

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

芦澤 恵太^{*} 山谷 克^{*} 斎藤 直樹[†]

^{*} 名城大学都市情報学部 [†] カリフォルニア大学デイヴィス校数学科

概要. 本論文では, 入力画像信号のブロック境界における勾配の高精度な近似を用いて, ブロック内データを予測し, 予測誤差を直交変換することによる画像圧縮アルゴリズムを提案する. 以前筆者らが提案した多重調和局所コサイン変換では, 隣接ブロックの直流成分の差で勾配を近似していたが, 本手法ではブロック境界に近い入力信号を複合的に用いて近似するため, ブロック歪および折り返し歪が軽減され, 客観的・主観的に画質が改善される.

The Method of Hierarchical Multi-neighbor Predictors and Residual Orthogonal Transformations and Its Application to Image Compression

Keita Ashizawa^{*} Katsu Yamatani^{*} Naoki Saito[†]

^{*}Department of Urban Science, Meijo University

[†]Department of Mathematics, University of California, Davis

Abstract. We propose new image compression algorithms by predicting each block using an improved gradient estimation at the block boundary followed by applying an orthogonal transformation to the prediction error. Compared to the previously proposed polyharmonic local cosine transform where the DC components of adjacent blocks were used for the gradient estimation, our new methods directly use multiple combinations of the neighboring pixel values to estimate the gradients at block boundary more accurately. Hence, we can improve the prediction of each block using such gradient information. Consequently, we can improve image reconstruction quality and reduce the blocking artifact and aliasing noise.

1. はじめに

デジタル画像の圧縮原理は, ある画素の状態(明るさや色)は周囲の状態との類似性が高いという性質に基づいている. このため, 画像圧縮には, 局所的な直交変換の貼り合わせによるアルゴリズムが多く用いられる. 例えば, 最も標準的なデジタル画像圧縮形式である JPEG 形式は, 8×8 画素ブロック単位の 2 次元離散コサイン変換(DCT)に基づいている [4, 9]. 一般に, 直交変換によって得られた係数は量子化され, エントロピー符号化により圧縮, 転送される. 係数の量子化方法は, 復元画像の品質および圧縮効率の両面

に大きな影響を及ぼすため，どの圧縮形式についても詳細に検討がなされている．例えば JPEG 形式においては，量子化行列と呼ばれる係数表によって，低周波数成分よりも高周波数成分の方が粗くなるような量子化ステップが定められている [4]．これは，低周波数成分と比較し，高周波数成分を多少粗く表現しても復元画像の主観的な品質にはそれほど影響しないという人間の視覚特性を利用している．しかしながら，画像内に含まれるオブジェクトと背景の境界などのエッジ部分では高周波数成分が重要であり，高周波数成分を粗く量子化し過ぎると，エッジを含むブロックにおいて，折り返し歪と呼ばれるギザギザしたノイズが現れる．また，高圧縮時には，低周波数成分に対しても，量子化ステップを大きくとる必要がある．この場合，変換ブロック毎に独立した基底関数を用いているために，ブロック境界における歪（ブロック歪）が問題となる．高解像度画像の圧縮によく用いられている JPEG2000 形式では，これらの問題を解決するため，高周波数成分の量子化ステップを比較的小さくするとともに，階層化された双直交ウェーブレット変換を採用している [7, Sec. 14]．ワンセグなどのデジタルテレビ映像の圧縮形式 H.264/MPEG-4 AVC においては， 4×4 ブロック単位の DCT および 2×2 ブロック単位のアダマール変換が階層的に適用されている [1, 3]．

JPEG2000 形式のようにウェーブレット変換を用いる方法の他にも，従来の JPEG 形式において，多項式曲面を用いた補間によってブロック歪を軽減する手法が提案されている [2], [4, Sec. 16.1]．これらの方法では，量子化によって失われた低周波数成分を，隣接ブロック間における原画像の連続性に関する仮定の下で，多項式曲面を用いて予測・復元している．一方，斎藤らは，原画像を調和関数とサイン級数に分解し，周波数成分の減衰速度を加速することによって画像の圧縮効率を高める方法（PHLST：多重調和局所サイン変換）を提案している [5, 6]．PHLST では，ブロック境界における画素値を周波数成分とは独立に保持するため，ブロック歪を軽減する効果がある．また，調和関数の代わりにポアソン方程式のノイマン境界値問題の解を用いる方法（PHLCT：多重調和局所コサイン変換）が山谷・斎藤により提案されており，最も標準的な静止画圧縮手法の JPEG と比較し，客観評価 PSNR，および，より主観に近い評価として提案されている MSSIM に関して優れた性能を確認している [11]．これらのアプローチでは，予測関数の変換係数と入力信号の変換係数との残差に着目するため，予測関数の精度が高いほど転送される情報量が少なくなることは明らかである．さらに，階層化された直交変換に基づく周波数成分に対し，局所的な平均画素値による線形予測を用いた方法も高橋・徳永により提案されている [8]．階層的な線形予測による補間によってブロック歪が軽減されることに加え，予測誤差を転送することにより圧縮効率を高める効果がある．本論文では，ブロック歪および折り返し歪の軽減を目的として，多項式を用いた予測手法と PHLCT に着目し， 2×2 画素ブロック単位で階層的に構成される新たな予測手法および画像圧縮アルゴリズムを提案する．

2. 交流成分予測アルゴリズム

本節では，簡単のため，最初に 1 次元信号を対象とした交流成分の予測アルゴリズムについて議論する．2 次元の信号については，1 次元信号に関する予測アルゴリズムのテンソル積によって実現できる．

2.1 準備

Fig. 1 に示すような区間 $I^{(t)} = (t, t+1) = \{y \mid t < y < t+1\}$, $t = 0, \pm 1, \pm 2, \dots$ 上のそれぞれ 2 点 $y = 1/4 + t, 3/4 + t$ で与えられた 1 次元の信号 $f_j^{(t)}$, $j = 0, 1$ について考える．ただし， $f_0^{(t)} := f(\frac{1}{4} + t)$, $f_1^{(t)} := f(\frac{3}{4} + t)$ ．このとき， $f_0^{(t)}$ と $f_1^{(t)}$ の平均値および差分に相当する

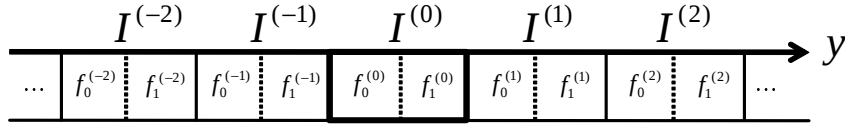


Fig. 1. The configuration of the adjacent blocks of $I^{(0)}$.

直流成分 $F_0^{(t)}$ および交流成分 $F_1^{(t)}$ は次の式で与えられる．

$$(2.1) \quad \begin{pmatrix} F_0^{(t)} \\ F_1^{(t)} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} f_0^{(t)} \\ f_1^{(t)} \end{pmatrix}.$$

また直流成分 $F_0^{(t)}$ および交流成分 $F_1^{(t)}$ が与えられとき，次式の逆変換により入力信号 $f_j^{(t)}$, $j = 0, 1$ を求めることが可能である．

$$(2.2) \quad \begin{pmatrix} f_0^{(t)} \\ f_1^{(t)} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} F_0^{(t)} \\ F_1^{(t)} \end{pmatrix}.$$

次に $U_1^{(t)}$ を予測交流成分とし，交流成分 $F_1^{(t)}$ の代わりに，以下に示す残差 $V_1^{(t)} := F_1^{(t)} - U_1^{(t)}$ を転送することを考える．

$$(2.3) \quad \begin{pmatrix} F_0^{(t)} \\ V_1^{(t)} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} f_0^{(t)} \\ f_1^{(t)} \end{pmatrix} - \begin{pmatrix} 0 \\ U_1^{(t)} \end{pmatrix},$$

$$(2.4) \quad \begin{pmatrix} f_0^{(t)} \\ f_1^{(t)} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} F_0^{(t)} \\ V_1^{(t)} \end{pmatrix} + \begin{pmatrix} U_1^{(t)} \\ -U_1^{(t)} \end{pmatrix}.$$

