

複素ウェーブレットの疎表現における係数選択手法の検討

A Study for Coefficients Selection Method on Sparse Representation of Complex Wavelet

高橋 良知[†]

Yoshitomo TAKAHASHI

渡辺 裕[†]

Hiroshi WATANABE

1. まえがき

離散ウェーブレット変換 (DWT) は方向分離性の乏しい変換である。一方で、複素ウェーブレット変換 (CWT) は高い方向分離性をもつ変換である。方向分離性をもつ変換は、エッジなどの曲線を方向性基底により近似するため、画像構造に基づく近似が可能である。しかし、この変換は変換後にデータ数が増加する冗長変換であることが一般的である。これは冗長性の削減という画像符号化の目的に反するため、画像符号化への応用はほとんどなされていない。冗長変換では疎表現と呼ばれる係数列が存在する。疎表現係数列は信号を最少数の非ゼロ係数で表現したものであり、画像符号化に有効な係数列である。

我々は、方向分離性と冗長性をもつ CWT の疎表現を用いた画像符号化の検討を行っている。本稿では、疎表現の近似解法である l^0 ノルムベース交互射影法に基づき、Noise Shaping における新しい係数選択手法を提案する。また、非ゼロ係数の個数に対する PSNR と全符号量により性能を評価する。

2. 複素ウェーブレット変換

CWT は 2 つの実数ウェーブレットを実部と虚部の基底とした複素数変換である [1]。これはフーリエ変換における正弦波をウェーブレットに置き換えた変換といえる。CWT はシフト不変性が成立し、2 次元 CWT は $\pm 15^\circ$, $\pm 45^\circ$, $\pm 75^\circ$ の 6 方向の画像エッジを異なるサブバンドに分離する。また、CWT は 2 つの最大間引きフィルタバンクを並列に配置した並列木構造により実装される。このため、2 次元の入力信号に対して 4 倍のデータ数が発生する冗長変換となる。

3. 冗長変換の疎表現

3.1 疎表現問題

n 次元ベクトル $\mathbf{y} \in \mathbb{R}^n$ を辞書 $\mathcal{D} = \{\phi_i \in \mathbb{R}^n, i = 1, \dots, m\}$ の線形結合により表現する問題を考える。

[†]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.

$$\mathbf{y} = \sum_i \alpha_i \phi_i \quad (1)$$

冗長変換では $m > n$ であり、辞書 \mathcal{D} は過完備な基底から構成されている。このとき、式 (1) の解 $\alpha \in \mathbb{R}^m$ は無数に存在する。このうち疎表現とは最少数の基底 (非ゼロ係数) を用いて信号 \mathbf{y} を表現した解である。これは式 (2) の最適化問題の解として求められる。

$$\text{Minimize } \|\alpha\|_0 + \lambda \|\mathbf{D}\alpha - \mathbf{y}\|_2^2 \quad (2)$$

式 (2) は誤差を含む場合の疎表現問題である。ここで、 \mathbf{D} は $n \times m$ 行列、 $\|\alpha\|_0$ は α の l^0 ノルムであり、非ゼロ係数の個数である。式 (2) は NP 困難のため、マッチング追跡や基底追跡などの近似解法が必要である。

3.2 l^0 ノルムベース交互射影法

本稿では、近似解法として l^0 ノルムベース交互射影法を用いる。この手法では式 (2) の代わりに式 (3) の問題より近似解を求める。

$$\hat{\alpha}(K) = \arg \min_{\alpha \in C(K)} \|\mathbf{D}\alpha - \mathbf{y}\|_2$$

$$\text{s. t. } C(K) = \{\alpha \in \mathbb{R}^m : \|\alpha\|_0 = K\} \quad (3)$$

解 $\hat{\alpha}(K)$ を求めるため、式 (1) の解集合 $S(\mathbf{D}, \mathbf{y})$ と $C(K)$ の間の交互射影を繰り返す [2]。ここで、 $S(\mathbf{D}, \mathbf{y})$ は式 (1) のアフィン集合で凸集合であるが、 $C(K)$ は凸集合ではない。そのため、収束解は局所解である。 $C(K)$ から凸集合 $S(\mathbf{D}, \mathbf{y})$ への射影は直交射影 $P_{S(\mathbf{D}, \mathbf{y})}^\perp$ が存在する。凸集合 $S(\mathbf{D}, \mathbf{y})$ から $C(K)$ への射影 $ht(\alpha, K)$ は $C(K)$ と $S(\mathbf{D}, \mathbf{y})$ のユークリッド距離を最小にするために最大値から K 個の係数を選択する。

3.3 CWT における Noise Shaping

Kingsbury らは CWT に対して、雑音変形手法である Noise Shaping を適用し疎表現を求めている [3]。Noise Shaping は反復処理に基づく手法であり、反復処理の一部で式 (4) の閾値処理がおこなわれる。

$$\hat{y}_r + j\hat{y}_i = \begin{cases} 0 & \text{if } |y_r + jy_i| < \theta \\ y_r + jy_i & \text{otherwise} \end{cases} \quad (4)$$

