

■2 群 (画像・音・言語) - 1 編 (画像処理)

2 章 画像処理の基本手法

(執筆者：児玉和也) [2012 年 8 月 受領]

■概要■

画像処理では、画像の有する何らかの不具合や劣化などの改善に加え、より後段の高次の解析などの前処理としての機能も求められる。実際、コントラストなどの階調表現や画像内容の幾何的な歪みなどの補正、画像中の雑音やボケによる劣化の抑制を含む品質調整など、本章で紹介する様々な基本手法が人間にも計算機などにも見やすい画像を構成し、そこから効率的に有用な情報を取り出し活用することが可能となっている。

近年、計算機性能の著しい向上と、最適化分野など信号処理周辺の多様な理論的發展が相俟って、画像処理としての性能がより優位であれば、はるかに複雑で演算量を要する反復型の手法などが十分に現実的となっている。ただし、そこで取り扱われる画像処理としての課題そのものは、上述のとおり基本手法が対象としてきた機能などから大きく外れるようなことはなく、その重要性はほとんど変わっていない。また、様々に提案され続ける高性能な手法でも多くの場合、例えば本章で中心となるフィルタ処理、信号変換などの基本手法が欠くことのできない構成要素として活用されている。

以上のことから本章は、画像処理の簡便な「基本手法」の紹介を通して、画像処理の「基本課題」に広く触れ、その効果や難しさなど各課題、各手法の特徴が合わせて確認できるよう、簡潔で概観しやすい記述にとどめた。ここから、より専門的な詳細、あるいは性能向上が著しい直近までの技術的展開につなげるならば、対応する他の頁はもちろん、映像情報メディア学会(編)“映像情報メディア工学大事典”(オーム社)などでも補うことができるはずである。

【本章の構成】

本章では、画像処理の基本手法として、画像の濃度変換 (2-1 節)、画像の補正・強調・復元 (2-2 節)、画像の空間・周波数分析 (2-3 節)、画像の特徴と分析 (2-4 節) について述べる。

■2 群-1 編-1 章

2-1 画像の濃度変換

(執筆：児玉和也) [2018 年 12 月 受領]

2-1-1 階調変換

撮像系、表示系といった画像入出力機器などのガンマ特性や光学特性などの制約に起因し、撮影などで得られた画像信号をそのまま扱うのは、対象の明るさ、色彩などを良好に再現したり、局所的な精細さと全体の整合性が維持された視覚情報を構成したりするのに、不十分である場合が多い。

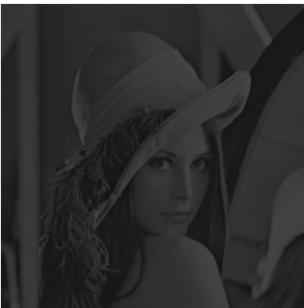
例えば、グレースケール（濃淡階調）画像の場合でも、階調特性の歪みなどから濃度分布に偏りが生じ、対象を精細に認識することが困難となりうる。これは画像の対象物そのものの持つ濃淡情報の分布の程度にも依存し、図 1・1 のように、広すぎれば画像信号が取りうる有限の階調に納めることができず白とびや黒つぶれが生じ、狭すぎれば画像信号の量子化誤差で情報が失われたりコントラスト不足で細部が認識困難となったりする。



望ましい階調（原画像）



階調不足（白とび，黒つぶれの発生）



コントラスト不足



ヒストグラム等化による補正

図 1・1 階調変換

実際、画像内容を十分に活用するには、画像信号に対しコントラスト変換など何らかの階調補正（変換）を施すことが、たいへん有効である。単純には、元の階調を x 、変換後の階調を

y とすれば、 $y = f(x)$ となる関数 f を定めることにより、階調変換は実現される。例えば、対象となる画像信号の階調の最小値から最大値までを、表示系などの最小値から最大値に線形に割り当てる関数 f を用いるだけでも、ある種の有効なコントラスト変換となりうる。また、 $g(x) = x^\gamma (0 \leq x \leq 1)$ なるガンマ特性を持つ表示系に対しては、 $f(x) = g^{-1}(x) = x^{1/\gamma}$ とする階調補正を施すことで、画像内容が適切に再現される。あるいは、画像内容に応じ、有限の階調を無駄なく均等に活用することを目的として、ヒストグラム等化と呼ばれる処理が用いられることもある。図1・1では、コントラスト不足の画像を対象とした結果を合わせて示したが、具体的には、対象となる画像信号の階調についてヒストグラムを求め、その累積値をとって更に正規化した関数 $h(x)$ に対し、 $f(x) = h^{-1}(x)$ と定める形の階調補正を行っている。原画像とは異なるが、細部まで容易に視認可能である。

以上では、グレースケール画像を例に挙げて説明したが、カラー画像の場合は赤、緑、青の三原色の間での整合性なども重要である。実際、階調補正に限らず様々な処理において、単純に各色を独立に扱うだけでは画像内容の色彩に関する再現性はしばしば損なわれることとなるので、必要に応じ、次々項でも見るような表色系の取り直しなども考慮しなければならない。

ところで、デジタル画像処理においては、8 bit 程度で階調の表現を行うことが多い。この場合、非常に暗い領域と非常に明るい領域を同時に含むような対象は、人間の視覚では比較的容易に認識できるにも関わらず、図1・1でも見たような白とびや黒つぶれで、いずれかの情報がまるごと損なわれることが多い。そこで近年、12 bit やあるいはそれ以上に階調を増加させた HDR (High Dynamic Range) 画像を用いることで、以上のような課題に対し様々な取り組みが展開されている。撮像系に関しては、例えば、同一の対象物について暗い領域、明るい領域それぞれに感度を合わせて得られた通常の 8 bit 画像 2 枚を、階調の整合性を維持しつつ 12 bit 階調の画像へと統合する処理などが提案されている。表示系に関しても、人間の視覚にならない、領域の明暗に応じ局所的に異なるコントラスト変換などを行いつつ、画像全体としても整合性を大きく損なわないような、適応的な階調補正などが検討されている。

2-1-2 限定色カラー表示・ディザ表示

表示デバイスや、画像情報を保持するメモリ装置を含めたその制御系の制約などから、限られた色数のみで画像を表示する必要があるがしばしば生ずる。実際、図1・2のとおり、赤、緑、青の三原色がそれぞれ 8 bit の階調を有する画像も、最低でも白、黒の 2 値があれば、その内容をある程度まで再現することが可能である。

一般には、より多くの色数、すなわち、図1・2の場合であれば、もとの三原色合計 24 bit で表される $256 \times 256 \times 256$ 色のうちから、8 bit で管理できる 256 色程度を代表色として選択し、表示を行うことが多い。例えば、各色について 0, 51, 102, 153, 204, 255 の 6 階調のみを選択すれば、 $6 \times 6 \times 6 = 216$ 色が得られることとなる。この中から各画素ごとに、元の色に最も近いものを割り当てれば、図1・2のように表示することができる。この場合、画像内容によらず代表色は固定されるが、再現性を改善するため、画像中で頻度が高い色を適応的に選ぶこともある。また、代表色の選択やもとの色に対する代表色の割当に際しては、次項で見えるような表色系の取り方もしばしば考慮される。



原画像 (RGB 各色が 8 bit の階調)



2 値によるディザ表示



限定色 (6×6×6 色) による表示



6×6×6 色によるディザ表示

図 1・2 限定色表示

以上の、限定色を単純に用いた表示に加え、印刷物における網点処理と同様の技術である、ディザ表示は再現性の向上にたいへん有効である。実際、カラー画像、グレースケール画像に関わらず、印刷物では非常に限られた色数を用いながらも、網目のように細かな各色の点の大きさを変化させることで、中間的な階調までが良好に再現されている。これは、人間の視覚では局所領域ごとに平均化された色が認識されていることによる。ディザ表示は中間調を表現するためこのような視覚特性に着目し、代表色への割当の際に生じる誤差が適当な局所領域内ではなるべく打ち消し合うように、各画素ごとに異なる処理を行う。

