

ウェーブレットによる画像処理

Wavelets for Image Processing

この節では、ウェーブレットが画像処理の分野でどのように使われているか、特に画像の近似、圧縮、ノイズ除去、特徴抽出に焦点を当てて解説する。ウェーブレットがこのような様々な画像処理に使われるのは、画像中の重要な特徴、特に鋭いエッジや角のような特異的な情報を、フーリエ変換やコサイン変換等に比べると、うまく捉え、かつコンパクトに表現できるのが主な理由である。

画像処理研究におけるウェーブレットの歴史は古く、1970年代末のマー (Marr) 等の視覚情報処理の研究や、1980年台初頭に提案されたバート (Burt) とエーデルソン (Adelson) のラプラシアン・ピラミッド (Laplacian Pyramid) は、マラー (Mallat) の多重解像度解析 (MRA) やドベシィ (Daubechies) の正規直交ウェーブレット基底の研究に多大な影響を与えた。

1. ウェーブレットによる画像近似・圧縮

コンパクト台を持つ一次元のウェーブレット関数 $\psi(x)$ は、そのバニシング・モーメント特性、即ち $\int x^p \psi(x) dx = 0, 0 \leq p \leq p^*$ (p^* は勿論 ψ に依存する) とその局所性のため、与えられた一次元信号が区分的に滑らかな場合、そのウェーブレット係数 (の絶対値) は、主に不連続点・特異点の近傍で大きくなり、他の滑らかな部分では無視できる程小さくなる。この性質のため、ウェーブレットは、区分的に滑らかな一次元信号の近似・圧縮に非常に有効である。画像を含む多次元信号処理には、一次元のウェーブレットのテンソル積 (例えば、二次元では、 $\varphi(x)\psi(y), \psi(x)\varphi(y), \psi(x)\psi(y)$) が使われることが多い。後述するように、このような一次元の実数値ウェーブレットのテンソル積は、一般の複雑な画像の処理には必ずしも最適ではないが、本項では、簡単のため、上記の一次元実数値ウェーブレットのテンソル積を使った画像の近似と圧縮について述べる。

1.1 画像近似

ここでは便宜上、与えられた画像信号 $f(\mathbf{x}), \mathbf{x} \in \Omega := [0, 1]^2 \subset \mathbb{R}^2$ が $L^2(\Omega)$ に属するものと仮定し、更に $L^2(\Omega)$ 上のある正規直交基底 $\{\phi_i(\mathbf{x})\}_{i=1}^{\infty}$ (例えばウェーブレット基底) を用いて f を近似することを考える。(双直交基底やフレームの場合は、それなりの変更をしなければならないが、基本的には同じ考え方である。) 特に f の N 項線型近似 f_N は、

$$f_N(\mathbf{x}) := \sum_{i=1}^N \langle f, \phi_i \rangle \phi_i(\mathbf{x}),$$

と表わすことができる。ここで注意することは、 N 個の基底関数 $\phi_1, \dots, \phi_N$ は与えられた f に依存しないことである。これに対し、 f の N 項非線型近似 $\tilde{f}_N$ は、

$$\tilde{f}_N(\mathbf{x}) := \sum_{i \in I_N} \langle f, \phi_i \rangle \phi_i(\mathbf{x}),$$

と表され、 $I_N = I_N(f) \subset \mathbb{N}$ は、(絶対値の意味で) 一番大きいものから N 番目までの展開係数のインデックスの集合で、与えられた画像 f に依存する。従って、図 1 から分るように、同じ基底と同じ項数 N を使うという条件下では、非線型近似が線型近似に比べ近似品質が悪くなることはありえないが、どの程度優れているかは、対象となる画像のクラスに依存する。そこで、画像の属する様々な関数空間全体を考え、その空間に属する任意の関数 (画像) f に対し、 N 項近似の品質を近似誤差の L^2 -ノルムで定量化し評価する考えが生れた。例えば、 f が $L^2(\Omega)$ のみならず有界変動関数 (functions of bounded variation) のクラス $BV(\Omega)$ に属するならば、