ここで、 y_r, y_i は複素数の実部、虚部である。 \hat{y}_r, \hat{y}_i は閾値処理後の実部、虚部である。

4. 提案手法

我々は, CWT における Noise Shaping が l^0 ノルムベース交互射影法と等価な手法であることを示した [4]. 等価性より, Noise Shaping における式 (4) の閾値処理は l^0 ノルムベース交互射影法における $ht(\alpha, K)$ に相当する. $ht(\alpha, K)$ は最大値から K 個の係数を選択する処理であるため, 式 (4) の閾値処理は複素数単位で選択する必要はない. そこで, 我々は実部と虚部に対して, 独立に閾値処理を適用する手法を提案する.

$$\hat{y}_r = \begin{cases} 0 & \text{if } |y_r| < \theta \\ y_r & \text{otherwise} \end{cases} \quad \hat{y}_i = \begin{cases} 0 & \text{if } |y_i| < \theta \\ y_i & \text{otherwise} \end{cases} \quad (5)$$

Noise Shaping はマッチング追跡と同様の貪欲アルゴリズムである. そのため, 基底の選択方法が重要となる. 提案手法では, 実部と虚部の基底を独立に選択することで, 複素数単位で選択する場合と比較し, 誤基底を選択する可能性を低減している.

また, 疎表現の性能は近似解法と辞書の定義により決まる. 方向性基底の複素ウェーブレットを $\psi_j^c = \psi_j^r + j\psi_j^i$ とすると, 辞書 \mathcal{D} は, 従来手法では $\mathcal{D}_c = \{\psi_j^c \in \mathbb{R}^n, j = 1, \dots, 2n\}$, 提案手法では $\mathcal{D}_p = \{\psi_j^r, \psi_j^i \in \mathbb{R}^n, j = 1, \dots, 2n\}$ となる. ψ_j^r と ψ_j^i はヒルベルト変換対であるためシフトの関係にある. このため, 提案手法ではより密な方向性基底の集合を辞書 \mathcal{D}_p として定義していることになる.

5. 実験結果

Barbara (512×512) において, 非ゼロ係数の個数に対する PSNR を図 1 に示す. 提案手法では, DWT や従来手法と比較し, 同数の非ゼロ係数で約 1.0–3.0 [dB] 高い PSNR が得られている. これより, 提案手法は従来手法よりも疎表現の近似解精度が高い. また, 全符号量における比較を図 2 に示す. CWT では各レベルのエントロピーの和に基づき全符号量を求めた. 量子化として, 提案手法にはスカラ量子化, 従来手法にはスカラ量子化または複素数のベクトル量子化を用いた. DWT は算術符号化なしの SPIHT 符号化を行った [5]. 図 2 より, 提案手法は SPIHT 符号化より約 0.2–0.5 [dB] 高い符号化効率となる. また, 提案手法は従来手法のスカラ量子化より高い符号化効率となるが, 従来手法のベクトル量子化よりは性能が劣る. これより, 提案手法の疎表現においても適したベクトル量子化を考慮する必要があり, 今後の課題である.

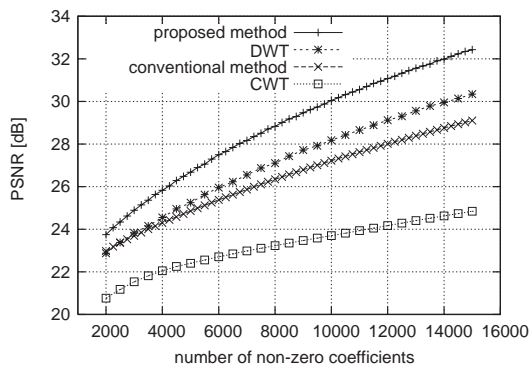


図 1: 非ゼロ係数へのエネルギー集約度

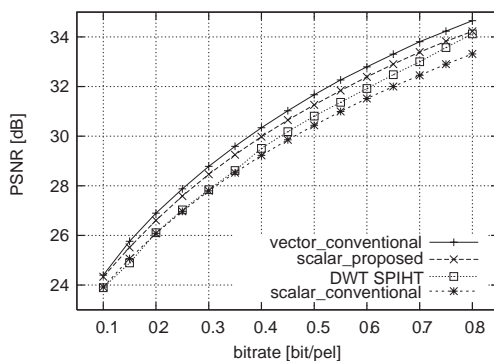


図 2: 全符号量

6. まとめ

本稿では, CWT における疎表現近似手法において新しい係数選択手法を提案した. 非ゼロ係数へのエネルギー集約度と全符号量による実験より, その有効性を確認した.

参考文献

- [1] I.W. Selesnick, *et al.*, “The dual-tree complex wavelet transform,” *IEEE Signal Processing Magazine*, Nov. 2005.
- [2] L. Mancera, *et al.*, “ l_0 -norm-based sparse representation through alternate projections,” *IEEE ICIP*, Oct. 2006.
- [3] N. G. Kingsbury *et al.*, “Iterative image coding with overcomplete complex transform,” *IEEE VCIP*, July 2003.
- [4] 高橋, 渡辺, “双対木複素ウェーブレットにおける疎表現の検討,” *AVM58-06*, Sep. 2007.
- [5] A. Said, *et al.*, “A new, fast, and efficient image codec based on set partitioning in hierarchical trees,” *IEEE Trans. CirSys Video*, June 1996.