画像データの符号化と復号化を想定したとき，ある区間 $I^{(t)}$ における予測交流成分 $U_1^{(t)}$ は次の要件を満たす必要がある：

要件 1：符号時： $U_1^{(t)}$ は $f_j^{(s)}$, $s = 0, \pm 1, \pm 2, \dots$ から構成．

要件 2：復号時： $U_1^{(t)}$ は $F_0^{(s)}, V_1^{(s)}$, $s = 0, \pm 1, \pm 2, \dots$ から構成．

要件 3：符号化 (2.3) および復号化 (2.4) で $U_1^{(t)}$ は同一．

2.2 節では，(2.3)，(2.4) および上の要件を満たす 2 通りの交流成分予測手法を提案する．

2.2 勾配推定を用いた予測法とその改善

この節では，従来の PHLCT で各ブロックが 2 標本点からなる場合の勾配推定手法について解説し，さらに新しい勾配推定法を用いた改善案を提案する．

2.2.1 従来の PHLCT における勾配推定

入力信号 $f_j^{(t)}$, $j = 0, 1$ に対し，区間 $I^{(0)}$ における交流成分 $F_1^{(0)}$ を予測することを考える．区間 $I^{(0)}$ の端点における入力信号の勾配 f_y を，隣接区間の入力信号の平均値を用いて次のように近似する．

(2.5)

$$f_y(0) \simeq \tilde{f}_y(0) = \frac{f_0^{(0)} + f_1^{(0)}}{2} - \frac{f_0^{(-1)} + f_1^{(-1)}}{2}, \quad f_y(1) \simeq \tilde{f}_y(1) = \frac{f_0^{(1)} + f_1^{(1)}}{2} - \frac{f_0^{(0)} + f_1^{(0)}}{2}.$$

予測交流成分 ${}^m U_1^{(0)}$ は，(2.5) を満足する 2 次曲線

$$(2.6) \quad u(y) = -\tilde{f}_y(0) \frac{(y-1)^2}{2} + \tilde{f}_y(1) \frac{y^2}{2}$$

を用いて次のように構成される．

$$(2.7) \quad \begin{aligned} {}^m U_1^{(0)} &:= \frac{1}{\sqrt{2}} \left(u\left(\frac{1}{4}\right) - u\left(\frac{3}{4}\right) \right) \\ &= -\frac{1}{4\sqrt{2}} (\tilde{f}_y(0) + \tilde{f}_y(1)) \\ &= \frac{1}{8\sqrt{2}} (f_0^{(-1)} + f_1^{(-1)} - f_0^{(1)} - f_1^{(1)}). \end{aligned}$$

(2.7) による予測は，始端区間および終端区間をのぞき，すべての区間 $I^{(t)}$ に適用可能である (Fig. 2(a) 参照)．ただし，始端区間では $\tilde{f}_y(0) = 0$ ，終端区間では $\tilde{f}_y(1) = 0$ として同様の予測が行なわれる．一方，(2.1)，(2.7) より明らかに

$$(2.8) \quad {}^m U_1^{(0)} = \frac{1}{8} (F_0^{(-1)} - F_0^{(1)})$$

であり，復元時においても，係数 $F_0^{(t)}$ から符号時と同一の ${}^m U_1^{(0)}$ を構成することができる (Fig. 2(b) 参照)．すなわち，3 つの要件すべてが満たされている．

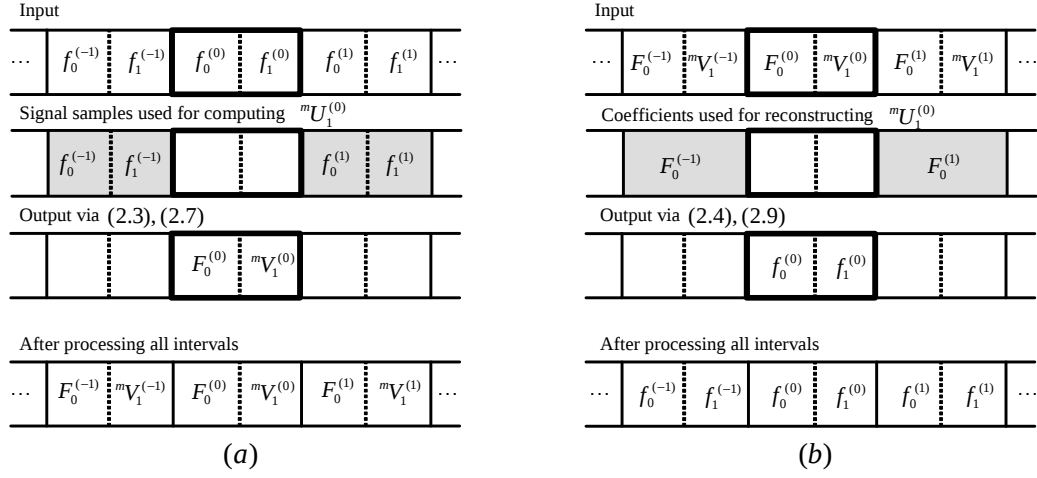


Fig. 2. 1D-PHLCT (2 sample points/block).

2.2.2 勾配推定の改善

本論文では，前節で述べた勾配近似の精度をさらに高めるため，隣接区間の入力信号の境界値を用いた次の近似を提案する．

$$(2.9) \quad f_y(0) \simeq \tilde{f}_y(0) = \frac{f_0^{(0)} - f_1^{(-1)}}{0.5}, \quad f_y(1) \simeq \tilde{f}_y(1) = \frac{f_0^{(1)} - f_1^{(0)}}{0.5}.$$

この場合，(2.6) および (2.9) より，予測交流成分 ${}^bU_1^{(0)}$ は次式で与えられる．

$$(2.10) \quad \begin{aligned} {}^bU_1^{(0)} &= \frac{1}{\sqrt{2}} \left(u\left(\frac{1}{4}\right) - u\left(\frac{3}{4}\right) \right) \\ &= -\frac{1}{4\sqrt{2}} (\tilde{f}_y(0) + \tilde{f}_y(1)) \\ &= -\frac{1}{2\sqrt{2}} (f_0^{(0)} - f_1^{(-1)} + f_0^{(1)} - f_1^{(0)}) \\ &= -\frac{1}{2\sqrt{2}} (\sqrt{2}F_1^{(0)} - f_1^{(-1)} + f_0^{(1)}). \end{aligned}$$

ここで注意しなければならないのは，右辺に係数 $F_1^{(0)}$ が含まれるため，このままでは要件 2 を満足しない．そこで，新たに ${}^bU_1^{(0)} \simeq F_1^{(0)}$ と仮定し，上式から $F_1^{(0)}$ を消去することによって近似的に得られる次の予測式を用いる．

$$(2.11) \quad {}^bU_1^{(0)} := \frac{1}{3\sqrt{2}} (f_1^{(-1)} - f_0^{(1)}) \simeq F_1^{(0)}.$$

しかし残念ながら，すべての区間で上記の (2.11) による変換を行うと，要件 3 を満たすことができない．実際，入力信号 $f_j^{(t)}$ を戻そうとしても，Fig. 3(b) に示したように，境界値

$f_1^{(t-1)}$, $f_0^{(t+1)}$ がわからないため, ${}^bU_1^{(t)}$ を構成できないことに気が付く.

