

2群(画像・音・言語) - 2編(パターン認識とビジョン)

2章 画像特徴抽出・照合

(執筆者: 石川 博)[2010年5月受領]

概要

画像から情報を引き出すには、しばしばその一部に注目してほかは無視することが必要である。例えば物体認識では、背景が何であろうと前景にある物体のみに注目する。あるいは類似画像を検索したい場合、一部が類似していればほかは無視して検索したいことが多い。ほとんどの応用において、画像の含む膨大な情報の中から目的に応じて注目する情報を抽出し、それに基づいて処理をする必要がある。

画像から引き出したい情報のうち、画像に写っている対象の位置や画像内での配置が最も基本的である。したがって、位置を正確に決定できる特徴、つまり特徴点の検出が重要である。画像認識の初期から、輝度あるいは色が急激に変化する点、すなわちエッジを検出する方法が数多く提案されてきた。エッジは一般に物体の縁などに検出されるが、これは1次元的に曲線として表れる。空間内での位置をより正確に決定するためには、点として表れる特徴が必要だが、物体の角などがこれにあたり、これを検出するコーナー検出器などの特徴点抽出器 (Interest Operator) が用いられる。

特徴点は画像のスケールによって変化しないことが望ましい。そのために画像をスケールによってパラメータづけられた族として表現し、すべてのスケールを同時に扱う Scale Space 理論が展開されるとともに、有限長の波形で画像を表現し、信号を空間及び周波数の両領域で局所的に特徴づけることができる Wavelet がパターン認識とビジョンにも導入されてきた。

画像パターン認識で最も基本的なのは、画像そのものの一部をパターンとして画像の部分と照合するテンプレートマッチングである。例えば、テンプレートを検出対象の標準画像とすることで、入力画像中の特定物体検出を実現する。より一般には、まず上記のように画像から特徴を抽出する。それらの特徴はそれぞれ例えば画像内での位置や方向などのパラメータをもつが、それらのパラメータ空間内での配置をパターンとして照合する。そのような照合方法には1次元的な弾性マッチング手法である DP マッチングや、特徴を頂点とし、特徴間の関係を辺とするグラフをつくり、グラフ間の対応づけを行うことにより特徴集間を照合するグラフマッチングがある。

【本章の構成】

本章では特徴抽出 (2-1 節) に関して、Interest Operator (2-1-1)、Scale space (2-1-2)、Wavelet (2-1-3)、SIFT (2-1-4) について、また照合 (2-2 節) の代表的な方法として、テンプレートマッチング (2-2-1)、DP マッチング (2-2-2)、グラフマッチング (2-2-3) について述べる。

2群 - 2編 - 2章

2-1 特徴抽出

(執筆者: 佐川立昌, 石川 博)[2009年9月受領]

画像中の特徴として、特徴点は最も基本となる要素である。ここで特徴点とは、輝度値あるいは色が周囲の画素と区別でき、その位置を正確に決定することができる点である。画像中の特徴点を抽出することにより、複数の画像間の対応づけが容易になるため、様々な応用に用いられる。このような特徴点を抽出する方法として、特徴点抽出器 (Interest Operator) が用いられる。特徴点は画像のスケールによって変化しないことが望ましい。そのために Scale Space 理論が展開されるとともに、応用数学から Wavelet がパターン認識とビジョンにも導入されてきた。現在最もよく使われる特徴量の一つである SIFT オペレータは、この分野におけるそのような発展の結果である。

2-1-1 Interest Operator

画像中で輝度値・色が一樣な領域では、周囲の画素と区別が困難となるため、位置を正確に決定することができない。したがって特徴点として、輝度値・色の変化が大きい部分が選ばれることになる。また、画像の2次元座標を正確に決定するためには、画像の縦・横の両方向に変化が大きい必要がある。画像中で角に見える部分が、このような特徴をもつため、コーナー検出器 (corner detector) とも呼ばれる。

