

残差周波数変換とその画像圧縮への応用

芦澤 恵太^{†*} 小川 順司^{††} 山谷 克[†]

†名城大学都市情報学部 ††名城大学大学院都市情報学研究科 *2009年4月1日より舞鶴高専

1 はじめに

近年、インターネットや携帯電話の普及に伴い、デジタル画像信号の高効率な圧縮技術が重要な研究課題となっている。デジタル画像信号の圧縮には、ブロック単位の局所的な信号に対して直交変換を適用する方法が広く用いられている [1-4]。これは、入力画像信号に対して直交変換を適用すると、原信号の大まかな特徴を反映する直流成分と画素値の局所的な変化を反映する交流成分とに分離することができるためである。画像内のある画素の状態は周囲の状態との類似性が高いことから、これらの成分の統計的分布には偏りが生じる。このため、視覚特性に応じた量子化が施された変換係数に、連長符号化やハフマン符号化などの可逆圧縮を併用することによって、大幅なデータ圧縮が可能となる。

我々は、ブロック単位の原信号にそのまま直交変換を適用するのではなく、隣接するブロックから予測される信号と原信号との残差に対して直交変換を適用する手法について研究している。本稿では、 8×8 画素ブロック単位の離散コサイン変換 DCT に基づく JPEG 標準方式に着目する。JPEG 標準方式は、最も普及している静止画像の非可逆圧縮方式である。本稿における提案手法では、ブロック境界における高次導関数を近似するように構成される予測信号との残差を利用する。

JPEG 標準方式における DCT を提案手法と置き換えたところ、代表的な標準画像に対して圧縮レートに依存することなく、JPEG 標準方式に対して圧縮画像の PSNR 値が向上することが数値実験により確認できた。

2 予測信号と原信号との残差に対する直交変換

予測信号と原信号との残差に対する直交変換は、原信号に対する変換係数と予測信号に対する変換係数との差をとることで実現できる。簡単のため、1次元離散信号 f_i に対する 2画素単位の DCT および逆 DCT を考える：

$$\begin{pmatrix} F_0 \\ F_1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} f_i \\ f_{i+1} \end{pmatrix}, \begin{pmatrix} f_i \\ f_{i+1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} F_0 \\ F_1 \end{pmatrix}.$$

ここで、交流成分 F_1 と予測信号の交流成分 U_1 との残差を V_1 とすると、予測信号と原信号との残差に対する直交変換、および、その逆変換は以下で与えられる：

$$\begin{pmatrix} F_0 \\ V_1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} f_i \\ f_{i+1} \end{pmatrix} - \begin{pmatrix} 0 \\ U_1 \end{pmatrix},$$

$$\begin{pmatrix} f_i \\ f_{i+1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} F_0 \\ V_1 + U_1 \end{pmatrix}.$$

この変換は、明らかに直交変換ではないことに注意しておこう。以下では、この変換を残差周波数変換と呼ぶこととする。図 1 に、相平面を用いて、残差周波数変換の概念図を示す。