0	8	2	10
12	4	14	6
3	11	1	9
15	7	13	5

2 値化の場合

(16 倍し 8 を足してから用いる)

2	28	9	34
40	15	47	21
12	37	5	31
50	25	44	18

6 階調とする場合

(階調間の幅 51 に合わせたもの)

図 1・3 ディザ行列による閾値のパターン

例えば、組織的ディザ法では、対象画像を 4×4 画素などに分割し、そのサイズに合わせてディザ行列と呼ばれる図 1・3 のような閾値パターンを定める。256 階調を 2 値化する場合であ

れば、単にこの閾値との大小で白か黒かを決定すればよい。6階調などとする場合でも、選択すべき階調のうち、各画素のもともとの階調を挟み込む隣り合う階調をいわば2値とみなしたうえで、対応する図1・3のような閾値を用いて同様の処理を行い、いずれの階調かに決定することで図1・2のとおり所望の結果が得られる。特に滑らかな領域では、このように閾値処理を画素ごとに可変にしたことで、誤差が良好に打ち消し合うため、限定色による単純な表示より望ましい効果が期待できるが、一方で、ディザ行列そのものが固定であることに起因し、繰り返しパターンが視認されがちである。このような固定パターンに依存せず、統計的に誤差を打ち消し合うよう閾値をランダムに発生させる方法もあるが、より一般には、注目画素が代表色に割り当てられる際に生じる誤差を後で処理する周辺の複数の画素から適当な比率で差し引くことで、以上の閾値による方法と同様に誤差の打ち消し合いの効果が生じる、誤差拡散法を用いた表示が安定な結果を与え好まれている。図1・2の2値によるディザ表示も、誤差拡散法を用いたものである。

2-1-3 色彩加工

カラーの画像信号を構成するにあたり、どのような表色系を選択するかは、対象の明るさや色彩を適切に扱ううえで、たいへん重要である。実際、人間の視覚や、撮影、表示に用いる画像入出力機器の特性も考慮し、様々な表色系が、印刷、放送、計算機などの各分野で用いられている。

プリズム分光などにより認識される様々な色は、物理的には連続的に波長が変化していく様子にすぎないとも言えるが、人間の網膜上にある基本的な3種類の錐体細胞がそれぞれ赤(R)、緑(G)、青(B)に相当する波長を中心に反応し、いわばその結果の組合せで、このように多様な色が知覚されている。本来、対象の明るさ、色彩を物理的に再現しようとするれば、画像信号はすべての波長についてその強度を保持する必要があるが、こういった視覚特性にならった撮影機器などの画像入力デバイスや、逆に所望の錐体を適切に刺激し人間の視覚に合わせて対象を再現する表示装置などの画像出力デバイスが実装可能なため、多くの場合、RGBの三原色の強度のみで構成された画像を基本としてよい。

直接的に発光することで画像を表現するTVモニターや計算機ディスプレイは、黒を地としており、そのままRGB表色系が適合している。逆に印刷物では白を地にインクなどで光を吸収させ、反射光を減ずる仕組みであることから3原色の補色であるシアン(C)、マゼンダ(M)、イエロー(Y)を基本としたCMYK表色系が用いられている。一方、従来からTV映像の伝送などでは、互換性に加え圧縮や劣化抑制などのため、YUV表色系やこれとほぼ同様であるYCbCr(YPbPr)表色系が標準となっていた。ここで、Yは明るさであり、視覚特性と関連してRGBのうちGに強く影響される。この信号は画像の内容にとって最も重要な情報であり、白黒テレビとの互換性もとることが可能となる。残りの2個のパラメータはそれぞれBとRに相当する色彩情報を保持するが、より具体的にはYUVとRGBの間の変換は線形であり、以下のような3×3行列などで簡単に表現できる。

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1 \cdot 1)$$

また、JPEGやMPEG2などの映像信号の符号化では、人間の視覚が明るさほどは色彩情報に

敏感でないことに着目し、残りの2個のパラメータについては画素数を縦横それぞれ半分の間引いた信号を構成してから圧縮を施す方法がとられている。カラー画像に関わる特性をうまく利用することで、視覚的な劣化を抑えながらも効率の良い圧縮符号化を実現しているのである。

このほかにも、色差の表現などを考慮し、従来からXYZ表色系やCIE L*a*b*表色系などが用いられてきたが、現在もCIE 2000による色差式をはじめ様々なものが提案され続けている。

画像処理において、特に重要なものとして、HSI表色系あるいは同様のHSL (HLS) 表色系、HSV表色系などがある。これらは、色相 (Hue)、彩度 (Saturation) と輝度あるいは明るさ (Intensity/Lightness/Value) を基本とした表色系であり、RGB表色系との対応で言えば、例えば、R/G/Bの3パラメータで表現される立方体において原点 (黒) と相対する頂点 (白) を結ぶ直線上をIの軸とし、これに直交する平面上で色彩 (HとS) を定める。色相 (H) はこの平面の原点であるI軸との交点から見た方向であり、彩度 (S) は交点に近付くほど混色が進み低下すると考える形をとる。いわば、Hで色の種類を、Sでその純粋さを表現していることに注意されたい。実際、画像内容の分析の際には、照明などの環境の影響を比較的受けにくい色相 (H) を用いることで、一般に、対象物の分類などがより安定に実現される。また、前々項でみたコントラストの強調、抑制や、次節以降でみるフィルタリングなどを、RGB表色系でR/G/B独立に適用すると、処理前後で不自然な色彩の変化が生じ所望の結果を得ることが非常に困難であるため、彩度 (S) に着目した強調、抑制や輝度 (I) に対する各種フィルタリングなどの後に、再びHSI表色系からRGB表色系に戻すことで、色彩の統一性の維持を図る場合も多い。

カラー画像においては、以降でみる各種の画像処理を適用する際に、不自然な色彩、いわゆる偽色が発生しないよう、上述のとおり表色系の取り方に注意を払うほかにも、例えば、復元処理であれば、色彩に関するなんらかの拘束条件を追加するなど、様々な工夫がたいへん重要となる。

■2 群-1 編-2 章

2-2 画像の補正・強調・復元

(執筆者：児玉和也) [2018 年 12 月 受領]

2-2-1 幾何補正

一般に、撮影などによって取得された画像そのものは、実際の光学系がモデルとなる理想的なピンホールカメラと相異なることから、様々な幾何学的な歪みを持つこととなる。これらの歪みは、例えば、正方形パターンの繰り返しとなるべき図 2・1 のような場合でも、画像中の直線性や直交性など撮影対象の重要な情報を失わせるため、本来のモデルに従うよう何らかの補正が必要である。具体的には、歪みを持った補正前の座標 (X, Y) に対し、歪みを抑制した補正後の座標 (x, y) を関係づける、線形あるいは非線形の幾何補正を施す。例えば、光軸に対する撮像面のずれや傾きなどから生ずる線形歪みに対しては、

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad (2 \cdot 1)$$

のアフィン変換を用いた補正が有効である。また、レンズ系からしばしば生ずる樽型歪み (Barrel Distortion) や、いわばその反対向きの糸巻型歪み (Pincussion Distortion) は非線形だが、

$$\begin{pmatrix} x \\ y \end{pmatrix} = (1 + k(X^2 + Y^2)) \begin{pmatrix} X \\ Y \end{pmatrix} \quad (2 \cdot 2)$$