$$\|f - f_N\|_2^2 = O\left(\|f\|_{BV} \|f\|_{\infty} N^{-1/2}\right),$$

$$\|f - \tilde{f}_N\|_2^2 = O\left(\|f\|_{BV}^2 N^{-1}\right)$$

が満たれることが知られている[2, Theorem 9.18]。ここで、 $\|f\|_{BV} := \int_{\Omega} |\nabla f(\mathbf{x})| d\mathbf{x}$ であり、不規則なテクスチャーのない比較的単純な画像 (例えば漫画画像等) の多くは $BV(\Omega)$ に属する。図 1 に、 128×128 画素の画像の 9/7-双直交ウェーブレットによる線型近似 (a) と非線型近似 (b) の違いを示す。ここでは、 $N = 1024$ 項、即ち全体の 6.25% の項数を使った。図 1 (c), (d) から明かなように、非線型近似の方が線型近似よりもスカーフの縞模様や唇の形をよく捉えていることがわかる。また、SN 比 ($20 \log_{10} (\|f\|_2 / \|f - f_N\|_2)$; 単位は dB) も、線型近似の 19.97dB に比べ、非線型近似では 24.91dB と改善している。

1.2 画像圧縮

上記のように、近似理論的には画像を Ω のような連続体上で定義された関数と考えるのが通常であるが、実際には離散画像、即ち Ω 内の有限個の格子点上での画素値データをどのように圧縮するか、という問題が画像圧縮の中心の問題となる。多くの場合、与えられた離散画素 (サンプル) 値 $\{f(\mathbf{x}_i)\}_{i=1}^n$ から、MRA の V_0 空間に属する関数 $f_s(\mathbf{x}) := \sum_{i=1}^n f(\mathbf{x}_i) \varphi(\mathbf{x} - \mathbf{x}_i), \mathbf{x} \in \Omega$ を構成し、

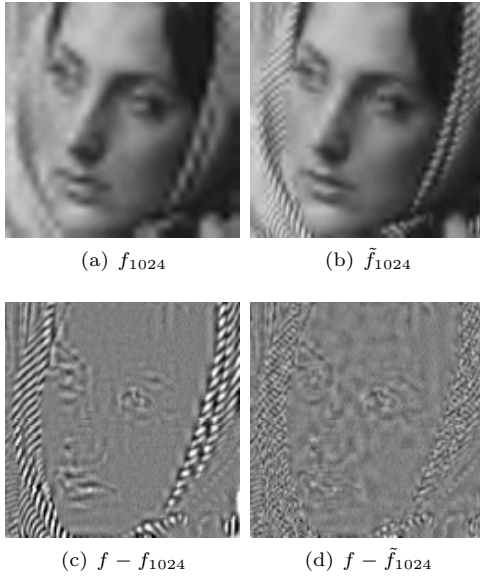


図 1 ウェーブレットによる画像の線型近似と非線型近似

この関数を近似・圧縮するということになる。これは、通常 $\{f(\mathbf{x}_i)\}_{i=1}^n$ から元々の $f(\mathbf{x})$, $\mathbf{x} \in \Omega$ を復元・推定するのが現実的には困難を伴うことが多いからである。画像圧縮とは、近似された画像を更にコンピュータ上でファイルとして記録できるような形に、即ちビット表現に変換することである。実際の圧縮には、近似画像で求められた展開係数の量子化と展開係数のインデックスも含めた情報の巧妙な符号化が必要で、画像圧縮率はそれらに強く依存する。量子化・符号化の詳細については、[2, Chap. 10] とそこに挙げられている参考文献を参照されたい。

この項では、ウェーブレットが採用された **JPEG2000** について概説する。デジタルカメラやコンピュータ上で多用されている静止画像（非可逆）圧縮の標準である JPEG 規格は、1992 年に国際規格 (IS) に採用された。この規格では、元の画像のブロック（ 8×8 画素）への細分化；各ブロック上での離散コサイン変換（Discrete Cosine Transform/DCT）；その展開係数の量子化（特別な量子化テーブルを用いる）；ハフマン符号による符号化；というプロセスを経る（詳しくは、[6], [2, Sec. 10.5] 等を参照のこと）。JPEG 規格では、画素あたり 0.25 ビット以上のレートで圧縮した場合は、通常満足の行く復元画像を得られるとされているが、それよりも低いレート（即ちより高い圧縮率）で圧縮した場合、ブロック歪と呼ばれるブロック境界における歪やモスキートノイズといった好ましくない現象が起こる[6]。これらの欠点を解消しようとして登場した