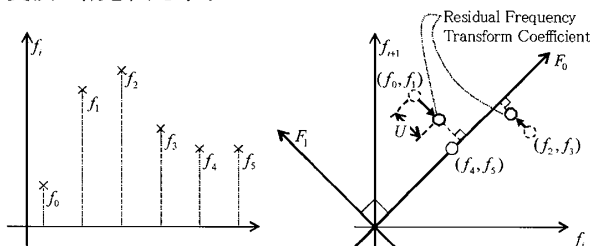


図 1: 残差周波数変換の概念図

3 DCT に基づく残差周波数変換

JPEG 標準方式においては、隣接ブロックの直流成分から構成される双 2 次多項式を用いて、低域の 5 つの交流成分を予測する手法が示されている (以下では QSFIT と表記) [3]。文献 [4] では、ブロック境界における入力信号の勾配の連続性を仮定することで、第 1 行および第 1 列の交流成分に対しては 2 次多項式による予測を、さらに残りの交流成分の予測にはポアソン方程式の解を利用する 2 段階の残差周波数変換 PHLCT が提案されている。本稿では、 8×8 画素ブロック単位の DCT 係数 $F_{(k_1, k_2)}$, $k_1, k_2 = 0, 1, \dots, 7$ における低域 5 交流成分 $(k_1, k_2) = (0, 1), (1, 0), (0, 2), (2, 0), (1, 1)$ に対して、我々が既に提案した勾配近似を用いた交流成分予測手法 [1][2] を複合的に用いることで、3 段階の残差周波数変換を提案する。隣接ブロックの DCT 係数を、例えば処理ブロックの右隣ならば $F_{(k_1, k_2)}^R$ のように、それぞれ右肩に添え字をつけて表記し (図 2 参照)、予測交流成分を $U_{(k_1, k_2)}$ とする。

Residual Frequency Transformation and Its Application to Image Compression

†* Keita ASHIZAWA (ashizawa@urban.meijo-u.ac.jp)

†† Junji OGAWA

† Katsu YAMATANI

Faculty of Urban Science, Meijo University (†)

Graduate School of Urban Science, Meijo University (††)

4-3-3 Nijigaoka, Kani-shi, Gifu 509-0261, Japan

提案手法の順変換および逆変換は以下：

順変換:

- 入力画像を 8×8 画素ブロックに分割.
- 各ブロックで DCT を適用し DCT 係数 $F_{(k_1, k_2)}$ を算出.
- 交流成分 $F_{(0,1)}, F_{(1,0)}$ に対し, 近似成分

$$U_{(0,1)} = -0.142(F_{0,0}^R - F_{(0,0)}^L),$$

$$U_{(1,0)} = -0.142(F_{0,0}^B - F_{(0,0)}^U) \text{ を算出.}$$
- 交流成分 $F_{(0,2)}, F_{(2,0)}$ に対し近似成分を算出:

$$U_{(0,2)} = 0.084(F_{(0,0)}^R - 2F_{(0,0)}^L + F_{(0,0)}^U) + 0.099(F_{(0,1)}^R - F_{(0,1)}^L),$$

$$U_{(2,0)} = 0.084(F_{(0,0)}^B - 2F_{(0,0)}^U + F_{(0,0)}^L) + 0.099(F_{(1,0)}^B - F_{(1,0)}^U).$$
- 交流成分 $F_{(1,1)}$ に対し近似成分を算出:

$$U_{(1,1)} = \frac{1}{8}(F_{(1,0)}^L - F_{(1,0)}^R + F_{(0,1)}^U - F_{(0,1)}^B) + \frac{1}{64}(F_{(0,0)}^{RU} - F_{(0,0)}^{LU} - F_{(0,0)}^{RB} + F_{(0,0)}^{LB}).$$
- 低域の 5 成分 $F_{(0,1)}, F_{(1,0)}, F_{(0,2)}, F_{(2,0)}, F_{(1,1)}$ に対し, 残差 $V_{(k_1, k_2)} = F_{(k_1, k_2)} - U_{(k_1, k_2)}$ を算出.
- 直流成分 $F_{(0,0)}$, 残差 $V_{(k_1, k_2)}$ および交流成分 $F_{(k_1, k_2)}$, $(k_1, k_2) \neq (0, 0), (0, 1), (1, 0), (0, 2), (2, 0), (1, 1)$ を出力.

逆変換:

- 直流成分 $F_{(0,0)}$, 残差 $V_{(k_1, k_2)}$ および交流成分 $F_{(k_1, k_2)}$, $(k_1, k_2) \neq (0, 0), (0, 1), (1, 0), (0, 2), (2, 0), (1, 1)$ を入力.
- 隣接ブロックの直流成分から, $U_{(0,1)}, U_{(1,0)}$ を算出し, $F_{(k_1, k_2)} = V_{(k_1, k_2)} + U_{(k_1, k_2)}$, $(k_1, k_2) = (0, 1), (1, 0)$ を再構成.
- 直流成分および再構成された $F_{(0,1)}, F_{(1,0)}$ から $U_{(0,2)}, U_{(2,0)}$ を算出し, $F_{(k_1, k_2)} = V_{(k_1, k_2)} + U_{(k_1, k_2)}$, $(k_1, k_2) = (0, 2), (2, 0)$ を再構成.
- 直流成分および再構成された $F_{(0,1)}, F_{(1,0)}$ から $U_{(1,1)}$ を算出し, $F_{(1,1)} = V_{(1,1)} + U_{(1,1)}$ を再構成.
- 各ブロックで逆 DCT を適用し入力信号を復号.