で近似的に補正可能である。それぞれ、 $k > 0$ あるいは $k < 0$ の場合に対応する。



線形歪み



樽型歪み

図 2・1 典型的な幾何歪み

以上の補正では、それぞれパラメータが必要となるが、カメラの内部パラメータが変更されなければ、それらは固定と考えてよく、テストチャートなどを用いてあらかじめ決めておくことができる。また、ステレオ視による奥行き認識処理などでは、これら補正の精度が前処理としてたいへん重要であるが、Tsai の方法、Zhang の方法など、あるパターンを撮影することで、上述のパラメータなどをまとめて効率良く推定する手法が提案されている。

なお、画像信号が離散的なことから、式(2・1)や式(2・2)の関係で直接に求まる座標に対応する画素は必ずしも存在しない。このため実際には、その近傍の画素の情報から対応する座標の情報を補間して得る処理が必要となる。これについては次々項で述べる。

2-2-2 鮮鋭化・平滑化

画像信号をその強調、抑制によって、更に活用しやすい形に加工するための最も基本的な処理として、鮮鋭化や平滑化がある。ここでは、各画素を含む局所領域ごとに何らかの処理を施すフィルタを用いた手法を示す。これらは、次々節でより一般的に述べるオペレータの一種とも考えられ、次節のみ見るような信号変換ほどの演算量やメモリ容量を必要とせず、画像中の空間座標のみを用いて具体的な演算が簡潔に表現できる。

画像信号 $f(x, y)$ に対する鮮鋭化、平滑化のためのフィルタ処理は、しばしば、

$$g(x, y) = \sum_{i, j} h(i, j) f(x - i, y - j) \quad (2 \cdot 3)$$

なる畳み込み演算として実装され、フィルタ $h(x, y)$ は、例えば鮮鋭化の場合、

$$\begin{pmatrix} h(-1, -1) & h(0, -1) & h(1, -1) \\ h(-1, 0) & h(0, 0) & h(1, 0) \\ h(-1, 1) & h(0, 1) & h(1, 1) \end{pmatrix} = \begin{pmatrix} 0 & -c & 0 \\ -c & 1 + 4c & -c \\ 0 & -c & 0 \end{pmatrix} \quad (2 \cdot 4)$$

及び、この範囲外で0と定めると、各画素に関し 3×3 個の近傍画素のみから処理結果の $g(x, y)$ を与える。ここで、 $c (> 0)$ は鮮鋭化の強度を表すパラメータであり、式(2・4)では強調すべき信号の変化を捉えるものとして、次々節に関連シラブラシアンオペレータ (フィルタ) を応用していることにも注意されたい。図 2・2 に $c = 1$ とした場合の処理結果を示した。前景の複雑なテクスチャが強調されるのみならず、焦点ぼけしていた背景も各エッジがシャープになることで各領域境界が明確に視認できる画像が得られている。すなわち、このような簡便な鮮鋭化でも、ぼやけた文字など画像内容を読み取りやすくするなどの効果が十分に期待できる。



図 2・2 鮮鋭化処理の例

平滑化も、同様の畳み込み演算で実現可能であり、例えば $h(x, y)$ を式(2・4)にならひ、

$$\frac{1}{256} \begin{pmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix} \quad (2 \cdot 5)$$

と、 5×5 個の近傍画素に対応させる形とすればよい。この $h(x, y)$ はガウシアンフィルタと呼ばれ、ある分散を持つガウス関数を近似したものとなっている。図 2・3 に、雑音を加算した画像に、これを繰り返し適用した結果を示した。変化の大きい高周波の信号は一律に抑制されるため、雑音は抑制されていくが、細かなテクスチャや領域境界もぼやけてしまっている。ガウシアンフィルタはその特性から、次節でもみるような焦点ぼけの表現などにも有効であるほか、解像度を下げる際にエイリアシング (Aliasing) を避けるため前処理として用いられることが多い。



雑音を加算した原画像



メディアンフィルタ



ガウシアンフィルタ (1回)



ガウシアンフィルタ (2回)

図 2・3 平滑化処理の例

一方、畳み込み演算とは異なり非線形の特長となるが、同様に 3×3 や 5×5 の近傍画素に注目し、その中間値を出力するメディアンフィルタも、外れ値を除外するという効果から、雑音除去などの平滑化でしばしば用いられる。実際、図 2・3 では 5×5 のメディアンフィルタによる結果を示したが、ガウシアンフィルタほど細部がぼやけてしまうことなく、雑音が抑制され

ていることがみてとれる。

2-2-3 画像復元・補間・超解像

焦点ぼけや雑音に起因する劣化などにより、対象とする画像信号中に所望の情報が直接には含まれていない場合、何らかの方法でそれを出来る限り安定に再構成しようとする処理を一般に画像復元と呼ぶ。前項でもみた鮮鋭化や平滑化も、ある種の画像復元である。また、前々項で必要性を述べた幾何補正においても、実際の標本点の情報から補正後の標本点の情報を再構成せねばならないが、これは補間処理と呼ばれ、同様の問題と考えるとよい。

画像信号が離散的であることから、様々な処理においてこのような画素補間は必要となるが、特に、画像信号を取得するイメージセンサ自身の制約に応じて、系統だった補間処理が用いられることに注意されたい。具体的には、イメージセンサで取得されているものに比して例えば2倍の解像度の画像を再構成しようとする超解像が、たいへん盛んに検討されている。このような処理は、単板でカラーイメージセンサを実現するため3原色に対応して画素をモザイク状に配置させて得られる情報から、各色ごとに本来の画素数分の標本点での情報を再構成するデモザイクングとも類似しており、実際、多くの技術的基盤を共有している。もちろん単に、イメージセンサ上の欠損画素を近傍画素から予測するような場合でも、補間処理は用いられている。一方で、画像内容の一部の消去など自然な編集を目的に、単一の画素単位ではなく、多数の画素の集合であるより大きな領域単位で画像信号を再構成しようとするインペインティングなども画像復元の先端的な研究として取り組まれているが、手法により、偏微分方程式による画像信号のモデル化やその反復解法など、様々な枠組が必要となっている。

ここでは以下、補間処理について簡潔に述べ、これに基づき、超解像などを含む画像復元における基本的な課題も概説しておく。まず、実際の標本点上の値 $f(i, j)$ から、任意の新たな標本点上の値 $f(x, y)$ を求めるには、それらの座標間の距離に関わる関数 $h(t)$ を用いたある種の畳み込み演算で、

$$f(x, y) = \sum_{i, j} h(x-i)h(y-j)f(i, j) \quad (2 \cdot 6)$$

とするのが一般的である。最も単純な、バイリニア (Bilinear, 双一次) 補間では、

$$h(t) = 1 - |t| (|t| < 1), 0 (|t| \geq 1) \quad (2 \cdot 7)$$

とすればよく、この場合、式(2・6)は実際には、 (x, y) を囲む近傍の 2×2 画素についてのみ和をとることとなる。図2・4に、図2・2で用いた原画像に式(2・5)のガウシアンフィルタを施した後1点おきにサンプリングして作成した縦横1/2ずつの低解像度の画像に対し、バイリニア補間を適用し、元のサイズへと画像拡大を行った結果を示した。原画像と比べれば滑らかではあるが、エッジなどを含め全体にぼやけた画像となっている。これは、補間の与える値が式(2・7)から分かるように近傍画素のとる値の範囲に収まり、信号の極値近辺やエッジなど変化の大きい領域をうまく捉えきれないことによる。より良好な補間のためにサンプリング理論に立ち返れば、シンク (sinc) 関数をもとにした、

$$h(t) = (n/\pi^2 t^2) \sin(\pi t) \sin(\pi t/n) (|t| < n), 0 (|t| \geq n) \quad (2 \cdot 8)$$