特徴点抽出法は、これまで様々な方法が提案されているが、よく使われる方法として下記の手法が挙げられる。

- SUSAN オペレータ¹⁾
- Harris オペレータ³⁾
- Shi と Tomasi の方法 (Good features to track)²⁾

まず、SUSAN (Smallest Univalued Segment Assimilating Nucleus) オペレータは、ある点の周りに円形マスク S を考え、中心点 p_0 とマスク内の画素 p を比較し、濃淡画像の輝度値差が小さい点の数を数える方法である。すなわち、 g と t を閾値パラメータ、 $I(p)$ を点 p の輝度値とすると SUSAN オペレータの結果 $R(p_0)$ は

$$R(p_0) = \max(0, g - n(p_0)), \quad n(p_0) = \sum_{p \in S} c(p_0, p), \quad c(p_0, p) = \begin{cases} 1 & \text{if } |I(p_0) - I(p)| \leq t \\ 0 & \text{if } |I(p_0) - I(p)| > t \end{cases} \quad (2.1)$$

で与えられ、値が大きい画素が特徴点として抽出される。

次の Harris オペレータと J. Shi と C. Tomasi の方法は、ほぼ類似した方法であり、近傍の点と区別しやすい特徴点として、ウィンドウ領域内の画像を微少シフトさせたとき、輝度値差の二乗和 (sum of squared difference, SSD) が大きくなる点を検出する。点 p を中心とするウィンドウ W において、 Δv だけ微少シフトさせたときの SSD を

$$S(p) = \sum_{q \in W} (I(q) - I(q + \Delta v))^2 \quad (2.2)$$

とし、シフトした画像をテイラー展開により 1 次近似すると、 $I(q)$ の x, y 方向の偏微分 $I_x(q)$ 、 $I_y(q)$ を用いて、

$$I(q + \Delta v) \approx I(q) + [I_x(q) \ I_y(q)] \Delta v \quad (2.3)$$

と表される。この 2 式をまとめると、

$$S(p) = \Delta v^T \begin{bmatrix} \sum_W I_x^2 & \sum_W I_x I_y \\ \sum_W I_x I_y & \sum_W I_y^2 \end{bmatrix} \Delta v = \Delta v^T H(p) \Delta v \quad (2.4)$$

となる。行列 $H(p)$ によって点 p 周りの輝度値分布の特徴が表されており、Harris オペレータと Shi と Tomasi の方法では、この行列の固有値を用いて特徴点を検出する。二つの固有値を λ_1, λ_2 とすると、両方とも大きい点の特徴点として適しているため、Harris オペレータでは特徴量として、

$$M = \lambda_1 \lambda_2 - \kappa (\lambda_1 + \lambda_2)^2 = \det(H) - \kappa \text{trace}^2(H) \quad (2.5)$$

が大きくなる点を抽出する。ここで κ はパラメータである。一方、Shi と Tomasi の方法では特徴量として、

$$M = \min(\lambda_1, \lambda_2) \quad (2.6)$$

を用いる。

2-1-2 Scale Space

自然画像は異なるスケールにおいては異なる構造を有する。例えば木の枝というものは数 cm からたかだか数 m のスケールでしか意味を持たず、nm や km のスケールで木の枝というものを考えることはできない。一方、画像から自動的に情報を得ようとするとき、どのスケールを見る必要があるか事前に知ることはできない。この問題に対処するため、A. Witkin⁷⁾ によって提案されたスケールスペース (scale space, 尺度空間) 理論は、画像をスケールによってパラメータづけられた族として表現し、すべてのスケールを同時に扱う。

画像のスケールスペース表現とは、与えられた画像を漸進的にぼかした画像を積み上げたもの (図 2.1) のことである。スケールに従ってぼかす方法にはいろいろ考えられるが、以下のようにガウスクアーネル (Gaussian kernel) を使えば、元の画像にない構造 (例えばゼロ交差) がぼかしたことにより現れることがないことが知られている。

与えられた画像を実平面上の関数 $I(x, y)$ と考えると、この画像のスケールスペース表現は、スケールを表すパラメータ $t > 0$ を加えた関数 $L(x, y, t)$ 、すなわち実平面とスケールパラメータの空間の直積空間上の関数 $L: \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$ として与えられる⁶⁾。

L において $t = 0$ としたものが元の画像である：

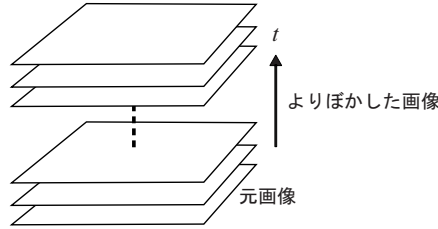


図 2.1 スケールスペース

$$L(x, y, 0) = I(x, y). \tag{2.7}$$

そして、パラメータ $t > 0$ に従って、分散 t のガウス関数

$$g_t(x, y) = \frac{1}{2\pi t} \exp\left(-\frac{x^2 + y^2}{2t}\right) \tag{2.8}$$