のが、2000 年にその核心部分が国際規格制定された JPEG2000 規格である。JPEG2000 規格には、パニング・モーメント、台のサイズ、正則性の考慮の結果、対称な 9/7 及び 5/3-双直交ウェーブレットが採用された。特に 9/7-双直交ウェーブレットは、様々な画像圧縮実験の結果、より優れた歪・レート特性を示すことが知られている。また、これらの双直交ウェーブレット変換・逆変換の実装には、リフティング法が効率的に用いられている[2, Sec. 10.5.2]。

1.3 画像境界の扱い上の注意点

与えられた画像を離散ウェーブレット変換（Discrete Wavelet Transform/DWT）するにあたり、常に注意しなければならないのは、画像境界（あるいは画像の外側）をどう扱うかということである。これは、与えられた画像は無限に広がっておらず、画像境界近くでは、DWT のコンボリューション・フィルタが画像の外側の値を必要とするからである。画像の外側を 0 と仮定したり、画像を無理に周期化したりすると、画像境界で画素値の不連続が発生し、（絶対値として）大きなウェーブレット係数が生じてしまう。これらのウェーブレット係数は、言わばこちらの都合で作られたものであり、画像の本質を表したものではないにもかかわらず、画像圧縮アルゴリズムの上では重要とみなされてしまう。これを避けるために最も実用的な方法は、画像を画像境界で鏡対称的に拡張することである。区間上でのウェーブレット（Wavelets on interval）[2, Sec. 7.5] も考案されたが、フィルタ係数が区間端に近づくにつれ変わるので、実装上面倒であることから、実際には余り使われていない。

2. ウェーブレットによる画像ノイズ除去

一般にイメージセンサ等で観測・記録された画像は、常に何らかのノイズ（例えば熱雑音等）を含んでおり、こうしたノイズの除去は、より効果的な画像圧縮・特徴抽出・認識のために重要である。ここでは、簡単のため、ノイズ無しの理想的離散画像を $\mathbf{f}$ 、ノイズ成分を $\boldsymbol{\eta}$ とし、観測画像 $\mathbf{g}$ を、この二つの和、 $\mathbf{g} = \mathbf{f} + \boldsymbol{\eta}$ で表し、これらはすべて $\mathbb{R}^n$ に属するベクトルとするモデルを考える（ n = 総画素数）。

まず、理想画像 $\mathbf{f}$ をある確率的画像モデルの実現値（標本値）、 $\boldsymbol{\eta}$ は画像モデルと統計的に独立と仮定する。このとき、ノイズ除去作用素 $D: \mathbb{R}^n \rightarrow \mathbb{R}^n$ を考える。通常、 D の性能はリスク（ノイズ無しの画像とノイズ除去の結果の平均 2 乗誤差）の値の大小で評価される。画像モデルの先験確率が分っており、かつ D をすべての非線型作用素（線型作用素も含む）の空間から選ぶことができれば、最適である

最小ベイズ・リスク (minimum Bayes risk) を達成できる D が存在し、それは、観測画像が与えられた場合のノイズ無し画像の条件付期待値と一致する。 D を線型作用素の中から求めることに制限したときに、最小ベイズ・リスクを達成する D を**ウィーナー推定量** (Wiener estimator) という。しかし実際には、最小ベイズ・リスクを達成するような D を求めることは、線型・非線型にかかわらず容易ではない。これは、1) 一般の画像 f を標本値とするような確率的画像モデルを設定する困難さ；2) たとえ多数の似た画像を集めることができ、それらを標本値とするようなモデルを考えたとしても、ウィーナー推定に必要な共分散行列を限られた数の標本から正確に推定することの困難さ；に起因する。

そこで、より現実的な手法であるウェーブレット係数の閾値によるノイズ除去法を考える。直交 DWT 基底ベクトルを $\{\psi_i \in \mathbb{R}^n\}_{i=1}^n$ とすると、 $g = \sum_{i=1}^n \langle g, \psi_i \rangle \psi_i$ と書くことができる。ここで、与えられた閾値 $T > 0$ に対し**ハード閾値作用素** (Hard Thresholding Operator) D_T を以下のように定義する。