を用いる Lanczos- n 補間や、これを3次関数で近似したともみなせる、

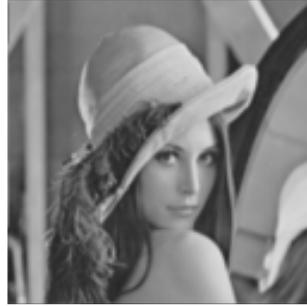
$$h(t) = 1 - (a+3)t^2 + (a+2)|t|^3 \quad (|t| < 1), a(-4+8|t|-5t^2+|t|^3) \quad (1 \leq |t| < 2), 0 \quad (|t| \geq 2)$$

(2・9)

を用いるバイキュービック (Bicubic) 補間が有効であり, 実際にそれぞれ $n = 3$, $a = -1.0$ とした場合の結果を図 2・4 に示したが, より多くの近傍画素を参照し負の値もとりうる $h(t)$ で信号の変化を精度良く捉えることにより, エッジなども含めたいへん滑らかな拡大画像が得られているのがみてとれる。



原画像 (低解像度)



バイリニア補間 (2×2 近傍)



バイキュービック補間 (4×4 近傍)



Lanczos-3 による補間 (6×6 近傍)

図 2・4 補間に基づく画像拡大 (2 倍) の例

一方, 以上の補間手法による単なる画像拡大のみでは, 先に述べた超解像などは実現されていないことにも注意されたい. 実際, ナイキスト周波数で特徴付けられる制約から, 上記の補間は制約の範囲で同一の信号に対し標本点の取り直しを行ってはいるが, 拡大後の高い解像度に応じた周波数の信号まで復元されているわけではない. 一般に超解像などでは, 反対に高解像度から低解像度への画像信号の劣化の過程をモデル化し, その逆問題を解く枠組が用いられる. 多くの場合, このような問題は正則でない, すなわち解が一意に定まらないため, 復元される信号に望まれる特徴などを何らかの形で与えることで, 解を絞り込めるようにする様々な正則化手法が提案されている. 例えば, TV (Total-Variation, 全変動) 正則化では, 信号の全変動と劣化過程への忠実度から評価関数を定めるなどし, これを出来る限り小さくする最適化問題として解を求め復元画像を得る. 変動が小さい信号の方が, 一般に画像としていけば存在確

率が高く、したがって解として望ましい、としているわけである。このほかにも、対象となる画像群がある程度まで限定される場合などは、次節で触れるように、それらの分析から基底を取り直すなどして、画像信号の表現のしやすさ、すなわちある種の確からしさなどを計るなど、多様な評価関数の設計法が試みられている。

上述した劣化過程のモデル化と評価関数による逆問題の正則化は、超解像のみならず焦点ぼけなどに対する鮮鋭化など、様々な画像復元の基本的な枠組となっていることに注意されたい。また、信号に対するノルムの定め方などを含め、評価関数の設計には、実際にそれを最適化する解を見つけるための仕組みが伴う必要があり、画像復元は、凸解析など収束性を保証する数理工学の理論とも密接に関連しつつ、信号処理の知見を総合して取り組まれる課題の一つともなっている。

■2群-1編-2章

2-3 画像の空間・周波数分析

(執筆者：児玉和也) [2018年12月 受領]

与えられた一般の画像信号を、複数の基本的な画像信号の和として表現することは、フィルタリングや符号化を含め様々な画像処理において、たいへん有用である。実際、画像内容を分析するうえで、その画像中にどのような信号が含まれているか、あるいは更に詳しく画像中の局所領域ごとにどのような信号が含まれているか、などを知らうとすれば、上述のとおり、元の画像信号を何らかの基本信号の和に自在に分解して表現する技術が必須である。

本節では、まず、画像信号の分解において最も基礎となる直交変換の考え方を示した後、実用上極めて重要なものとして具体的に、画像中にどのような周波数を有する信号がどの程度まで含まれているかを分析するフーリエ変換、画像中の空間座標ごとに同様の分析を局所的に行うウェーブレット変換を紹介する。

フーリエ変換、ウェーブレット変換は、いわばあらかじめ与えられた基本信号を基底として画像を表現する手法とみなせる。これに対し、大量の画像群が自在に蓄積、通信される計算機ネットワークの著しい発展などを背景として、対象データ群から統計的に適切な基底を求めるPCA (Primary Component Analysis, 主成分分析) やICA (Independent Component Analysis, 独立成分分析) が画像信号の分析にも導入されつつある。そこでは、与えられた基底ではなく、画像群ごとに最適な基底を求め、これに基づきより高い品質での画像の分析が試みられている。このような先端的な画像信号処理技術へと理解を進めるうえでも前提となる基本知識の概要を示すことが本節の目的である。

2-3-1 直交変換

与えられたグレースケール画像の輝度値が、 (x, y) をその画像中の空間座標として、 $f(x, y)$ で表されるとする。

まず、以下の条件を満たす、基本的な画像信号 $\phi_{u,v}(x, y)$ を用意する。

$$\sum_{x,y} \phi_{u_1,v_1}(x, y) \phi_{u_2,v_2}^*(x, y) = \begin{cases} 1 & ((u_1, v_1) = (u_2, v_2)) \\ 0 & ((u_1, v_1) \neq (u_2, v_2)) \end{cases} \quad (3 \cdot 1)$$

ここで、一般に $\phi_{u,v}(x, y)$ は複素数の値をとり、 $\phi_{u,v}^*(x, y)$ はその複素共役の値を表す。なお、加算の範囲は画像全体である。

いま、もとの画像信号 $f(x, y)$ を基本的な画像信号 $\phi_{u,v}(x, y)$ の重み付け和として表すことを考える。すなわち、各 (u, v) に対し $F(u, v)$ をその重みとして、

$$f(x, y) = \sum_{u,v} F(u, v) \phi_{u,v}(x, y) \quad (3 \cdot 2)$$

と表現する。

式(3・1)のとおり直交性を有する $\phi_{u,v}(x, y)$ については、この重み $F(u, v)$ をたいへん簡単に求めることができ、

$$F(u, v) = \sum_{x,y} f(x, y) \phi_{u,v}^*(x, y) \quad (3 \cdot 3)$$

である。実際、式(3・3)において、右辺に式(3・2)を代入すれば式(3・1)の条件から、左辺が得ら

れることが確認できる．なお， $\phi_{u,v}(x,y)$ のとり方により， $F(u,v)$ も複素数の値となることに注意されたい．

以上のように，直交性を有する $\phi_{u,v}(x,y)$ を用いた重み付け表現を前提として，元の画像信号 $f(x,y)$ からその重み付け係数を表す新しい信号 $F(u,v)$ を求めることを，直交変換と呼ぶ．

2-3-2 フーリエ変換

前項で述べた，直交性を有する基本的な画像信号，すなわち画像の直交基底として，

$$\phi_{u,v}(x,y) = e^{2\pi i(ux+vy)/N} \quad (3 \cdot 4)$$

をとることを考える．簡単のため，画像サイズは $N \times N$ としており， $-N/2 \leq u, v < N/2$ に対応する基底のみで議論できる．なお，式(2・10)の条件をそのまま満たそうとすれば，式(3・4)の右辺に $1/N$ を乗じて正規化を行う必要があるが，直交性のみが重要であることから慣例的に式(3・4)を用いる．この基底を用いた場合の式(3・3)は，

$$F(u,v) = \sum_{x,y} f(x,y) e^{-2\pi i(ux+vy)/N} \quad (3 \cdot 5)$$

と書け，これを DFT (Discrete Fourier Transform, 離散フーリエ変換) と呼ぶ．このようにして得られた係数 $F(u,v)$ を用いて元の画像 $f(x,y)$ を表現する式(3・2)は，正規化を省略した式(3・4)に対応させ，

$$f(x,y) = \frac{1}{N^2} \sum_{u,v} F(u,v) e^{2\pi i(ux+vy)/N} \quad (3 \cdot 6)$$