との畳み込みによりぼかした画像が $L(x, y, t)$ として与えられる：

$$L(x, y, t) = (g_t * I)(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_t(u, v) I(x - u, y - v) du dv. \tag{2.9}$$

したがって、 t が大きくなるほど、元の画像のよりぼかした画像が与えられる。

特徴抽出においては信号の局所的な変化を検出する微分処理が重要であるが、任意のスケール t におけるスケールスペース微分、すなわちそのスケールに対応してぼかした画像の微分は、

$$\frac{\partial}{\partial x} L(x, y, t) = \left(\frac{\partial}{\partial x} g_t * I\right)(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial}{\partial x} g_t(u, v) I(x - u, y - v) du dv \tag{2.10}$$

のように、スケールごとの微分カーネルとの畳み込みで表される。実際の処理においては、与えられた画像をぼかしてから微分するより、あらかじめデータとして用意した線形フィルタをかけるだけでよいことは大きな利点である。これは次項で紹介するウェーブレット解析とも関係する。

2-1-3 Wavelet

空間的に一様な三角関数を基底として使うフーリエ解析には、空間領域で信号を局所的に特徴づけることができないという問題がある。これに対して、ウェーブレット (wavelet) 解析は有限長の波形で信号を表現し、信号を空間及び周波数の両領域で局所的に特徴づけることができる方法である¹³⁾。画像においては特に特徴の局所性が重要であるので、フーリエ解析からウェーブレット解析への発展は画像応用において特に重要である。

D. Gabor¹²⁾は三角関数にガウス関数をかけて空間的に局所化したフィルタ

$$g_{\lambda, \theta, \varphi, \sigma, \gamma}(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \varphi\right) \tag{2.11}$$

$$(x', y') = (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta) \tag{2.12}$$

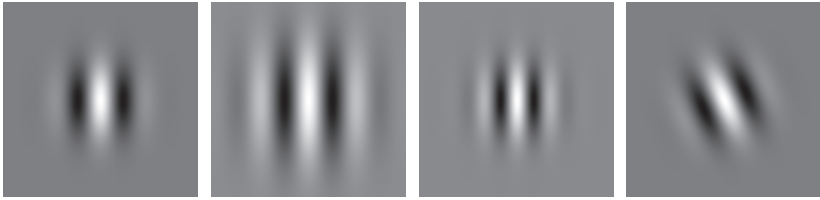


図 2・2 Gabor フィルタ $g_{\lambda,\theta,\psi,\sigma,\gamma}$. 左端を標準に，順に σ, λ, θ を変化させた例．

を使って信号の局所的な解析を試みた (図 2・2)．しかしこの Gabor フィルタは，周波数領域における広がり固定されているところに問題があった．

一般に， L^1 及び L^2 ノルムが有限，すなわち

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi(x, y)| dx dy < \infty, \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi(x, y)|^2 dx dy < \infty \quad (2\cdot13)$$

を満たす可積分関数 $\psi(x, y)$ をマザーウェーブレットと呼ぶ．このとき

$$W_s f(x, y) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) \psi\left(\frac{u-x}{s}, \frac{v-y}{s}\right) dudv \quad (2\cdot14)$$

を関数 $f(x, y)$ の連続ウェーブレット変換と呼ぶ．A. Grossman と J. Morlet⁸⁾は連続ウェーブレット変換を研究し，最初に地震学上のデータの解析に応用した．

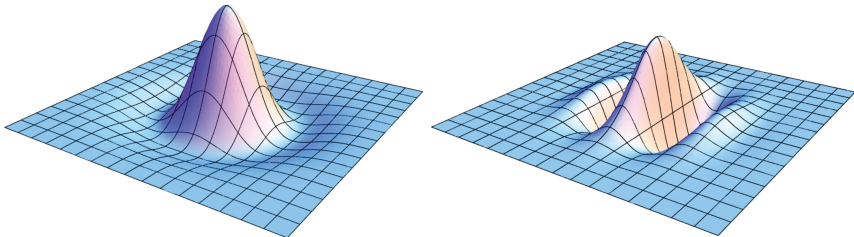


図 2・3 Laplacian of Gaussian(左) と Gabor Filter(右)