$$D_T g := \sum_{i \in I_T} \langle g, \psi_i \rangle \psi_i$$

$$I_T = I_T(g) := \{1 \leq k \leq n : |\langle g, \psi_k \rangle| \geq T\}.$$

定義から明かなように D_T は非線型作用素である。ここで、どのように T を決めるのかが重要な問題となる。ドノホー (Donoho) とジョンストーン (Johnstone) は、 η がガウス性白色雑音 $\mathcal{N}(\mathbf{0}, \sigma^2 I_n)$ に従うという仮定のもとで、 $T = \sigma\sqrt{2 \ln n}$ とした時、

$$E\|f - D_T g\|^2 \leq (1 + 2 \ln n) \cdot \left(\sigma^2 + \sum_{i=1}^n \min(|\langle f, \psi_i \rangle|^2, \sigma^2) \right)$$

となることを示した[2, Theorem 11.7]。実際には、 σ は未知であり、細かいスケールのウェーブレット係数から推定することになる。また、通常の DWT がシフト不変でないことに着目し、冗長ではあるがシフト不変の「**定常**」ウェーブレット変換 (Stationary Wavelet Transform/SWT) を使った閾値法がコイフマン (Coifman) とドノホー (Donoho) により提案され、より優れたノイズ除去効果を示している。様々な閾値決定方法のみならず、ウェーブレットやそれに関連する変換を用いたノイズ除去については、ソフト閾値法、ブロック閾値法等を含め、膨大な数の研究がなされてきた。それらに関しては、[2, Chap. 11] とそこに挙げられている文献を参照されたい。

図 2 に、いくつかのノイズ除去法の結果を示す：(a) 原画像の画素値を $[0, 1]$ に正規化したものに

ガウス性白色雑音を加えたもの ($\sigma = 0.08$; SN 比 = 15.90dB)；(b) ウィーナー推定によるノイズ除去 (SN 比 = 22.36dB; ノイズ無しの原画像の情報も用いた)；(c) 9/7-双直交ウェーブレット係数の閾値法 ($T = 3\sigma$; SN 比 = 19.55dB)；(d) 定常 9/7-双直交ウェーブレット係数の閾値法 ($T = 3\sigma$; SN 比 = 22.24dB)。



図 2 ウェーブレットによる画像ノイズ除去の例

3. ウェーブレットによる画像特徴抽出

一般の画像の基本的特徴としては、局所的なエッジ、角、それらの統合されたリッジ (稜線)、テクスチャー等が挙げられよう。これらの画像特徴の抽出とウェーブレットとの関係は、冒頭でも述べたように視覚神経情報処理論を通して、歴史的にも深い。多重スケールエッジ検出作用素として、マー (Marr) 等によるラプラシアン・オブ・ガウシアン (Laplacian of Gaussian/LoG) フィルタやその近似である差分ガウシアン (Difference of Gaussians/DoG) フィルタ、またテクスチャー解析用のツールとして、様々なスケール・方向性・空間周波数を持つことができるガボール (Gabor) フィルタ等が用いられてきた。これらは、すべてウェーブレットの前身とも言え、画像のウェーブレット係数の絶対値はウェーブレット・フィルタによる画像特徴抽出の結果と考えることもできる。

ここで注意しなければならないのは、一次元実数値ウェーブレットのテンソル積を二次元ウェーブレット

関数として使った通常の冗長性のない(双)直交 DWT は、一般の複雑な画像の処理には必ずしも最適ではない、ということである。これは、単に3つの実数値ウェーブレット関数、 $\varphi(x)\psi(y)$, $\psi(x)\varphi(y)$, $\psi(x)\psi(y)$ の伸長・平行移動を使う DWT では、画像に多く存在する水平・垂直以外の方向性をもったエッジやリッジ等の特異曲線、及びテクスチャー等の振動的な特徴を記述するのに、多数のウェーブレット係数が必要となるからである。こうした欠点を解消するために、ウェーブレットフレーム (wavelet frame)、複素ウェーブレット変換 (complex wavelet transform/CWT)、カーブレット (curvelets)、バンドレット (bandlets) 等、様々な画像変換・表現法が提案されている (その一部は, [2, Sec. 5.5, 9.3] 等を参照のこと)。また、ウェーブレット・パケット (wavelet packet) が提案された元々の理由も、ウェーブレットの振動的な特徴に対する弱点を解消するためであった。これらの変換の多くは冗長度のある画像変換であり、特徴抽出や認識のみならず画像近似・圧縮・ノイズ除去にも有効である。