となるが，これはフーリエ変換で得られる係数 $F(u,v)$ から元の信号 $f(x,y)$ を再構成する演算であることから，逆フーリエ変換と呼ばれる．画像は2次元信号であるから2次元フーリエ変換が必要であるが，式(3・5)から分かるように，各列に対し1次元フーリエ変換を施したのち各行に対し1次元フーリエ変換を施せば実現できることに注意されたい．これは，式(3・6)の逆フーリエ変換についても同様である．また，式(3・5)と式(3・6)の変換はほぼ同じ演算であるが，どちらも実際には，バタフライ演算と呼ばれる工夫を組み込んで演算量を削減し高速化を施した FFT (Fast Fourier Transform, 高速フーリエ変換) として実装することができる．この場合，DFT では $o(N^3)$ の演算量が，FFT によって $o(N^2 \log N)$ にまで高速化される．

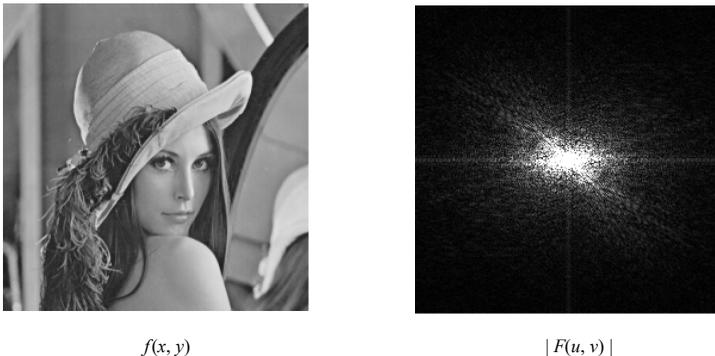


図 3・1 フーリエ変換の例

図 3・1 にフーリエ変換の例を示す．原画像 $f(x, y)$ は実数であるが，そのフーリエ変換で得られる $F(u, v)$ は複素数であることから，ここではノルム $|F(u, v)|$ の大きさに適当に明暗をつけて示した．中央近辺の (u, v) の小さな領域にエネルギー $|F(u, v)|^2$ が集中しており，原画像 $f(x, y)$ は低周波数の信号を中心として構成されていることがみてとれる．一般に図 3・1 に限らず自然画像は，多くの部分が低周波に対応する滑らかな領域で占められるため，同様の傾向を有する．このような， $F(u, v)$ の振幅 $|F(u, v)|$ について画像内容に起因してみられる傾向は，位相については特に存在せず，以下でも確認するように，フーリエ変換で得られる情報は，図 3・1 のとおりノルムに着目して考察することが重要である．

さて，前節でもみた画像のフィルタ処理に関し，フーリエ変換は畳み込み演算に対する以下のたいへん有用な性質を持つ．すなわち，原画像 $f(x, y)$ にフィルタ $h(x, y)$ を畳み込んで得られる画像を $g(x, y)$ とすれば，

$$g(x, y) = h(x, y) * f(x, y) \quad (3 \cdot 7)$$

と書けるが，それぞれをフーリエ変換したものの間で，

$$G(u, v) = H(u, v)F(u, v) \quad (3 \cdot 8)$$

が成立するのである．まず，式(3・7)の畳み込み演算では，フィルタのサイズ $M \times M$ が大きくなるに従い $o(M^2N^2)$ の計算量が課題となるのに対し，式(3・8)の右辺自体は計算量 $o(N^2)$ の複素乗算のみであるから，原画像 $f(x, y)$ のフーリエ変換 $F(u, v)$ に $H(u, v)$ を乗じて $G(u, v)$ としたものに逆フーリエ変換を施して $g(x, y)$ を得る手順全体を考慮しても，有意な場合が多い．また，フィルタ $h(x, y)$ をフーリエ変換した $H(u, v)$ は，原画像 $f(x, y)$ とは関係なくあらかじめ計算しておくことができる利点もあるうえに，そのフィルタが，図 3・1 で見たような原画像を構成する各周波数の成分にどのように作用する特性を持っているか，簡潔に分析することができる．

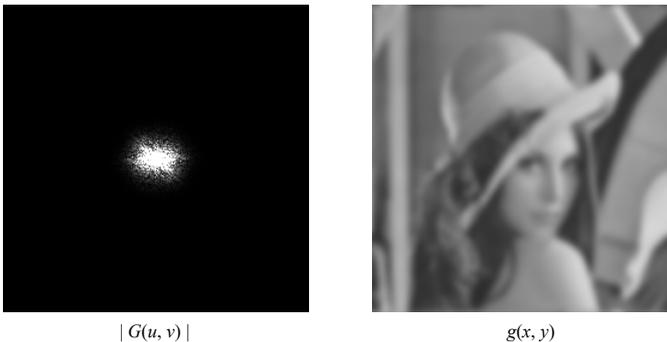


図 3・2 フーリエ変換に基づくフィルタ処理の例

図 3・2 に，図 3・1 の画像に対し上述の手順でフィルタ処理を施した例を示す．フィルタとしては，

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3 \cdot 9)$$

のガウス関数を用いた。これは、焦点ぼけのモデル化とみなせるが、適切な σ を与え 5×5 のフィルタとして近似したうえで式(3・7)を適用すれば、前節でもみたガウシアンフィルタの畳み込みに一致する。一方、式(3・8)によるフィルタリングを実現するためには、フィルタ $h(x, y)$ のフーリエ変換 $H(u, v)$ が必要であるが、この場合は解析的に求めることもできて、

$$H(u, v) = e^{-\frac{\sigma^2(u^2+v^2)}{2N^2}} \quad (3 \cdot 10)$$

を得る。この特性はたいへん簡単であり、高周波数の成分ほど減衰させるフィルタとなっており、高周波に対応する強いテクスチャほど減衰させ、低周波に対応する滑らかな領域はほぼそのまま維持される、という焦点ぼけの効果をよく表している。特に、 $H(0, 0) = 1$ であるから、直流成分に対応する画像信号の平均値はそのまま保たれる、という特性にも注意されたい。実際、図3・2では $\sigma = 3.0$ とした場合を示したが、図3・1の $|F(u, v)|$ と比べ、 $|G(u, v)|$ は高周波ほど減衰させたものであることがみてとれ、その逆フーリエ変換 $g(x, y)$ では、焦点ぼけをよくモデル化した、強いテクスチャが抑制された画像が得られている。

フィルタ処理に関しフーリエ変換を用いた式(3・8)を導入することの更に重要な利点として、これも前節でみた画像復元などへの適用がある。すなわち、式(3・8)を変形して、

$$F(u, v) = H(u, v)^{-1}G(u, v) \quad (3 \cdot 11)$$

とすれば、 $H(u, v) \neq 0$ であれば、図3・2のような焦点ぼけ画像から図3・1の元の画像を復元する鮮鋭化処理が考えられる。実際、前節でみた鮮鋭化フィルタは、適当な $H(u, v)^{-1}$ について逆フーリエ変換して得たフィルタを、 3×3 などのサイズで近似したものとみなせることに注意されたい。上でみたガウスフィルタなどは正則で $H(u, v) \neq 0$ を満たすが、一般の $h(x, y)$ に対し $H(u, v)$ は零点を有するうえ、対象とする画像中にも式(3・7)から様々な誤差が存在するため、 $|H(u, v)|$ が0に近づくほど式(3・9)による単純な復元は著しく不安定となる。例えば、通常のカメラなどから得られる焦点ぼけ画像は、式(3・7)による焦点ぼけ過程の後に雑音が重畳されている。白色ガウス雑音などは、図3・2の $|G(u, v)|$ 上すべての周波数成分へ影響を与えるため、特に $|H(u, v)|$ が0に近づく高周波において式(3・11)の復元画像上に目障りな品質低下、すなわちアーティファクトが生ずる。このように、焦点ぼけや雑音などの劣化が組み合わさった現実的な問題設定で、式(3・11)のみから $F(u, v)$ を安定に復元することは困難であり、前節でより一般的な、超解像含む画像復元の課題として述べたとおり、復元しようとする画像信号に対し、何らかの望ましい性質をどれほど満たしているかなどの観点で評価関数を定め、これをできるだけ最小化する形で解を得るような、問題に正則化を施す手法がさかんに研究されている。