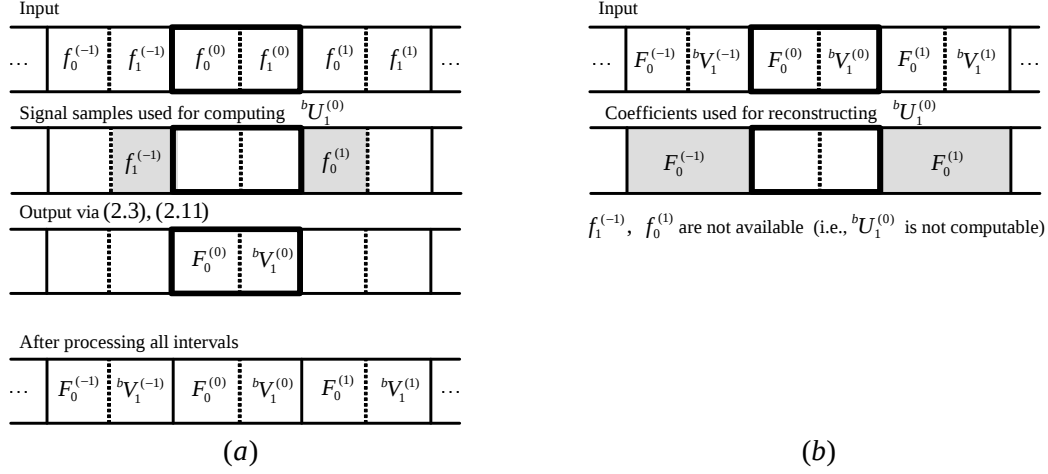


Fig. 3. A straightforward implementation of improved gradient estimators does not necessarily leads to successful reconstruction.

そこで境界値による勾配近似を用いる際には, 要件 2 および要件 3 を満たすために, Fig. 4(a) のように隣接区間では平均値による交流成分予測を行うこととする. その結果, 残差は bV_1 と mV_1 で交互に与えられる. 実際, $F_0^{(-1)}$, ${}^mV_1^{(-1)}$, $F_0^{(0)}$, ${}^bV_1^{(0)}$, $F_0^{(1)}$, ${}^mV_1^{(1)}$ が与えられたとすると, まず区間 $I^{(-1)}$ および区間 $I^{(1)}$ の端点における入力信号の勾配を, (2.5) に示すように, 隣接区間の入力信号の平均値を用いて近似することで, $f_0^{(-1)}$, $f_1^{(-1)}$, $f_0^{(1)}$, $f_1^{(1)}$ が復元される. よって (2.11) より ${}^bU_1^{(0)}$ を構成することができ, $f_0^{(0)}$, $f_1^{(0)}$ を復元することが可能になる (Fig. 4(b) 参照).

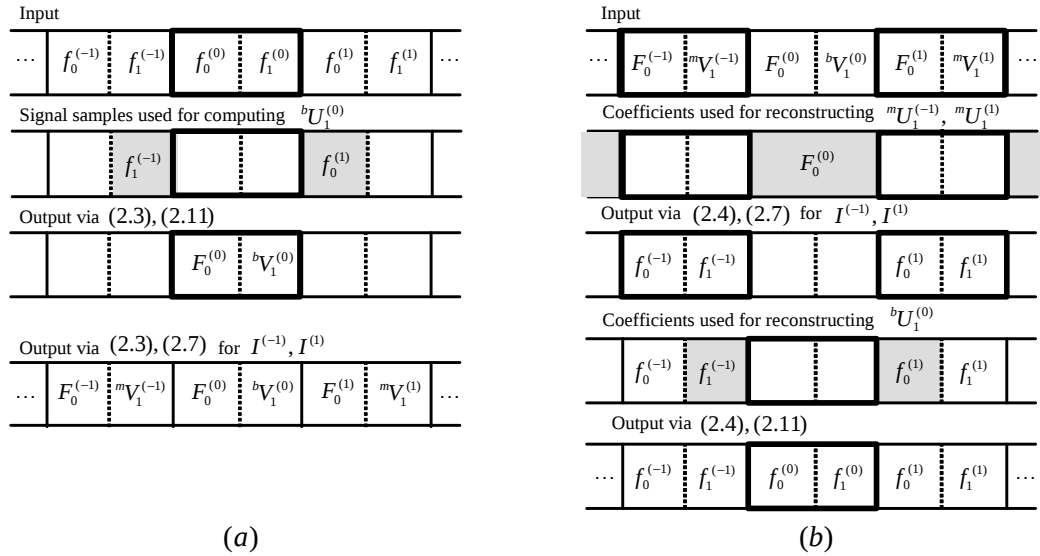


Fig. 4. Our proposed algorithm : 1D-MPROT.

2.3 2次元信号への拡張

次に，領域 $\Omega^{(0,0)} = (0,1)^2 = \{(x,y) \mid 0 < x < 1, 0 < y < 1\}$ 上の4点 (x_i, y_j) で与えられた2次元の入力信号について考える．ただし $x_i = \frac{1}{4} + \frac{i}{2}$, $y_j = \frac{1}{4} + \frac{j}{2}$, $i, j = 0, 1$ ．Fig. 5に示すように $\Omega^{(0,0)}$ の周囲の9つのブロック領域を考え，それぞれの領域 $\Omega^{(s,t)}$ において，入力データを $f_{i,j}^{(s,t)} := f(x_i + s, y_j + t)$, $i, j = 0, 1$ と表記する．なお 2×2 画素ブロック単位では，種々の直交変換，離散コサイン変換，離散フーリエ変換，アダマール変換，ハール変換は一致し，次式で示す変換となる [7]．

$$(2.12) \quad \begin{pmatrix} F_{0,0}^{(s,t)} \\ F_{0,1}^{(s,t)} \\ F_{1,0}^{(s,t)} \\ F_{1,1}^{(s,t)} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} f_{0,0}^{(s,t)} \\ f_{0,1}^{(s,t)} \\ f_{1,0}^{(s,t)} \\ f_{1,1}^{(s,t)} \end{pmatrix}.$$

また，次式の逆変換により，係数 $\{F_{i,j}^{(s,t)}\}$, $i, j = 0, 1$ から入力信号 $\{f_{i,j}^{(s,t)}\}$, $i, j = 0, 1$ を復元することが出来る．

$$(2.13) \quad \begin{pmatrix} f_{0,0}^{(s,t)} \\ f_{0,1}^{(s,t)} \\ f_{1,0}^{(s,t)} \\ f_{1,1}^{(s,t)} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} F_{0,0}^{(s,t)} \\ F_{0,1}^{(s,t)} \\ F_{1,0}^{(s,t)} \\ F_{1,1}^{(s,t)} \end{pmatrix}.$$

ここで，画素データの平均値に相当する係数 $F_{0,0}^{(s,t)}$ を直流成分，水平方向の差分値に対応する係数 $F_{0,1}^{(s,t)}$ ，垂直方向の差分値に対応する係数 $F_{1,0}^{(s,t)}$ ，交差方向の差分値に対応する係数 $F_{1,1}^{(s,t)}$ の3つの成分ををまとめて交流成分と呼ぶ．

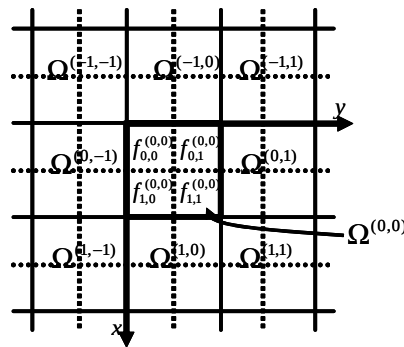


Fig. 5. The configuration of the adjacent blocks of $\Omega^{(0,0)}$.