その後，I. Daubechies⁹⁾はある程度滑らかでコンパクトな台をもつ (つまり値が 0 でない領域が有界な) 正規直交ウェーブレット基底を構成し，S. Mallat^{10, 11)}は Y. Meyer とともに多重解像度解析 (multiresolution analysis) と呼ばれる，ウェーブレット基底を構成する一般的方法を提案した．ウェーブレットは信号処理，画像解析，通信システム，レーダ，システム制御など，多彩な分野で応用されている．

特徴点検出で使われる代表的なウェーブレットとして，上記の Gabor フィルタのほかに，次項に述べる SIFT にも使われる Laplacian of Gaussian (LoG) や Difference of Gaussian

(DoG) がある．式 (2・8) のガウス関数に Laplacian を作用させたもの

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) g_r(x, y) \quad (2 \cdot 15)$$

が LoG で，画像 $f(x, y)$ との畳み込み関数のゼロ交差としてエッジを，極値として “blob” すなわち周りより明るい暗い点を検出することができる．

2-1-4 SIFT

D. Lowe によって提案された SIFT (Scale-Invariant Feature Transform) 特徴⁵⁾は，回転・スケール変化に不変，照明変化に頑健という特徴をもち，近年最もよく用いられる特徴量の一つである．SIFT アルゴリズムは下記のステップから構成されている．

1. 特徴点のスケール・位置の検出
2. 特徴点の絞り込み
3. 特徴点の向きの正規化
4. 特徴量の計算

まず SIFT アルゴリズムでは，スケール空間理論⁶⁾に基づいて，画像中の様々な大きさをもつ特徴点を見つける．様々なスケールの LoG フィルタをかけたとき，特徴点として用いるのに適切なスケールが極値をもつことを利用するものである．実際には，LoG フィルタの代わりに，標準偏差を様々に変えたガウス関数を用いて画像を平滑化し，その差を計算する DoG フィルタを用いて LoG フィルタを近似する．図 2・4 に示すように，ガウス平滑化とダウンサンプリングを併用してスケール空間を構築する．特徴点として適した画素では，X-Y-スケールの 3 次元から構成されるスケール空間において DoG フィルタの結果が極値をもつことを利用し，特徴点候補の位置とスケールを決定する．位置だけでなく，スケールも決定することにより，画像中での見えのスケールが変化した場合でも，同一の特徴量を計算でき，スケール不変の特徴量とすることができる．

上述の方法で検出された特徴点の候補には，エッジ上の点や，コントラストが小さいといっ

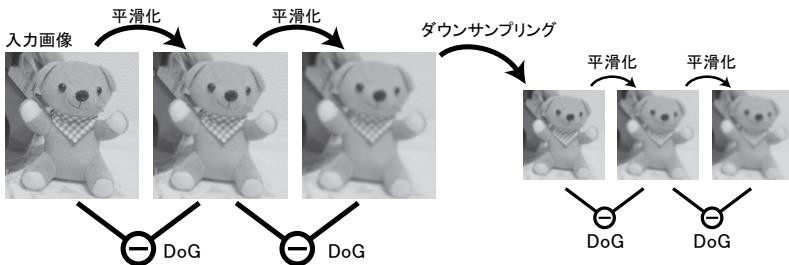


図 2・4 DoG フィルタによるスケール空間の構築

た、特徴点として適さないものも含まれている。そこで、これらを除外し、候補の絞り込みを行う必要がある。まず、DoG フィルタの結果を $D(p)$ とすると、エッジ上の点を取り除くため下記のヘッセ行列を計算し、その固有値を調べる：

$$H = \begin{bmatrix} D_x^2 & D_x D_y \\ D_x D_y & D_y^2 \end{bmatrix}. \quad (2.16)$$

Harris オペレータの場合と同様に、二つの固有値が共に大きい場合に特徴点として適している。SIFT アルゴリズムでは、固有値を計算する代わりに $\text{trace}^2(H) / \det(H)$ が閾値より大きい場合に、特徴点の候補から除外する。また、 $D(p)$ が小さいものを除外することにより、コントラストが小さい候補を除外する。