なお、式(3・7)からの誤差としては、画像端の問題もあるが、これについては窓関数の適用で比較的簡単に対応することができる。すなわち、フーリエ変換では画像を垂直、水平それぞれ周期 N の信号とみなしているため、右端と左端、上端と下端の間で回り込む際にそれぞれステップ状の強い不連続性が生じ、その結果、図3・1の $|F(u, v)|$ でも u 軸、 v 軸の上へのみ高周波まで、これを表現するための成分が存在していることがみてとれる。式(3・8)そのものを用いて求められた $|G(u, v)|$ では、このような成分は存在しないが、これは $g(x, y)$ において両端の間で焦点ぼけも回り込み不連続性が解消されることに対応しており、通常のカメラなどで取得される焦点ぼけ画像とは状況が異なる。実際、 $g(x, y)$ の画像端は、回り込みではなく、画像の範囲外で生じている焦点ぼけの影響を受けるのが一般的であるため、 $f(x, y)$ と同様の不連続性を

有しており、それをフーリエ変換した $G(u, v)$ においても、図 3・1 の $|F(u, v)|$ で見られた成分が高周波まで存在する。したがって、上で述べた雑音の影響と同じく、式(3・11)をそのまま用いれば、復元画像に広範囲の強いアーティファクトが誘導されることとなる。これについては、復元前の画像信号 $g(x, y)$ に対しあらかじめ、画像中央から画像端に向かって正の値から 0 に近付く窓関数と呼ばれる重み付けを施し、画像端での不連続性を解消する前処理が有効である。この処理後のフーリエ変換 $G(u, v)$ では、 u 軸、 v 軸上の成分が抑制され、式(3・11)による復元画像中でも、大きなアーティファクトは発生しない。窓関数の適用による影響で、画像端そのものは良好には復元されないが、いわばその代償として、それ以外の領域で品質の良い復元結果を得るのである。窓関数としては、ハミング窓、ハンニング窓、ブラックマン窓など、様々な特性を持ったものが提案されている。

以上でみた、画像信号のフーリエ変換により低周波に成分が集中する特性は、その圧縮符号化などにおいても、たいへん有効に活用されている。このような圧縮においては、まず画像内容の骨格に相当する極めて低い周波数成分を維持するため、例えば画像全体を 8×8 画素単位などのブロックに分割してから、それぞれに変換を適用する手順をとっている。その変換も、これまでで前提としていた複素数ではなく、実数の範囲で簡潔に表現するなどして様々な効率を高めるため、DCT (Discrete Cosine Transform, 離散コサイン変換) が一般に採用される。この場合の DCT は、 $K \times K$ 画素単位の離散信号 $b(x, y)$ を第 1 象限の奇数座標に配置しそれ以外の偶数座標は 0 としたうえで、他の象限には各軸対称の信号を定めてから $4K \times 4K$ 単位のフーリエ変換を施す、といった手順などで構成できる。すなわち、各軸対称の信号は、式(3・4)の各基底の虚部が打ち消され、結局、

$$B(u, v) = 4 \sum_{x=0}^{K-1} \sum_{y=0}^{K-1} b(x, y) \cos(\pi u(2x+1)/2K) \cos(\pi v(2y+1)/2K) \quad (3 \cdot 12)$$

となつて、実数部の余弦のみで処理が完結する。また、対応する逆変換は、 $B(u, v)$ が $4K \times 4K$ の範囲で持つ周期性、対称性などを考慮しながら、軸上は場合分けを行って、

$$\begin{aligned} b(x, y) &= \frac{1}{K^2} \left(\frac{1}{4} B(0, 0) \right. \\ &+ \frac{1}{2} \sum_{v=1}^{K-1} B(0, v) \cos(\pi v(2y+1)/2K) + \frac{1}{2} \sum_{u=1}^{K-1} B(u, 0) \cos(\pi u(2x+1)/2K) \\ &\left. + \sum_{u=1}^{K-1} \sum_{v=1}^{K-1} B(u, v) \cos(\pi u(2x+1)/2K) \cos(\pi v(2y+1)/2K) \right) \end{aligned} \quad (3 \cdot 13)$$

で与えられる。ここで、 $B(u, v)$ も $K \times K$ 単位で扱えて、すべて実数で構成されることに注意されたい。

図 3・3 に $K=8$ のもと、式(3・12)の DCT 変換により得られる係数と、このうち絶対値の大きな順に 1/10 の係数だけ選択し、それ以外は 0 とした場合からの式(3・13)による再構成結果を示した。図 3・3 の各ブロックでは、DCT 係数の絶対値について、左上から右下に向け高い周波数に対応するよう表現した。強いテクスチャを含むブロックのみ高周波成分が存在し、また、DCT 係数の選択によってかなりの情報が削減されているにも関わらず、再構成結果では僅かにブロック境界が視認できる程度で画像内容が十分に保持されていることも、図 3・3 からみてとれる。実際、このような特性に基づき、例えば、標準方式の一つである JPEG では、高周波ほど粗い量子化を施したうえで、各ブロックを低周波から高周波に向けジグザグスキャンした

信号にハフマン符号化を適用することで、有効な画像圧縮を実現している。

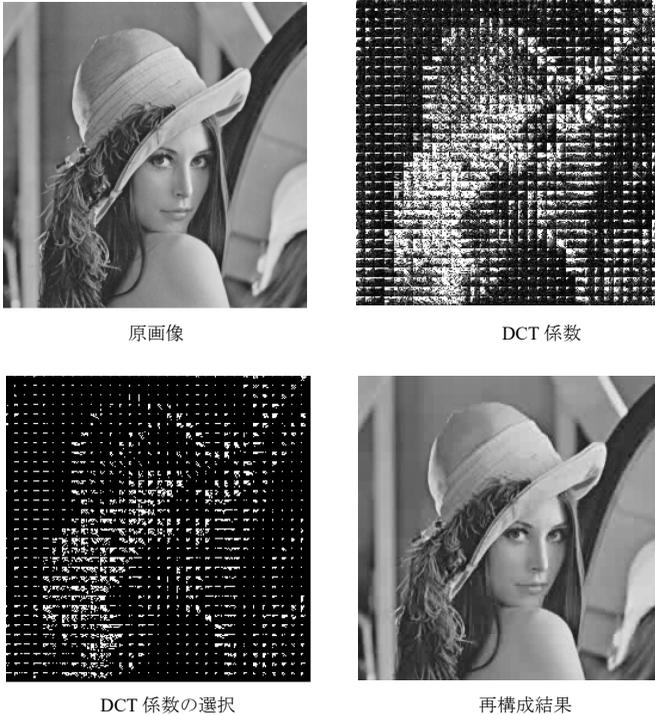


図 3・3 DCT 変換の例

2-3-3 ウェーブレット変換

前項でみたフーリエ変換は、三角関数で構成される周期信号を基底としており、いわば画像信号全体をひとまとめに分析、処理する形となっていた。一方、ウェーブレット変換では、基底として範囲の小さな波を空間方向にシフトさせ、画像内の局所的な特性をそれぞれ分析、処理することで、より柔軟な機能を実現している。式(3・8)以降でみたような、畳み込み演算の複素乗算への置き換えと、それに基づく周波数特性によるフィルタの特徴づけなどは、フーリエ変換と異なり、用いることができないが、画像信号を扱うにあたって周波数ごとの処理だけでなく、合わせてその空間座標ごとの適応的な処理が可能となる意義はたいへん大きい。