交流成分 $F_{0,1}^{(0,0)}, F_{1,0}^{(0,0)}, F_{1,1}^{(0,0)}$ を予測するため，2次元信号 $f_{i,j}^{(s,t)}$ に対し，2.2.1節の予測アルゴリズムを，水平（垂直）方向に適用後，得られた係数に対し，さらに垂直（水平）方向に同じ予測アルゴリズムを適用する．これをテンソル積 **PHLCT** (TPHLCT) と呼ぶ

(Fig. 6(a) 参照). 同様に 2.2.2 節の予測アルゴリズムを適用したものを , その一般化性も考慮して , 多近傍情報による予測とその残差に対する直交変換 (Multi-neighbor Predictors and Residual Orthogonal Transformations : MPROT) と呼ぶ (Fig. 6(b) 参照). MPROT 適用時の 2×2 画素データブロックにおける勾配近似の組み合わせは次の 4 通りとなる :

- (A) 水平方向 : 平均値, 垂直方向 : 平均値
- (B) 水平方向 : 境界値, 垂直方向 : 境界値
- (C) 水平方向 : 平均値, 垂直方向 : 境界値
- (D) 水平方向 : 境界値, 垂直方向 : 平均値

ここで , A , B , C , D はそれぞれ Fig. 6 内のブロックに対応している . それぞれのケースに

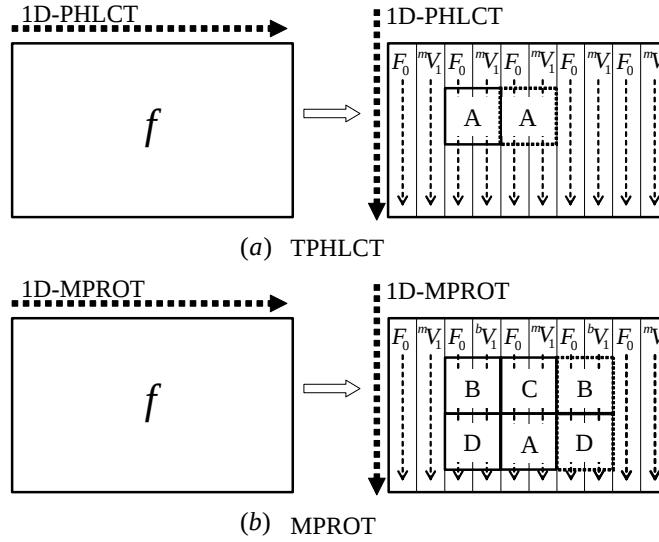


Fig. 6. Our proposed algorithm : (a) TPHLCT; (b) MPROT.

における 2×2 画素データブロックの交流成分を ${}^A F_{i,j}^{(0,0)}$, ${}^B F_{i,j}^{(0,0)}$, ${}^C F_{i,j}^{(0,0)}$, ${}^D F_{i,j}^{(0,0)}$, $(i, j) \neq (0, 0)$ をとすると , 予測交流成分 ${}^A U_{i,j}^{(0,0)}$, ${}^B U_{i,j}^{(0,0)}$, ${}^C U_{i,j}^{(0,0)}$, ${}^D U_{i,j}^{(0,0)}$, $(i, j) \neq (0, 0)$ は , 2 次元入力信号 $f_{i,j}^{(s,t)}$ を用いて , 次の式で表される (Fig. 7 参照).

$$\begin{cases} {}^A U_{0,1}^{(0,0)} = {}^C U_{0,1}^{(0,0)} = \frac{1}{16}(f_{0,0}^{(0,-1)} + f_{0,1}^{(0,-1)} + f_{1,0}^{(0,-1)} + f_{1,1}^{(0,-1)} - f_{0,0}^{(0,1)} - f_{0,1}^{(0,1)} - f_{1,0}^{(0,1)} - f_{1,1}^{(0,1)}), \\ {}^B U_{0,1}^{(0,0)} = {}^D U_{0,1}^{(0,0)} = \frac{1}{6}(f_{0,1}^{(0,-1)} + f_{1,1}^{(0,-1)} - f_{0,0}^{(0,1)} - f_{1,0}^{(0,1)}), \end{cases}$$

$$\begin{cases} {}^A U_{1,0}^{(0,0)} = {}^D U_{1,0}^{(0,0)} = \frac{1}{16}(f_{0,0}^{(-1,0)} + f_{0,1}^{(-1,0)} + f_{1,0}^{(-1,0)} + f_{1,1}^{(-1,0)} - f_{0,0}^{(1,0)} - f_{0,1}^{(1,0)} - f_{1,0}^{(1,0)} - f_{1,1}^{(1,0)}), \\ {}^B U_{1,0}^{(0,0)} = {}^C U_{1,0}^{(0,0)} = \frac{1}{6}(f_{1,0}^{(-1,0)} + f_{1,1}^{(-1,0)} - f_{0,0}^{(1,0)} - f_{0,1}^{(1,0)}), \end{cases}$$

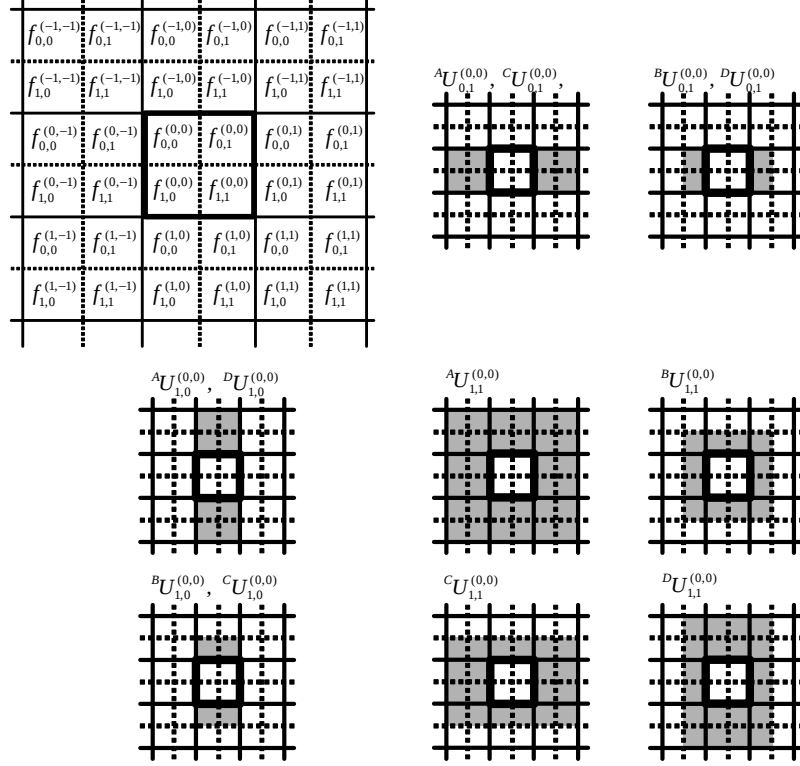


Fig. 7. Signal samples in the adjacent blocks used for estimating gradients. Note that multiple combinations of the adjacent samples are used.

$$\left\{ \begin{aligned}
 A_{U_{1,1}}^{(0,0)} &= \frac{1}{16} (f_{0,0}^{(0,-1)} + f_{0,1}^{(0,-1)} - f_{1,0}^{(0,-1)} - f_{1,1}^{(0,-1)} - f_{0,0}^{(0,1)} - f_{0,1}^{(0,1)} + f_{1,0}^{(0,1)} + f_{1,1}^{(0,1)} \\
 &\quad + f_{0,0}^{(-1,0)} + f_{0,1}^{(-1,0)} - f_{1,0}^{(-1,0)} - f_{1,1}^{(-1,0)} - f_{0,0}^{(1,0)} - f_{0,1}^{(1,0)} + f_{1,0}^{(1,0)} + f_{1,1}^{(1,0)}) \\
 &\quad - \frac{1}{128} (f_{0,0}^{(-1,-1)} + f_{0,1}^{(-1,-1)} + f_{1,0}^{(-1,-1)} + f_{1,1}^{(-1,-1)} - f_{0,0}^{(1,-1)} - f_{0,1}^{(1,-1)} - f_{1,0}^{(1,-1)} - f_{1,1}^{(1,-1)} \\
 &\quad - f_{0,0}^{(-1,1)} - f_{0,1}^{(-1,1)} - f_{1,0}^{(-1,1)} - f_{1,1}^{(-1,1)} + f_{0,0}^{(1,1)} + f_{0,1}^{(1,1)} + f_{1,0}^{(1,1)} + f_{1,1}^{(1,1)}), \\
 B_{U_{1,1}}^{(0,0)} &= \frac{1}{6} (f_{0,1}^{(0,-1)} - f_{1,1}^{(0,-1)} + f_{1,0}^{(0,1)} - f_{0,0}^{(0,1)} + f_{1,0}^{(-1,0)} - f_{1,1}^{(-1,0)} + f_{0,1}^{(1,0)} - f_{0,0}^{(1,0)}) \\
 &\quad - \frac{1}{18} (f_{1,1}^{(-1,-1)} - f_{0,1}^{(-1,-1)} - f_{1,0}^{(-1,1)} + f_{0,0}^{(-1,1)}), \\
 C_{U_{1,1}}^{(0,0)} &= \frac{1}{6} (f_{1,0}^{(-1,0)} + f_{0,1}^{(1,0)} - f_{1,1}^{(-1,0)} - f_{0,0}^{(1,0)}) \\
 &\quad - \frac{1}{16} (f_{0,0}^{(0,-1)} + f_{0,1}^{(0,-1)} - f_{1,0}^{(0,-1)} - f_{1,1}^{(0,-1)} - f_{0,0}^{(0,1)} - f_{0,1}^{(0,1)} + f_{1,0}^{(0,1)} + f_{1,1}^{(0,1)}) \\
 &\quad + \frac{1}{48} (f_{1,0}^{(-1,-1)} + f_{1,1}^{(-1,-1)} - f_{0,0}^{(1,-1)} - f_{0,1}^{(1,-1)} - f_{1,0}^{(-1,1)} - f_{1,1}^{(-1,1)} + f_{0,0}^{(1,1)} + f_{0,1}^{(1,1)}), \\
 D_{U_{1,1}}^{(0,0)} &= \frac{1}{6} (f_{0,1}^{(0,-1)} + f_{1,0}^{(0,1)} - f_{1,1}^{(0,-1)} - f_{0,0}^{(0,1)}) \\
 &\quad - \frac{1}{16} (f_{0,0}^{(-1,0)} + f_{1,0}^{(-1,0)} - f_{0,1}^{(-1,0)} - f_{1,1}^{(-1,0)} - f_{0,0}^{(1,0)} - f_{1,0}^{(1,0)} + f_{0,1}^{(1,0)} + f_{1,1}^{(1,0)}) \\
 &\quad + \frac{1}{48} (f_{0,1}^{(-1,-1)} + f_{1,1}^{(-1,-1)} - f_{0,0}^{(-1,1)} - f_{1,0}^{(-1,1)} - f_{0,1}^{(1,-1)} - f_{1,1}^{(1,-1)} + f_{0,0}^{(1,1)} + f_{1,0}^{(1,1)}).
 \end{aligned} \right.$$

