

# 複素ウェーブレット変換を用いた画像符号化に関する一検討 A Study on Image Coding with Complex Wavelet Transform

高橋 良知†

河村 圭†

渡辺 裕†

Yoshitomo TAKAHASHI

Kei KAWAMURA

Hiroshi WATANABE

## 1. まえがき

離散ウェーブレットを用いた2次元多重解像度解析(2D-DWT: Discrete Wavelet Transform)は、冗長性を持たない変換であるため、画像符号化に広く利用されている。一方で、シフト不変性の欠如、HH成分における $\pm 45^\circ$ 方向のエッジの混在、という画像品質の劣化に影響を与える特性を有する。シフト不変性の欠如は、画像中のエッジ付近のウェーブレット係数がLevel間で大きく変化する原因となる。また、HH成分における方向性の混在は、低ビットレートにおけるチェッカーボードアーティファクトの原因となる[1]。

シフト不変性問題を解決する手法として、複素ウェーブレットが提案されている[2]。複素ウェーブレットを用いた多重解像度解析は、 $m$ 次元の信号に対し $2^m:1$ という冗長性を持たせることで、近似的にシフト不変性が成立し、 $\pm 15^\circ$ 、 $\pm 45^\circ$ 、 $\pm 75^\circ$ の6方向の画像エッジを独立に持つ高域サブバンドを生成する。

著者らは、上記の特性を利用し、複素ウェーブレットを画像符号化へ適用する検討を行なっている。本稿では、複素ウェーブレットのシフト不変性を利用し、補間による高域サブバンドのエントロピー削減手法について検討を行う。

## 2. 2次元複素ウェーブレット多重解像度解析

### 2.1 複素ウェーブレット

画像圧縮に用いる離散ウェーブレットの大部分は実数型ウェーブレットであるが、この実数型ウェーブレットの単一位相がシフト不変性の欠如の原因である。そこで、このシフト不変性の欠如を解決する手法として、式(1)のように、2種類の異なる実数型ウェーブレット $\psi_r(t)$ 、 $\psi_i(t)$ を実数部、虚数部に配置した複素ウェーブレットが提案されている[2]。また、 $\psi_r(t)$ 、 $\psi_i(t)$ に対応するスケーリング関数を $\phi_r(t)$ 、 $\phi_i(t)$ とすると、式(2)により、複素ウェーブレットのスケーリング関数が定義される。

$$\psi_c(t) = \psi_r(t) + j\psi_i(t) \quad (1)$$

$$\phi_c(t) = \phi_r(t) + j\phi_i(t) \quad (2)$$

実数部、虚数部に配置する実数型ウェーブレットには様々な組み合わせが考えられ、ヒルベルト変換ペア[3]や双直交Splineウェーブレットを用いたRI-Splineウェーブレット[4]などが提案されている。

### 2.2 複素ウェーブレット多重解像度解析の構造

複素ウェーブレットによる2次元多重解像度解析(2D-CWT: Complex Wavelet Transform)の処理フローを図1に示す。入力画像に対し、実数部、虚数部、それぞれのウェーブレットを用いて、行方向、列方向の順に1次元多重解像度解析を行なうことで実現される。2D-CWTでは、1つのLevelについて、12種類の高域サブバンドと4種類の低域サブバンドの計16種類のサブバンドが生成されるため、4:1の冗長性を持つ。

2D-CWTの後に得られた12種類の高域サブバンドについて、式(3)、(4)による変形を行い、得られたサブバンドを式(5)、(6)のように、実数部、虚数部に配置することで、2つの複素数が得られる。ここで、 $i=1, 2, 3$ であり、 $\phi_{1,i}$ 、 $\phi_{2,i}$ は、図1における $R^*R$ 、 $I^*I$ の3つの高域サブバンド、 $\phi_{3,i}$ 、 $\phi_{4,i}$ は、 $R^*I$ 、 $I^*R$ の3つの高域サブバンドである。

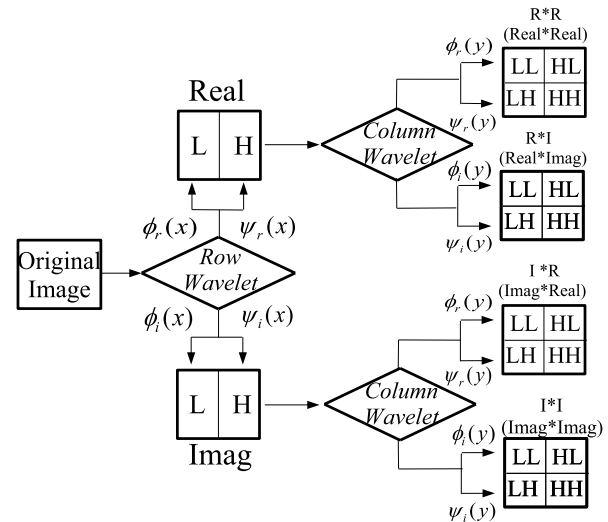


図1: 2D-CWTにおけるLevel1への分解

$$\begin{cases} \psi_i^{+,r}(x, y) = \frac{1}{\sqrt{2}}(\phi_{1,i}(x, y) - \phi_{2,i}(x, y)) \\ \psi_i^{-,r}(x, y) = \frac{1}{\sqrt{2}}(\phi_{1,i}(x, y) + \phi_{2,i}(x, y)) \end{cases} \quad (3)$$