1 次元信号を対象としたウェーブレット変換では、適当な条件を満たすウェーブレットと呼ばれる関数 ψ をもとに、

$$\phi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (3 \cdot 14)$$

を基底としてとる。この基底のパラメータは 2 つで、それが一つのフーリエ変換と異なり、 a がいわば周波数、 b が新たに空間座標に対応することに注意されたい。ここでは、離散的な信号 $f(x)$ に対象を限定し、これを表現する最も簡単なものとして、

$$\psi(t) = 1(0 \leq t < 1/2), -1(1/2 \leq t < 1), 0(t < 0, 1 \leq t) \quad (3 \cdot 15)$$

のハール (Haar) ウェーブレットを考える. $j(> 0), k$ を整数として $a = 2^j, b = 2^k$ と離散化した基底を,

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k) \quad (3 \cdot 16)$$

と改めて書けば, 直交性は式(3・15)から簡単に確認でき,

$$f(x) = C + \sum_{j,k} w_j(k) \psi_{j,k}(x) \quad (3 \cdot 17)$$

の形で, 元の信号 $f(x)$ を直流成分 C を除き, ウェーブレット係数 $w_j(k)$ により特徴づけることが考えられる. 実際, 直交性から, これらの係数は,

$$w_j(k) = \sum_x f(x) \psi_{j,k}(x) \quad (3 \cdot 18)$$

で求められるが, 一般に, より効率的な演算が存在する. すなわち, 特に簡単な実装では, $s_0(m) = f(m)$ から始めて, ウェーブレット関数に対応して定まるパラメータ p_n, q_n を用い,

$$\begin{aligned} s_{j+1}(m) &= \sum_k p_{k-2m}^* s_j(k), \\ w_{j+1}(m) &= \sum_k q_{k-2m}^* s_j(k) \end{aligned} \quad (3 \cdot 19)$$

なるある種の畳み込み演算を繰り返し実行することで, 所望の係数が $w_1(m), w_2(m), \dots$ の順に求っていく. 式(3・15)のハールウェーブレットの場合は, $p_n = q_n = 0 (n \neq 0, 1)$ 及び $p_0 = p_1 = q_0 = -q_1 = \alpha$ なる定数を与えることになり, $\alpha = 2^{-1/2}$ とすれば正規化される. 画像信号処理においては2次元のウェーブレット変換が必要となるが, フーリエ変換と同様, 各列を処理した後, 各行に対し同じ処理を施せばよい. ただし, 式(3・19)の演算ごとにその処理後の結果は, $s_{j+1}(m), w_{j+1}(m)$ の順に左右または上下に並べるものとする.

図3・4に, 以上に述べた手順を2段階まで行ったウェーブレット変換の例を示した. 図3・4中では見やすさのため, 係数部分については灰色を0として表現し, それ以外の左上部分は1/4を乗じて原画像と明るさを一致させている. なお, 本節のここまでと異なり, 原画像には白色ガウス雑音を加算したものをを用いていることに注意されたい. 先に述べたように, ウェーブレット変換は雑音除去のような空間座標ごとの処理に適しており, 実際, 図3・4でも示すように, 絶対値が適当な閾値未満となる係数を0で置き換えたとしても, 対応する局所領域にのみその効果を限定できる. このような処理はウェーブレット縮退 (Shrinkage) と呼ばれ,

$$s_{j-1}(m) = \sum_k (p_{m-2k} s_j(k) + q_{m-2k} w_j(k)) \quad (3 \cdot 20)$$

の逆変換により再構成した図3・4の結果では, 画像全体に大きな影響を与えることなく, 雑音が抑制されているのがみてとれる.

以上ではハールウェーブレットを前提に説明したが, テクスチャなどの画像認識には解析特性の良いガボール (Gabor) ウェーブレットが, JPEG 2000 などの圧縮を含めた様々な処理には視覚特性の良いドブシィ (Daubechies) ウェーブレットがしばしば応用される. また, 単純な基底の取り換えでなく, 画像信号の有する方向性なども考慮しながらウェーブレット変換の枠組そのものを拡張するような試みが, ある種の多重解像度解析として提案され続けており, 今

後とも多様な展開が期待されている。



原画像



ウェーブレット変換



ウェーブレット係数の縮退



再構成結果

図 3・4 DCT 変換の例ウェーブレット変換の例

■2群-1編-2章

2-4 画像の特徴と分析

(執筆者：久保田彰) [2018年12月受領]

2-4-1 近傍演算

ある画素から近い距離にある画素の集合をその画素の近傍領域と呼ぶ。例えば、あの画素を中心とした、縦横のサイズが3画素である正方領域はよく使われる。これは、 3×3 近傍などと呼ばれる。近傍演算とは、入力画像 I_{in} のある画素 (m, n) の近傍領域 R の画素の輝度値を入力として、ある定まった処理 f により出力を算出し、それを出力画像 $I_{out}(m, n)$ の輝度値に割り当てる処理のことをいう。

$$I_{out}(m, n) = f(\{I_{in}(m+i, n+j) \mid (i, j) \in R\}) \quad (4 \cdot 1)$$

近傍演算では、一般に近傍領域 R と処理 f は画素 (m, n) に依存しないことが多い。処理 f には、例えば、最小値や中間値、平均値などを出力する処理がある。また、広義の意味では、近傍処理のことをフィルタリングと呼ぶことがある。

2-4-2 フィルタとオペレータ

画像の線形フィルタリングは最も代表的な近傍演算である。フィルタリングは、通常、 3×3 や 5×5 などの近傍領域に対して定義された係数 $H(i, j)$ を用いて、近傍領域の画素値の線形和を算出し、出力画像の画素値に割り当てる処理をいう。

$$I_{out}(m, n) = \sum_{i,j} H(i, j) I_{in}(m+i, n+j) \quad (4 \cdot 2)$$

係数 $H(i, j)$ をフィルタマトリックスあるいは単にフィルタと呼ぶ。フィルタマトリックス $H(i, j)$ の値を変えることによって、画像のエッジを強調したり、ノイズを除去したりするといった様々な処理が可能である。それぞれの目的に合ったフィルタ $H(i, j)$ が提案されており、特にオペレータと呼ぶ。

フィルタ $H(i, j)$ を反転した $H'(i, j) = H(-i, -j)$ を用いると、式(4.2)は

$$I_{out}(m, n) = \sum_{i,j} H'(i, j) I_{in}(m-i, n-j) \quad (4 \cdot 3)$$

と表される。すなわち、出力画像 I_{out} は H' と入力画像 I_{in} との畳み込み和で表される。このとき、 H' もフィルタあるいはオペレータと呼ぶが、特にフィルタカーネルと呼ばれる。

上記の処理はどちらも画像のフィルタリングと呼ばれるが、一般的には、畳み込み和による処理のことを指すことが多い。ただし、対称なフィルタを用いる場合は、どちらの処理を用いてもよい。

2-4-3 エッジ・線の検出

エッジの検出は、画像のフィルタリングを利用して行われる。エッジは、ある方向に対して輝度値が急激に変化する画素を表している。そのため、輝度値の変化（輝度勾配）を捉えるためのオペレータ（フィルタ）が多く提案されている。