なお上式の計算は，Fig. 6 に示すように (2.3) のテンソル積で表されるため，実際の計算量は低く抑えられる．

3. 階層化

前節までは 2×2 画素ブロック単位の処理について述べてきたが、本節では、 2×2 画素ブロック単位の処理を階層的に適用することによって、圧縮効率を改善する方法について述べる。ここで言う階層化とは、 2×2 画素ブロック単位の処理によって生成された直流成分に対し再帰的に 2×2 画素ブロック単位の処理を繰り返すことを意味する。階層化処理によって、入力信号の局所的な相関をより広い範囲で利用することが可能となる。実際、JPEG などの標準的な画像圧縮手法では、 8×8 ブロック単位の DCT が採用されており、これは 2×2 画素ブロック単位の処理を 3 階層で適用した場合に相当するものと考えられる。さらに、2.3 節で提案した予測手法を 3 階層で適用した場合について、各階層の係数ごとに既存の予測手法と提案手法の予測性能を比較した結果についても述べる。

$2^L \times 2^L$ 画素ブロック単位での階層的な処理を考えると、各階層の入力信号は次のように与えられる。

第 1 階層 : $2^L \times 2^L$ 原画像信号。

第 2 階層 : 第 1 階層を、 2×2 ブロックに分割した各ブロックの直流成分。

⋮

第 ℓ 階層 : 第 $\ell - 1$ 階層を、 2×2 ブロックに分割した各ブロックの直流成分。

⋮

ある階層において入力信号 $2^\ell \times 2^\ell$ に対して MPROT (TPHLCT) を行い、その結果として生成された直流成分からなる信号に対しさらに $2^{\ell-1} \times 2^{\ell-1}$ の上位層を構成するという処理を L 回くりかえす。以下ではこの処理を階層的 MPROT (Hierarchical MPROT: HMPROT)、階層的 TPHLCT (HTPHLCT) と呼ぶ。例えば、 $2^3 \times 2^3$ サイズの画素データに着目すると、Fig. 8 に示すような 3 階層が構成され、各階層のブロックサイズは下から 8×8 、 4×4 、 2×2 となる。この場合、 8×8 画素ブロックで与えられた原画像信号は、Fig. 9 に示す 6 4 個の基底の重ね合わせとして表される。言い換えれば、我々の目的は、直流成分を除く 6 3 個の基底に対応する係数を前節で提案した方法によって予測することである。予測の精度が高ければ、必然的に転送される残差も小さくなる。

我々の提案する二つの交流成分予測手法 HMPROT および HTPHLCT の性能を数値的に検証した結果、および、本予測手法の性能を客観的に評価するため、交流成分予測に関する先行研究との比較結果について述べる。性能比較の対象とする手法には、我々の提案手法と同様に、段階的に生成された直流成分に対する再帰的な処理での交流成分予測を行う高橋・徳永によって提案された手法を採用した。高橋・徳永予測法では、変換ブロックの 4 近傍の直流成分による線形補間を用いて交流成分の予測を行っている [8]。テスト画像としては、各成分に対する予測特性を詳細に調べるため、Fig. 10 に示す特徴的な図形を含むような画像を用いた。Fig. 10 中の (g) Baboon および (h) Barbara は、画像処理アルゴ

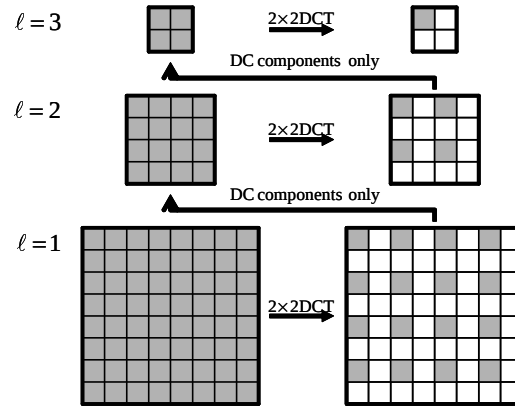


Fig. 8. Hierarchical MPROT for a block of $2^3 \times 2^3$ pixels.

リズムの性能評価において標準的に用いられるテスト画像である． 2×2 DCT を 3 階層で適用したとき，下位の階層の基底は，Fig. 9 に示したように，上位の階層の基底のスケールを落としたものとなっている．したがって，各階層の基底によって表現されるのは，それぞれ解像度の異なった入力信号の差分情報となるため，予測性能の比較は，各階層での交流成分ごとに行うこととする．

まず，Fig. 9 に示すように，直流成分を除く 63 種の基底を H_1 から D_3 の 9 つのグループに分類する．ここで $H_1, H_2, H_3 / V_1, V_2, V_3$ は各階層における水平/垂直方向の差分を意味し， D_1, D_2, D_3 は交差した差分を意味している．Fig. 11 は，各グループにおける係数 F の絶対値の平均，および HMPROT，HTPHLCT，高橋・徳永 (T & T) 予測それぞれの予測法を適用することによって得られた予測後の残差 V の絶対値の平均をプロットしたものである．各グループにおける残差がもとの係数よりも小さければ，予測は成功したものとする．逆に，残差がもとの係数よりも大きくなる場合には，交流成分予測を行わない方が良いであろうと考えられる．実際，Fig. 10 に示す 8 種の画像中 (e) を除く 7 種で，残差がもとの係数より小さくなっており，交流成分予測が成功したことを示している．

ここで特に D_1 グループについて詳細な考察を行うこととする．これは Fig. 7 に示したように，予測に用いる入力信号の組み合わせが，HMPROT と HTPHLCT において異なり，また残差係数の絶対値が D グループの中で最も大きいため，予測による削減の影響が見やすいからである．Fig. 11 からわかるように，画像 (e) 以外の 7 画像に関しては，HMPROT による予測が最も成功している．HTPHLCT の D_1 グループにおける予測は，(b) (c) (d) (e) の 4 画像において失敗しているものの，(a) (f) (g) (h) については成功している．一方， $H_1, H_2, H_3 / V_1, V_2, V_3$ についての結果より，HTPHLCT と高橋・徳永手法が水平/垂直の差分に対しては，まったく同等の予測性能をもつことがわかる．また，すべての予測法が画像 (e) において失敗した原因は，画像 (e) が主に縦線と横線で構成されているにも拘わらず，予測に不必要な周囲のブロック情報まで用いてしまうためであると

[illegible]

Fig. 9. All the basis vectors used in HMPROT and HTPHLCCT for the block of $2^3 \times 2^3$ pixels.

