{ij} と補間画素 y_{ij} との誤差電力 E は

$$E = E[(x_{ij} - y_{ij})^2]$$

となる。補間画素の平均値と元のブロック内平均値が一致するような条件のもとで誤差電力 E が最小となるようにラグランジェの未定乗数法を用いて係数を決定する。

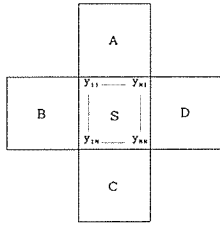


図3 ブロック配置図

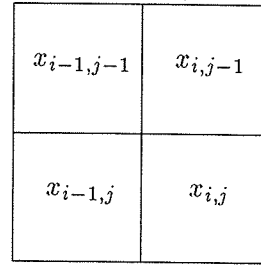


図4 予測画素

5. 符 号 化

本方式の符号化ブロック図を図5に示す。

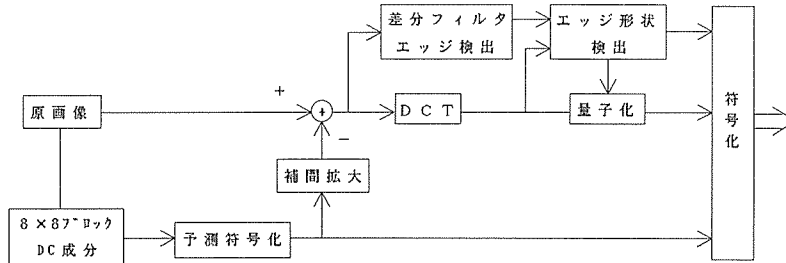


図5 符号化ブロック図

5.1 直流成分の符号化

各ブロックの直流成分間には、多くの相関が残っているので予測符号化によって圧縮をおこなう。各ブロックの直流成分を $x_{i,j}$ (図4) とすると、その縦、横、斜めのブロックの直流成分より次式を用いて平面予測し、原信号との差分をハフマン符号化する。

$$\hat{x}_{i,j} = \rho_h x_{i-1,j} + \rho_v x_{i,j-1} - \rho_{hv} x_{i-1,j-1}$$

ただし、 ρ_h , ρ_v , ρ_{hv} は水平、垂直、斜めの相関係数 $\rho_h = \rho_v = 0.86$, $\rho_{hv} = 0.70$

5.2 交流成分の量子化、符号化

交流成分は、エッジの形状によってブロック内のエネルギー分布が変わる。横エッジでは図6の水平成分に、縦エッジでは垂直成分にエネルギーが集中する。したがって、エッジ情報に基づいて横エッジの場合は水平成分の、縦エッジの場合は垂直成分の量子化ステップを小さくして（細

かく量子化) する。各クラスごとに量子化テーブルを用意して量子化をおこなう。交流成分を図7の様にジグザグスキャンをして, JPEG方式と同様にハフマン符号化を行う。

6. 実験結果

使用した画像は, LENA (8 bit 白黒, 256×256 画素) である。

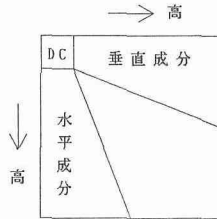


図6 DCT 係数分布

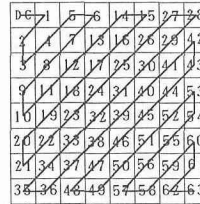


図7 ジグザグスキャン

SEGMENTATION: 本方式

JPEG: JPEG デフォルトの量子化テーブルを用いて符号化

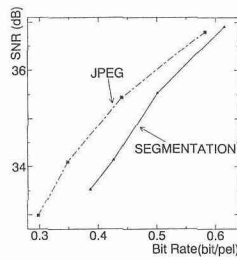


図8 実験結果(LENNE)

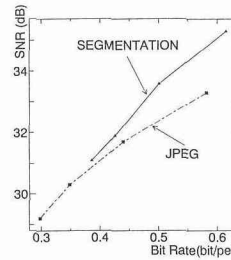
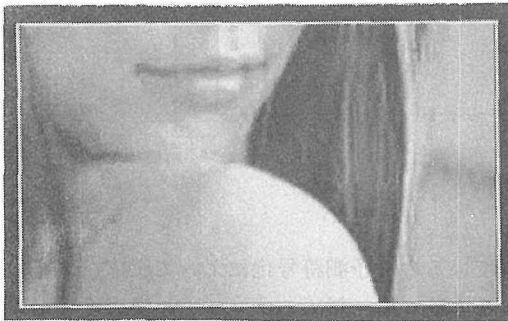


図9 エッジ部分の SN 比(LENNA)



JPEG (0.348 bit/pel)

図10:再生画像



SEGMENTATION (0.386 bit/pel)

図11:再生画像

画像全体の SN 比では JPEG 方式の方がすぐれているが, エッジ部分だけみると本方式の方が SN 比がよくなってエッジ部分のぼけが改善されていることが確認された。また, ブロックノイズも肩の平坦な部分 (図10, 11 参照) が, かなり改善されていることがわかる。

7. ま と め

本方式は簡単ではあるが、エッジ部分の再現性を向上させるのに有効な方法である。また、直流成分画像より補間拡大することによってブロックノイズを（特に平坦な部分）低減することができる。現在画像符号化方式としてよく用いられている JPEG 方式にも本方式を応用することも可能である。しかし、エッジ情報などの付加情報が多く、その効率的な符号化法を検討する必要がある。

参 考 文 献

- 1) 森, 板倉: 画像認識の基礎 [II], オーム社(1990)
- 2) D. S. Kim, S. U. Lee: Image Vector Quantizer Based on a Classification in the DCT Domain, IEEE Trans. Commun. COM-39, 4, pp. 549-556 (Apr. 1991)
- 3) 木原, 伊東, 宇都宮: 隣接ブロックの平均値を用いた最適線形補完法とそのブロック符号化への応用, 信学技報, IE 87-100, pp 71-78, (1987)