ここでは、これらの変換のうち、特に**二重木複素ウェーブレット変換** (Dual Tree Complex Wavelet Transform/DT-CWT)[5] について述べる。この名前の由来は、入力が一次元信号の場合、実数値 DWT を2回行うことから来ている。さて、一次元複素ウェーブレット関数を $\psi_c(x) = \psi_r(x) + i\psi_i(x)$ と表記しよう。本来は、理想的な複素ウェーブレット関数は解析的 (analytic)、即ちその虚部が実部のヒルベルト変換 (Hilbert transform), $\psi_i(x) = \mathcal{H}\psi_r(x)$, になることが望まれる。この場合 $\xi < 0$ の領域で $\psi_c(\xi) = 0$ となる。実数値の場合と同様に、二次元の複素ウェーブレット関数も、一次元複素ウェーブレット関数のテンソル積 $\varphi_c(x)\psi_c(y)$, $\psi_c(x)\varphi_c(y)$, $\psi_c(x)\psi_c(y)$ で表されるが、大きな違いは、 $\psi_c(x)$ と $\psi_c(y)$ の上記の解析性により、これらのテンソル積の空間周波数領域における台が単一の象限に局在化されることである。このために、二次元複素ウェーブレット関数は、水平・垂直以外の様々な方向を選択することができるのである。この性質以外にも、実数値 DWT に比べ、CWT は、位相情報を抽出できる、展開係数がほぼシフト不変になる、といった長所が挙げられる。残念ながら、この解析性はウェーブレット変換に要請される完全再構成性と両立できないことが知られている。従って、DT-CWT では、完全再構成性を満たしながらウェーブレット関数を出来るだけ解析関数に近づけ、かつ効率的に計算できるアルゴリズムが採用されている。尚、DT-CWT の冗長度は入力画像の4倍、即ち n 画素の実数値画像は $2n$ 個の複素数値係数に変換される。

DT-CWT による非線型近似・特徴抽出の様子を図3に示す。(a) わずか $N = 50$ 項の非線型近似で、鼻、唇、スカーフのエッジ等の方向性をもった特徴がよく表されている；(b) $N = 300$ 項を使うとスカーフの縞模様が捉えられ始めているのが分る。ウェーブレット変換やその様々な拡張により抽出さ

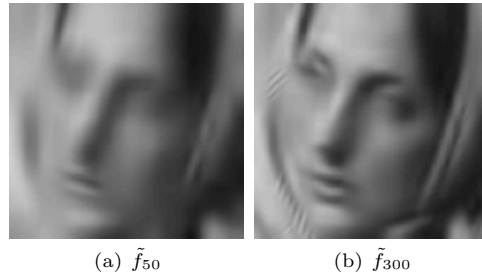


図3 DT-CWT による画像の非線型近似

れた特徴の画像認識への応用は、胸部 X 線写真、マンモグラフィー等の医用画像診断、虹彩認識、顔認識、指紋認識等の生体認証、画像検索、等枚挙にいとまがない。その一端は、例えば[1]等を参照して頂きたい。

謝 辞

図1, 2は、ペイレ (Peyré) の[3], 図3は、セレスニック (Selesnick) の[4], に記載されている MATLAB コードを用いて作成した。

[斎藤直樹]

参 考 文 献

- [1] A. BOVIK, ed., *Handbook of Image and Video Processing*, 2nd ed., Academic Press, 2005.
- [2] S. MALLAT, *A Wavelet Tour of Signal Processing*, Academic Press, 3rd ed., 2009.
- [3] G. PEYRÉ, *A numerical Tour of Signal Processing*. <http://www.ceremade.dauphine.fr/~peyre/numerical-tour>.
- [4] I. SELESNICK ET AL., *Matlab Implementation of Wavelet Transforms*. <http://eeweb.poly.edu/iselesni/WaveletSoftware>.
- [5] I. W. SELESNICK, R. G. BARANIUK, AND N. G. KINGSBURY, *The dual-tree complex wavelet transform*, IEEE Signal Processing Magazine, 22 (2005), pp. 123–151.
- [6] K. YAMATANI AND N. SAITO, *Improvement of DCT-based compression algorithms using Poisson's equation*, IEEE Trans. Image Process., 15 (2006), pp. 3672–3689.