考えられる．特に予測が成功し，残差が小さくなっているのは，(f) Gabor function であり，このことから画像中のグラデーション部分への交流成分予測が効果的であることがわかる．実際に画像処理法の評価に用いられている画像 (g) Baboon, (h) Barbara は，グラデーションをはじめ幾何学的に多様な特徴を含むため，周囲の情報をより多く，よりの確に用いた，交流成分予測の効果が認められる．

そこで次節では，この交流成分の予測性能の違いが，圧縮の効率にどのように影響を及ぼすのかを考察するために，3 階層の HTPHLCT, HMPROT を画像データに対して適用する実験をおこなう．

4. 画像圧縮への応用

本節では，我々の提案する予測アルゴリズム; HTPHLCT, HMPROT を画像圧縮アルゴリズムに適用したときの圧縮性能を，直交変換 DCT に基づく画像圧縮手法として，最も汎用的に普及している画像圧縮アルゴリズムである JPEG standard と比較することで，定量的，視覚的の両面から評価する．なお，ここでいう JPEG standard とは，量子化時に必要となる量子化テーブルには，JPEG 形式で最も標準的に用いられる輝度用量子化テーブルを，符号化アルゴリズムには，Huffman 符号化アルゴリズムを用いたものを指す．

一方，本論文の主な目的は，HTPHLCT, HMPROT による予測の画像圧縮における効果を検証することである．現在，我々の提案する変換法に最適な量子化およびエントロピー符号化のアルゴリズムを検討中であるが，ここでは暫定的に，一様量子化および Huffman 符号化アルゴリズムを用いることにする．

試験用の原画像として，前節で用いた標準テスト画像である Baboon および Barbara を使用する (Fig. 10(g), (h) 参照)．定量的評価には，以下の 2 種類の指標を用いる．

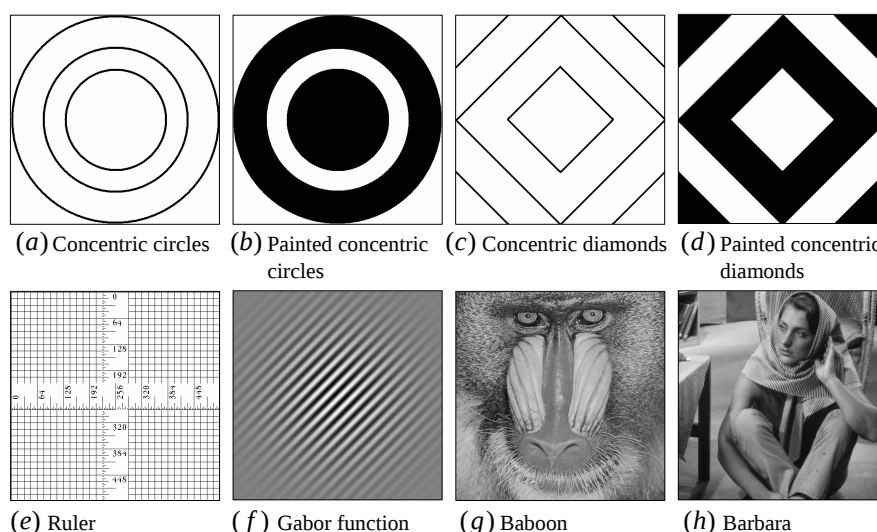
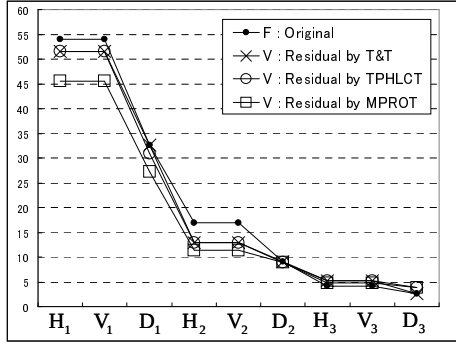
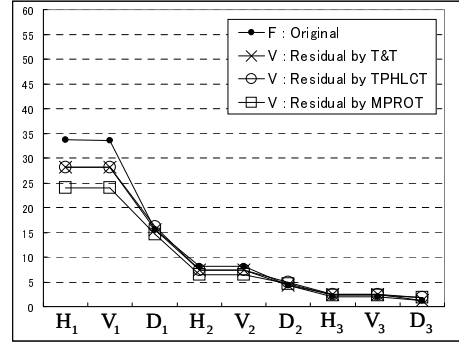


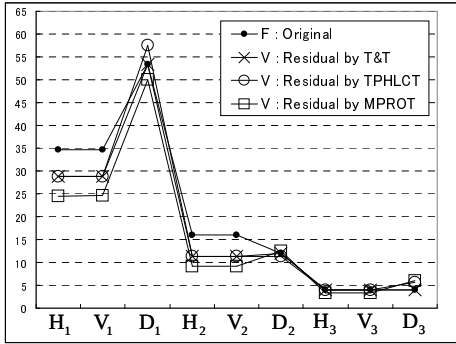
Fig. 10. Original images (512 × 512 pixels, 8 bits/pixel grayscale).



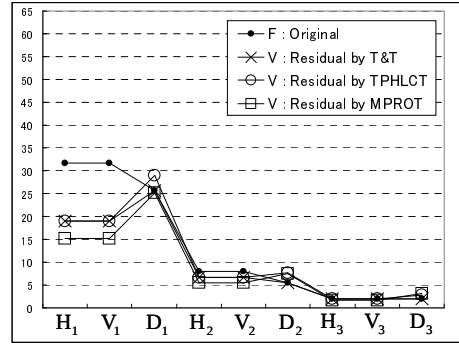
(a) Concentric circles



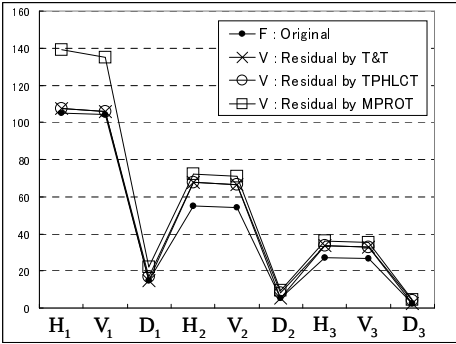
(b) Painted concentric circles



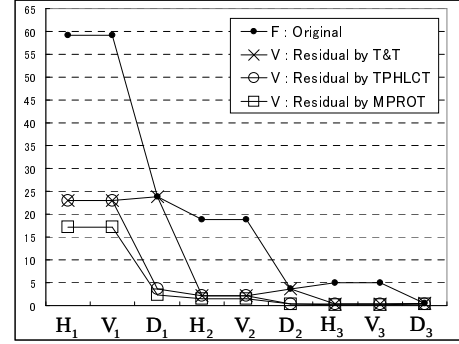
(c) Concentric diamonds



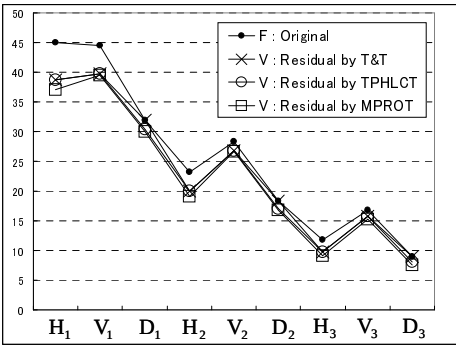
(d) Painted concentric diamonds



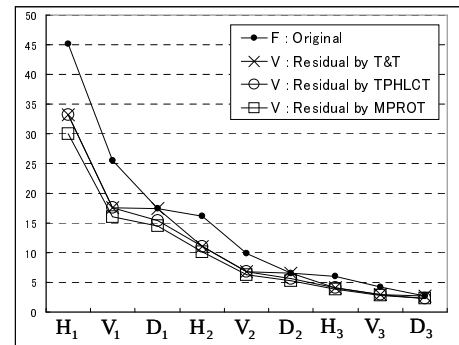
(e) Ruler



(f) Gabor function



(g) Baboon



(h) Barbara

Fig. 11. Averages of the absolute value of residual coefficients of various prediction methods.