$$\begin{cases} \psi_i^{+,i}(x, y) = \frac{1}{\sqrt{2}}(\phi_{3,i}(x, y) - \phi_{4,i}(x, y)) \\ \psi_i^{-,i}(x, y) = \frac{1}{\sqrt{2}}(\phi_{3,i}(x, y) + \phi_{4,i}(x, y)) \end{cases} \quad (4)$$

$$\psi_i^+(x, y) = \psi_i^{+,r}(x, y) + j\psi_i^{+,i}(x, y) \quad (5)$$

$$\psi_i^-(x, y) = \psi_i^{-,r}(x, y) + j\psi_i^{-,i}(x, y) \quad (6)$$

2つの複素数のうち、 $\psi_i^+(x, y)$ の絶対値は、 $15^\circ, 45^\circ, 75^\circ$ 方向の画像エッジ付近で大きい値を持ち、 $\psi_i^-(x, y)$ は、 $-15^\circ, -45^\circ, -75^\circ$ 方向の画像エッジ付近で大きい値を持つ。

また、式(5)、(6)により得られた2つの複素数において、実数部、虚数部、それぞれ独立ではシフト不変性は成立しないが、複素数の絶対値に関して近似的にシフト不変性は成立する。

## 3. 複素ウェーブレットによる画像符号化

### 3.1 符号量の削減方法

2D-CWTは4:1の冗長性を持つ変換であるため、このまま符号化へ適用すると符号量が増加する。そこで、このような冗長信号に基づく信号圧縮は、解析信号の特性を良く表す係数を順に抽出してゆくことで行なわれる[5]。

この手法を複素ウェーブレットに適用したNoise Shaping法が提案されている[6]。Noise Shapingは、複素ウェーブレット変換後の係数に対し、係数間のエネルギーの移動を行なうことで、非ゼロ係数を減少させる手法である。

次節で提案する高域サブバンドのレベル間相関の除去は、このような手法と組み合わせることで、さらにエントロピーを削減することを想定している。

†早稲田大学大学院 国際情報通信研究科  
Graduate School of Global Information and Telecommunication Studies, Waseda University.

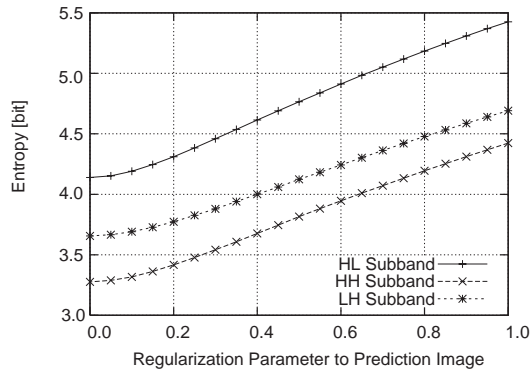


図 2: 2D-DWT の Level1 における高域サブバンドのエントロピー

表 1: 2D-CWT 後の Level1 高域サブバンドのエントロピー

Subband	15°	45°	75°	-15°	-45°	-75°	Total
Predict	3.60	2.46	2.76	3.70	2.72	3.14	18.38
Not Predict	4.50	3.00	3.43	4.61	3.23	3.80	22.57

表 2: 2D-DWT 後の Level1 高域サブバンドのエントロピー

Subband	0°	± 45°	90°	Total
DWT	4.14	3.28	3.66	11.08

### 3.2 離散ウェーブレットにおけるレベル間相関

実数型ウェーブレットにより得られた高域サブバンドは、シフト不変性が欠如しているため高域サブバンドの Level 間の相関性が低いと考えられる。図 2 は、2D-DWT の後、Level2 の高域サブバンドを線形補間して得られた予測バンドと Level1 の高域サブバンドの間の差分係数のエントロピーを示している。入力画像は Lena (512×512)、フィルタは Daubechies 8,8 フィルタを用いた。横軸は、予測係数のエネルギーを正規化する係数である。それぞれの高域サブバンドにおいて、サブバンド間の予測が適切になされなため差分係数のエントロピーは単調に減少している。そして、予測を行わない場合にエントロピーが最小となる。

### 3.3 複素ウェーブレットにおけるレベル間相関

式 (5), (6) の絶対値において、Level  $j$  のウェーブレット係数の分布は、画像中のエッジ中心からの距離  $k$  に応じて、式 (7) のガウス分布で近似できることが知られている [7]。

$$L_j(k) = K e^{-\frac{1}{2}(\frac{k}{\sigma_n})^2} \quad (7)$$

ここで、 $K$  はガウス関数の頂点の高さ、 $\sigma_n$  はガウス関数の分散である。この近似は Level に関係なく成立するが、 $K$ ,  $\sigma_n$  はともに高い Level ほど大きくなる。2D-CWT 後の高域サブバンドは、この特性があるため Level  $j$  の高域サブバンドを Level  $j+1$  の高域サブバンドから予測することが可能である。