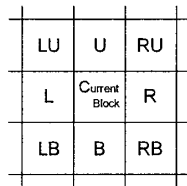


図 2: ブロックの配置

4 数値実験

提案手法の有効性を確認するため, DCT 係数の低域 5 交流成分を, 前節で述べた QSFIT, PHLCT を用いて残差と置き換えた場合に対して, 提案手法との圧縮画像の画質比較を行なう. 全ての手法に対して, JPEG 用の一般的な設定の量子化テーブル (文献 [3], Table 4-1. を参照), および, ハフマン符号化を適用することとする. テスト画像には, 図 3 に示す代表的な標準画像 (512×512pixels, grayscale) を用いる. 客観的な評価指標として次式で定義される PSNR 値を用いた比較を行う: $PSNR = 20 \log_{10} (\text{Peak}/\text{RMSE})$. PSNR 値は, 2 乗誤差を入力信号のピーク値で正規化

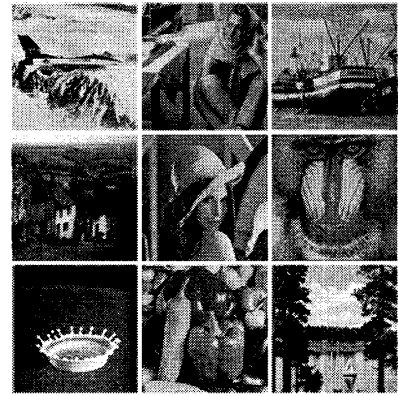


図 3: テスト画像

することによって得られる誤差評価指標であり, ここで Peak は圧縮画像の最大画素値, RMSE は原画像と圧縮画像の 2 乗平均平方根誤差である. PSNR 値の算出においては, 各手法の性能を公平に評価するために, テスト画像の最も外側のブロックは除外し, 9 種類の標準画像の平均値を用いた. 図 4 に JPEG 標準方式に対する各手法の優位性: $PSNR \text{ Gain} = PSNR \text{ 値 (各改善手法を適用)} - PSNR \text{ 値 (JPEG 標準方式)}$ を示す.

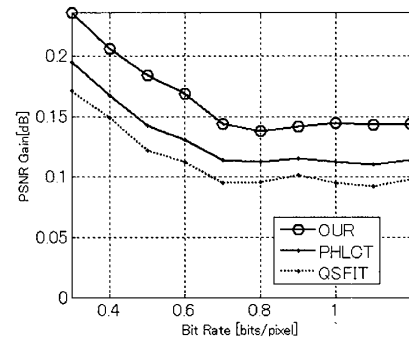


図 4: 圧縮画像の画質評価

5 まとめ

本稿では, 8×8 画素ブロック単位の離散コサイン変換における低域の 5 交流成分に対し, 勾配近似を用いた交流成分予測手法を複合的に用いることによる, 3 段階の残差周波数変換を示した. 標準画像に対する比較数値実験においては, 客観的な評価指標である PSNR 値により, 提案手法の優位性が示された.

参考文献

- [1] 芦澤恵太, 山谷克, 斎藤直樹, “多近傍情報による予測と残差直交変換の階層化およびその画像圧縮への応用,” 日本応用数学会, vol.17, pp.239–257, 2007.
- [2] 芦澤恵太, 山谷克, “高階導関数に着目した段階的な DCT 係数予測と画像圧縮への応用,” 信学論 (A), vol.J91-A, pp.808–816, 2008.
- [3] W.B. Pennebaker and J.L. Mitchell, JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, New York, 1993.
- [4] K. Yamatani and N. Saito, “Improvement of DCT-based compression algorithms using Poisson’s equation,” IEEE Trans. Image Process., vol.15, pp.3672–3689, 2006.