1) PSNR: 2乗誤差を入力信号のピーク値で正規化することによって得られる誤差評価指標であり, 次式で定義される:

$$20 \log_{10} \left(\max_{(x,y) \in \Omega} |f^Q(x,y)| / RMSE \right),$$

Ω は一枚の画像全体を表す領域, f^Q は再構成画像データ, $RMSE$ は原画像と再構成画像の2乗平均平方根誤差を表す. 単位は (dB) であり, その値が大きいほど, 良好な結果とみなす.

2) MSSIM: 画像情報に特有の幾何学的構造, すなわち, 一般に, 一枚の画像は有限個のオブジェクトから構成されていることに着目し, 原画像と再構成画像の局所的な類似性に重みを付加した誤差評価指標である. MSSIM の値が大きいほど良好な結果とみなし, その最大値は1である. PSNR と比較し, より視覚的評価に近い結果を与えるものと考えられる [10].

Fig. 12 および Fig. 13 は, それぞれ横軸に圧縮レート, 縦軸に PSNR, MSSIM の値をグラフ化したものである. 同一の圧縮レートに対する再構成画像の精度を, 定量的に評価している. これらの結果から次の3点が読み取れる: a) HMPROT を適用した場合 MSSIM および PSNR とともに通常の DCT(JPEG standard) からの改善がみられる; b) HTPHLCT と HMPROT の二つの手法において, PSNR では差が見えにくい, MSSIM では HMPROT の方が HTPHLCT より画像によっては良好な結果を示す; c) HTPHLCT を適用した場合 MSSIM の値が (b) Barbara では改善されていないが, PSNR については通常の DCT(JPEG standard) からの改善がみられる.

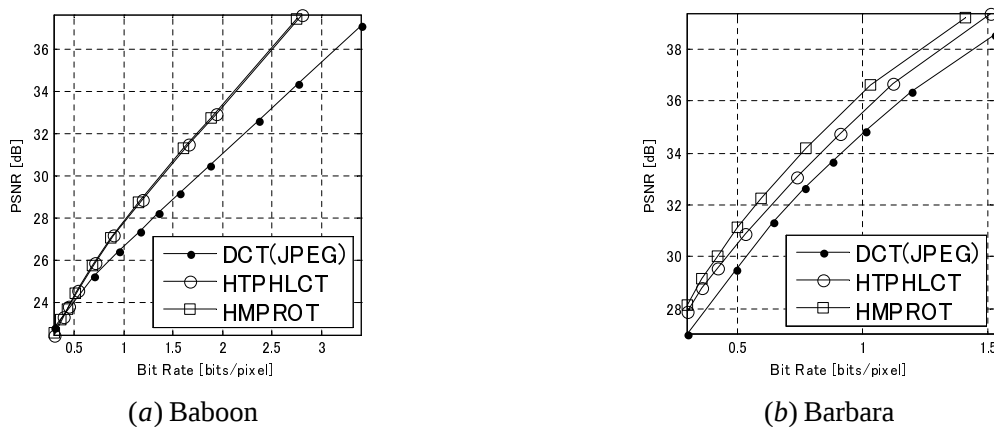
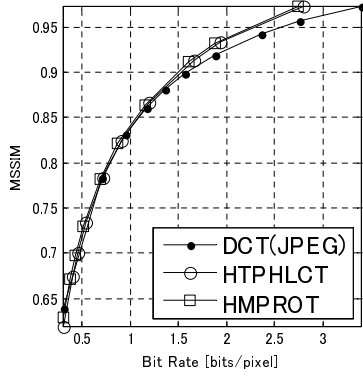
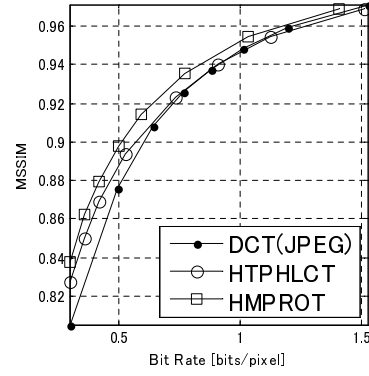


Fig. 12. Comparison of the compression performance with PSNR between HTPHLCT, HMPROT, and DCT(JPEG Baseline method).

そこで, 視覚的評価として, 圧縮レート 0.3 bpp における再構成画像を次の Fig. 14 から Fig. 17 に示す. まず Fig. 14 では, 視覚的比較の材料として, 画像全体ではなく, 画像内で最も観測者の注意を引く Barbara の顔の周囲を, 縦横 4 倍に拡大し表示している.



(a) Baboon



(b) Barbara

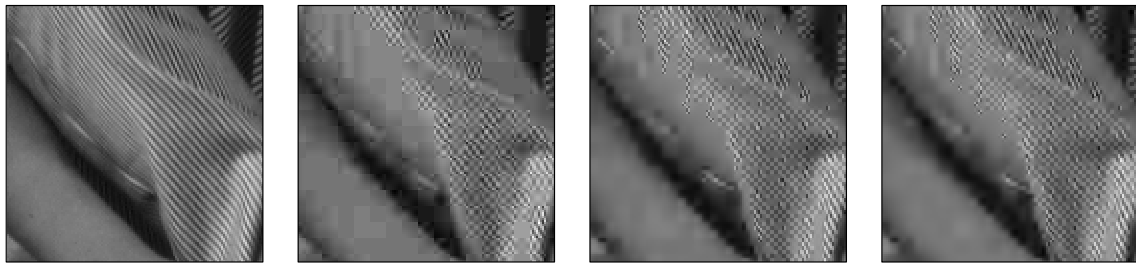
Fig. 13. Comparison of the compression performance with MSSIM between HTPHLCT, HMPROT, and DCT(JPEG Baseline method).

DCT による再構成画像に比べ，HTPHLCT および HMPROT でのブロック歪の軽減などは明らかであり，視覚的評価としても二つの本提案手法発明の効果は確認できる．また，HTPHLCT と HMPROT による再構成画像を比較すると，HMPROT による再構成画像の方が HTPHLCT に比べ，顔の輪郭がはっきりしていることがわかる．次に DCT による



(a) Original (b) DCT, 27.00 dB. (c) HTPHLCT, 27.90 dB. (d) HMPROT, 28.18 dB.

Fig. 14. The face part of the reconstructed Barbara images.



(a) Original (b) DCT, 27.00 dB. (c) HTPHLCT, 27.90 dB. (d) HMPROT, 28.18 dB.

Fig. 15. The arm and leg part of the reconstructed Barbara images.

表現が困難とされている斜めのラインの考察を行うために，Fig. 15 に，手と足による斜

めのラインおよびズボンの縞模様を含む範囲を拡大表示した．前述の顔の輪郭のように，手の輪郭は HMPROT による再構成画像が最もはっきりしている．またズボンの縞模様に関しては，DCT による再構成画像では斜めの縞が崩れてしまっているが，HTPHLCT と HMPROT による再構成画像では縞模様が保存されている．加えて，ズボン部分で布の質感を感じさせるしわに着目すると，DCT ではしわ自体がつぶれてしまう部分でも，HTPHLCT と HMPROT による再構成画像では表現ができている．一方，Fig. 16 および Fig. 17 に示す，テスト画像 Baboon に対する比較でも，Barbara 同様に，DCT による再構成画像に比べブロック歪の軽減は明らかである．また，Fig. 16 での DCT による再構成画像では，鼻孔周辺に顕著な折り返し歪が生じているが，HTPHLCT と HMPROT による再構成画像では，鼻孔部分での折り返し歪は，ほとんど気にならない．また HTPHLCT と HMPROT の違いとしては，Fig. 16(a) Original において，はっきりと確認できる鼻の頭のクレーター状の特徴が，HTPHLCT では跡形もなく周囲に溶け込んでしまっているのに対し，HMPROT では白いブロックとして残っていることが挙げられる．これは Fig. 7 に示したように，HTPHLCT と HMPROT での予測に用いる入力信号の組み合わせが異なるためであると考えられる．しかし，Fig. 16(c), (d) をよくみると，鼻孔周辺の DCT で折り返し歪が生じていた部分において，解像度が周囲と異なる，まるで離れ小島のような小さなタイル状のノイズが観察できる．このノイズは原画像に存在するものが，周囲は溶け込んでしまったにも拘わらず，局所的に保存され構成されたノイズである．そこで原因として考えられるのは，HTPHLCT と HMPROT に対して現在暫定的に行っている一様量子化である．すなわち低周波と高周波を削減するバランスの悪さが，解像度の違う‘離れ小島ノイズ’の原因であると考えられる．このノイズは，Fig. 17(c), (d) において，よりはっきりと確認できる．最後に，Fig. 17 において，頬髭の太さに注目しよう．DCT による再構成画像に比べ，HTPHLCT と HMPROT による再構成画像では，より細かい頬髭が表現できていることがわかる．これは階層的な処理に加え，各階層での近似が上手くいった結果であると考えられる．このような詳細な部分の表現は残しつつ，‘離れ小島ノイズ’を視覚的に気にならないように周囲に溶け込ませていくことは今後の課題の一つである．