ここでは、Level  $j+1$  の高域サブバンドを 1:2 にアップサンプリングし、これにローパスフィルタを施したものを予測バンドとする。これと Level  $j$  の高域サブバンドとの差分を計算する。ローパスフィルタは、

- 平均値フィルタ (線形補間)
- ガウシアンフィルタ (分散: 1.7, 窓長: 11)
- 複素ウェーブレットの実数部の再構成スケーリング関数の 3 種類を比較する。予測は複素数の絶対値で行なうが、再構成時には、複素数の実数部と虚数部の各々の値が必要となる。そこで、複素数の絶対値で求めた残差係数を実数部と虚数部に分配する。このとき、予測前の位相を保存するように、実数部、虚数部へ残差係数を分配することで、完全再構成を保持することができる。また、複素数の絶対値で得られた残差係数は、正負のいずれかの符号が発生するため、残差係数には符号フラグを付加する。

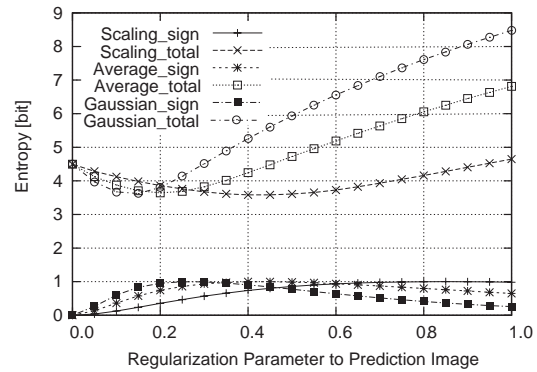


図 3: 2D-CWT の Level1 における 15° 方向高域サブバンドのエントロピー

## 4. 実験結果・考察

本稿では、複素ウェーブレットとして、RI-Spline ウェーブレット [4] を用いて実験を行なった。15° 方向の高域サブバンドの実験結果を図 3 に示す。図 3 は、3 種類のローパスフィルタについて、高域サブバンド間の予測後の残差係数の符号フラグのエントロピーと、実数部、虚数部の残差係数と符号フラグの合計のエントロピーを表している。

いずれのローパスフィルタを用いた場合でも、合計のエントロピーが極小値を持つ。これは、適切な正規化係数を用いて高域サブバンド間の予測を行なうことで、高域サブバンド間の相関を除去し、エントロピーを削減できることを示している。図 3 では、合計のエントロピーの最小値は、ローパスフィルタの種類によらずほぼ同じである。しかし、符号フラグを除いた残差係数のみのエントロピーは、スケーリング関数を用いた場合が最も小さいため、Noise Shaping と組み合わせると他に 2 つのローパスフィルタよりもエントロピーを小さくできると考えられる。また、最小のエントロピーを与える正規化係数の値は、用いるローパスフィルタの種類により異なるが、同一のローパスフィルタの場合、全方向の高域サブバンドにおいてほぼ同じ値となる。

2D-CWT 後の Level1 の全高域サブバンドにおける予測の有無によるエントロピーの違いを表 1 に示す。高域サブバンド間の予測を行なうことで、全体で約 4.0[bit] のエントロピーを削減できる。また、2D-DWT 後の Level1 の全高域サブバンドのエントロピーを表 2 に示す。

## 5. むすび

本稿では、複素ウェーブレットのシフト不変性を利用し、補間を用いた高域サブバンド間予測を行なうことで、高域サブバンドのエントロピーを削減する手法について検討を行なった。実験より、予測サブバンドを用いることで、高域サブバンド間の相関を除去し、エントロピーを削減できることを確認した。

今後は、本手法を Noise Shaping と組み合わせることで、離散ウェーブレット変換を用いた画像符号化方式よりも優れた符号化効率を実現することを目指す。

## 参考文献

- [1] I.W. Selesnick, R.G. Baraniuk, and G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," IEEE Signal Processing Magazine, pp. 123-151, Volume 22, No. 6, Nov. 2005.
- [2] N.G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," Applied Computational Harmonic Anal., Vol. 10, no. 3, pp. 234-253, May 2001
- [3] I.W. Selesnick, "Hilbert Transform Pairs of Wavelet Bases," IEEE Signal Processing Letters, 8-6, pp. 170-173, 2001
- [4] H. Kawabata, H. Toda, Z. Zhang, and H. Fujiwara, "A new complex wavelet transform by using RI-spline wavelet," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), Montreal, vol. 2, pp. 937-940, May 2004
- [5] S.G. Mallat, Z. Zhang, "Matching pursuits with time-frequency dictionaries," IEEE Trans. Signal Processing, Vol. 41, pp. 3397-3415, Dec. 1993.
- [6] T.H. Reeves and N.G. Kingsbury, "Overcomplete image coding using iterative projection-based noise shaping," ICIP 02, Rochester, NY, Sept 2002.
- [7] T.H. Reeves and N.G. Kingsbury, "Prediction of Coefficients from Coarse to Fine Scales in the Complex Wavelet Transform," in IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing, 2000.