エッジの検出には、水平方向の変化を捉えるオペレータ H_x と垂直方向の変化を捉えるオペレータ H_y を用いる。まず、入力画像 I に対して、それぞれのオペレータによるフィルタリングを行い、出力画像を得る。

$$I_x = H_x * I, \quad I_y = H_y * I \quad (4 \cdot 4)$$

次に、得られた画像から、勾配の大きさを以下のように計算し、エッジを検出した画像 I_{edge} を得る。

$$I_{\text{edge}} = \sqrt{I_x^2 + I_y^2} \quad (4 \cdot 5)$$

代表的なエッジ検出方法として **Prewitt** オペレータや **Sobel** オペレータを用いた方法がある。以下に **Sobel** オペレータを示す。

$$H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad H_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (4 \cdot 6)$$

2-4-4 零交差オペレータ

画像に2次微分をかけて、その結果の零交差点(符号が変化する点)を見つけることにより、エッジを検出する処理である。これにより、正確にエッジの位置を特定することができる。ただし、2次微分により画像のノイズを強調するので、あらかじめ平滑化処理を行う。代表的な零交差オペレータとして、**Laplacian-of-Gaussian (LoG)** がある。

2-4-5 2値化閾値選択

濃淡画像を2値化し、前景領域と背景領域を分離するためには、最適な閾値を選ぶ必要がある。2値化には、すべての画素で一定な閾値を用いる方法(大域的閾値選択法)と画素ごとに变化する閾値を用いる方法(局所的閾値選択法)がある。

大域的閾値選択法では、画像のヒストグラムの統計量を用いて閾値を決定する。代表的な例として、大津の方法やエントロピー最大化法などがある。大津の方法では、濃淡画像が前景と背景の各クラスから発生した画素から構成されると仮定し、クラス間の分散が最大となるような閾値を決定する。

局所的閾値選択法では、各画素の近傍領域内の統計量を利用して閾値を決定する。例えば、**Bernsen**の方法では、近傍領域の最大値と最小値の平均値を閾値として用いる。

2-4-6 ハフ変換

エッジなどの特徴点が並んでできる直線を検出する方法である。検出は、直線を表すパラメータを見つけることによって行われる。直線のパラメータとして、通常、その傾きと原点からの距離 ρ を用いる。特徴点の座標を (x_i, y_i) とする。それを通る直線は無数に存在するため、直線の傾きを θ と仮定し、そのときの距離 ρ を計算する。距離 ρ は、

$$\rho = x_i \cos \theta + y_i \sin \theta$$

により算出される。 θ を変えながら、 ρ を計算し、 θ - ρ 座標に値(例えば +1)を投票する。この操作をすべての特徴点で行う。その結果、 θ - ρ 座標に投票値が積算される。投票値の多い座

標を求め、直線の傾きと原点からの距離を算出する。これにより、最も多くの特徴点を通る直線を検出できる。この計算は、 θ - ρ 座標を細かく分割して行われる。

同じ原理を円や楕円などの一般の図形の検出に応用した一般化ハフ変換も確立されている。

2-4-7 細線化と骨格化

2値画像において前景領域の画素を境界から順次削除するような処理を細線化と呼ぶ。細線化は、領域の連結性を保ちつつ行われ、最終的には前景領域は幅が1画素の線分に変換される。

前景領域の境界までの距離が極大となる画素の集合を骨格と呼び、これを求める処理を骨格化と呼ぶ。

細線化及び骨格化は、前景領域の基本的な形状を調べるための前処理として用いられる。

2-4-8 形状解析

2値画像における前景領域の形状は、重心、面積、境界の長さ(周囲長)、外接長方形の各辺の長さ、主軸方向などによって特徴を記述することができる。

ほかには、境界画素の2次元座標を複素平面上の座標とみなし、それらを離散フーリエ変換して得られるデータも形状特徴として用いられる。これをフーリエ記述子という。低域成分のみを用いて逆離散フーリエ変換すると、領域の概形が得られる。

2-4-9 膨張収縮

前景領域(値が1の黒画素)と背景領域(値が0の白画素)から成る2値画像に対して、前景領域の近傍にある白画素を黒画素に変換して前景領域を拡大する処理を膨張、逆に背景領域の近傍にある黒画素を白画素に変換して前景領域を縮小する処理を収縮という。

どの範囲が注目画素の近傍であるかは、構造要素と呼ばれる2値画像で定める。近傍と定める画素の値を1、それ以外の画素の値を0とする。例えば、注目画素の上下左右の4画素を隣接として定義したい場合、構造要素は注目画素を中心に十字型の3×3の画像となる。

2-4-10 ラベリング

2値画像において、繋がっている黒画素の領域を分類し、それぞれに異なる番号(ラベル)を付ける処理をラベリングと呼ぶ。黒画素の繋がりは、連結性と呼ばれ、4連結と8連結の2つが主に使われる。4連結では、上下左右に隣接した画素が連結されていると定義し、8連結では、4連結の画素に加えて、斜め方向に隣接した画素も連結されていると定義する。

2-4-11 モルフォロジー

処理対象の2値画像と構造要素との集合演算により、2値画像の処理を行うことを総称してモルフォロジー演算と呼ぶ。構造要素は、様々な形状を表す2値画像である。その形状が処理結果に強く反映される。

モルフォロジー演算では、構造要素を移動させながら、2値画像との近傍演算を行い、その結果に応じて画素値を反転させて、前景領域の形状を変形する。

膨張と収縮及びそれらの組合せは代表的なモルフォロジー演算である。膨張の後に収縮を行う処理をクローズング、収縮の後に膨張を行う処理をオープニングと呼ぶ。これらを更に組み

合わせるとノイズ除去に有効であることが分かっている。

2-4-12 セグメンテーション

画像を同じ特徴を持つ領域に分割する処理のことをセグメンテーション（領域分割）という。特徴量としては、色の輝度やテクスチャ特徴が用いられる。セグメンテーションは、エッジ検出に基づいた方法と領域特徴に基づいた方法に分けられる。前者は、検出したエッジを用いて、それらで囲まれた領域を抽出する。ユーザが与えた閉曲線からエッジを逐次推定するスネーク法や背景と対象領域との識別を効率良く行うグラフカット法などがある。後者は、隣接した画素あるいは領域の特徴の類似性に基づいて、それらを統合していく方法である。

2-4-13 テクスチャ

テクスチャとは、布目の模様のように規則的に輝度に変化しているパターンのことをいう。規則的なパターンは、パワースペクトル、輝度ヒストグラムの統計量、同時生起行列の特徴量などによって特徴が分析され、識別される。

2-4-14 特徴抽出

画像中の対象などについて、カラー、テクスチャ、形状、構図といった、その内容を特徴づける何らかの情報を、画像から抽出する処理を「特徴抽出」という。多くの場合、複数の特徴量を画像から抽出し、後述するように画像内容の分類などに活用する。

2-4-15 不変量

画像内容を特徴づけるにあたっては、本質的でない条件が変動しても一定の特徴量が抽出されることが望ましい。例えば、対象の位置や姿勢はもちろん視点、照明などの撮影時の環境などが多様な条件下であっても不変となる特徴量を活用できれば、極めて安定に画像内容の分類などが可能となる。特に、射影不変量、アフィン不変量といった幾何学的な不変量はよく調べられており、画像内容の平行移動や回転、拡大縮小などに対し安定に本質的な特徴量が抽出されている。

2-4-16 動き特徴

画像の時系列や、複数の視点から取得した画像群などにおいて、画像中の対象が画像間でどのように動くに着目した特徴を「動き特徴」という。動き特徴により、単一の画像からでは得ることが難しい、対象の様々な性質が捉えられる。一般に、勾配法のオプティカルフローやブロックマッチングなどにに基づき、画素、ブロック、領域と多様な単位で動き特徴が抽出、活用される。

2-4-17 分類

画像から抽出された特徴量を用いて、その内容についてカテゴリー分け、すなわち分類することができる。一般に、複数の特徴量で構成される多次元空間上で、このようなカテゴリー分けを行うため、機械学習などの分野で様々な理論が整備され続けている。