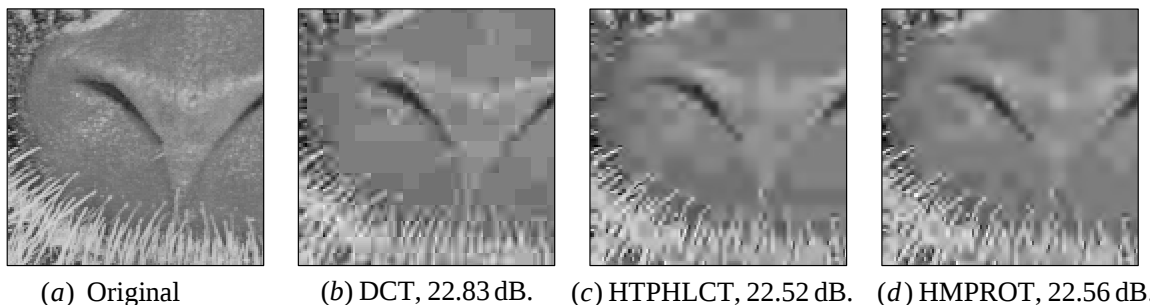


Fig. 16. The nose part of the reconstructed Baboon images.

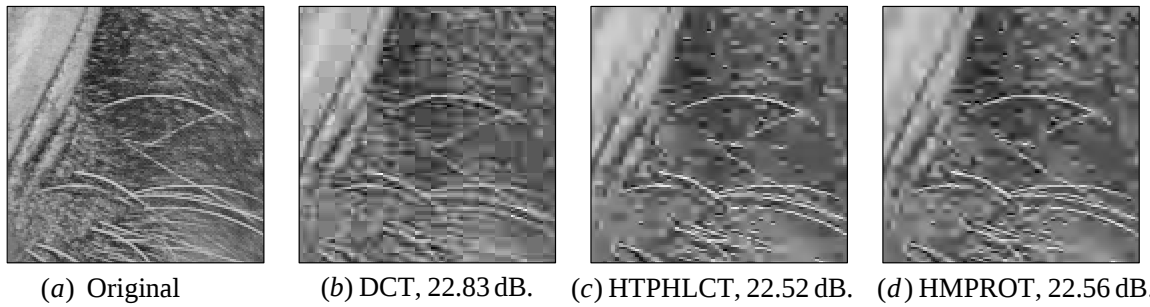


Fig. 17. The whisker part of the reconstructed Baboon images.

5. 結論

本論文では、 2×2 ブロック単位の直交変換によって得られた係数に対し、変換ブロック周囲の入力信号を複合的に用いる新たな予測法を提案した。また、 2×2 ブロック単位の直交変換を階層的に適用した場合にも本予測法が有効であることを数値的に示した。さらに、本予測法に基づく新たな画像圧縮アルゴリズム HMPROT を提案し、その有効性を種々の数値実験によって検証した。HMPROT では、ブロック境界に近い入力信号を複合的に用いて係数を予測し、係数との残差を転送することにより、標準的な画像圧縮方式 JPEG に比べてブロック歪および折り返し歪が軽減され、客観的・主観的にも画質が改善されることを示した。

画像圧縮への応用例としては単純にすべての係数に対してまったく同一の量子化ステップを用いた場合についてのみ述べたが、係数の量子化方法は復元画像の品質および圧縮効率の両面に大きな影響を及ぼすため、本アルゴリズムに適した量子化ルールの構成は、今後の重要な検討課題である。また、本論文では 2×2 画素ブロック単位での直交変換に対して HMPROT の適用を考察しているが、多近傍情報による係数の予測およびその残差に対する直交変換の、例えば 3×3 など、他のブロックサイズへの適用についても現在検討中である。

謝辞 本研究は科学研究補助金（萌芽研究，課題番号：17654022），ONR N00014-06-1-0615, ONR N00014-07-1-0166 および NSF DMS-0410406 の補助を受けて行われた。ここに記して謝意を表する。

参考文献

- [1] 亀山渉, 花村剛, デジタル放送教科書（上）, インプレス R & D, 東京, 2004.
- [2] G. Lakhani, Improved equations for JPEG's blocking artifacts reduction approach, IEEE Trans. Circ. Syst. Vid., 7 (1997), 930–934.
- [3] 大久保榮, 角野眞也, 菊池義浩, 鈴木輝彦, H.264/AVC 放送教科書, インプレス R & D, 東京, 2006.

- [4] W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, New York, 1993.
- [5] N. Saito and J.-F. Remy, A new local sine transform without overlaps: A combination of computational harmonic analysis and PDE, in Wavelets: Applications in Signal and Image Processing X, M. A. Unser, A. Aldroubi, and A. F. Laine, Eds., Proc. SPIE **5207** (2003), 495–506.
- [6] N. Saito and J.-F. Remy, The polyharmonic local sine transform: A new tool for local image analysis and synthesis without edge effect, Applied and Computational Harmonic Analysis, **20** (2006), 41–73.
- [7] K. Sayood, Introduction to data compression, 2nd ed., Morgan Kaufmann Publishers, San Francisco, 2000.
- [8] 高橋隆史, 徳永隆治, 画像ブロックの平均値から交流成分を予測する高速演算アルゴリズム, 電子情報通信学会論文誌, **J81-D-II** (1998), 778–780.
- [9] G. K. Wallace, The JPEG still picture compression standard, IEEE Trans. Consumer Electronics, **38** (1992), xviii–xxxiv.
- [10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: From error measurement to structural similarity, IEEE Trans. Image Process., **13** (2004), 600–613.
- [11] K. Yamatani and N. Saito, Improvement of DCT-based compression algorithms using Poisson’s equation, IEEE Trans. Image Process., **15** (2006), 3672–3689.

芦澤 恵太 (正会員) 〒509-0261 岐阜県可児市虹ヶ丘 4-3-3 ashizawa@urban.meijo-u.ac.jp
 2006 年静岡大学大学院理工学研究科博士後期課程修了。博士 (理学)。現在, 名城大学総合研究所研究員。応用数学の研究に従事。日本数理生物学会会員。

山谷 克 (正会員) 〒509-0261 岐阜県可児市虹ヶ丘 4-3-3 yamatani@urban.meijo-u.ac.jp
 1997 年 大阪大学大学院工学研究科博士後期課程修了。博士 (工学)。現在, 名城大学都市情報学部准教授。数値計算法の研究・開発に従事。日本数学会会員。

斎藤 直樹 (正会員) One Shields Avenue, Davis CA 95616 USA saito@math.ucdavis.edu
 1984 年 東京大学大学院工学系研究科計数工学専攻修士課程修了。日米でのシュルンベルジェ社勤務を経て, 1994 年 イェール大学数学科博士課程修了。Ph.D. 現在, カリフォルニア大学デイヴィス校数学科教授。応用・計算調和解析の研究に従事。IEEE, SIAM 会員。