次に、特徴量を計算する前に、特徴の基準となる方向を求めることにより、特徴の向き正規化を行う。これにより、後述の方法によって計算する特徴量が、画像の回転に対して不変な特徴量とすることができる。上述方法によって決定したスケールのガウス関数によって平滑化した画像を L とし、特徴点周りのガウス窓内の点に対して、次の勾配強度 $m(p)$ と勾配方向 $\theta(p)$ を計算する：

$$m(p) = \sqrt{(L(p + \Delta x) - L(p - \Delta x))^2 + (L(p + \Delta y) - L(p - \Delta y))^2}, \quad (2.17)$$

$$\theta(p) = \tan^{-1} \left(\frac{L(p + \Delta y) - L(p - \Delta y)}{L(p + \Delta x) - L(p - \Delta x)} \right). \quad (2.18)$$

勾配方向を 36 段階に離散化して、勾配強度をヒストグラムに加算する。得られたヒストグラムの最大値をもつ方向を特徴の方向として決定する。ほかの方向に最大値の 80% 以上のものがある場合には、複数の方向を特徴の方向として選ぶ。

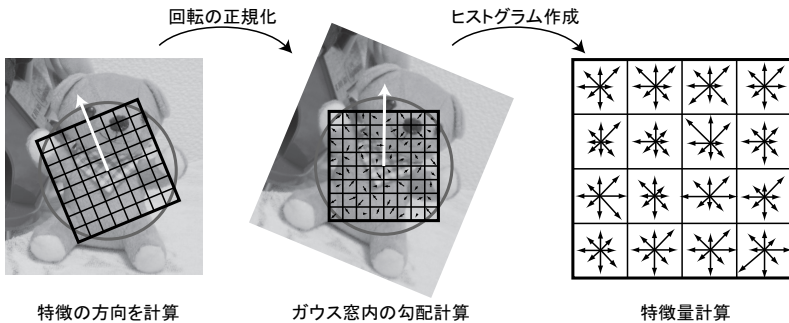


図 2.5 SIFT 特徴量の計算

最後に、マッチングなどに利用するための特徴量を計算する。図 2.5 に示すように、特徴量を計算する前に、上述の方法で計算した特徴の方向を基準として画像を回転しておく。次にガウス窓内の各画素の勾配を計算する。ガウス窓内の画素をブロックに分割し、ブロック内の勾配方向のヒストグラムを作成する。ここでは勾配方向を 8 方向に離散化し、ブロック数は 4×4 ブロックが良いとされている。ヒストグラムの各要素をベクトルとして表すことに

より, SIFT 特徴量とする. すなわち, $8 \times 4 \times 4 = 128$ 要素からなるベクトルとなる. また, 照明変化に頑健な特徴量とするため, ベクトルの長さを単位長に正規化したものを用いる.

2群 - 2編 - 2章

2-2 特徴照合

(執筆者：内田誠一，石川 博)[2009年9月受領]

画像から抽出された特徴を照合することにより目的のパターンを見つけるのがパターン照合である．ここでは代表的な方法として，テンプレートマッチング，DP マッチング，グラフマッチングを述べる．

2-2-1 テンプレートマッチング

(1) 原理

テンプレートマッチング(Template Matching)の一般的な目的は，画像 T (サイズ $M_T \times N_T$) と類似した部分を入力画像 I (サイズ $M_I \times N_I$) 中で検出することにある．具体的には，図 2-6 のように， T を I 上でずらしながら，重なった領域の類似度 S (もしくは相違度 D) を計算する．位置 (i, j) において，その値がある程度以上大きく (小さく) なったならば， (i, j) を検出位置とする．なお，画像 T のことを，検出対象の“型板”(テンプレート)と呼ぶ．

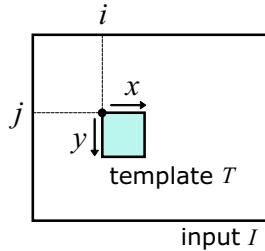


図 2-6 テンプレートマッチング

類似度として代表的なものは相互相関 (cross-correlation)

$$S_{CC}(i, j) = \sum_{(x,y) \in \mathcal{T}} I(i+x, j+y)T(x, y) \quad (2-19)$$

である．ここで \mathcal{T} は T の画素インデックスの集合 $\{[1, M_T] \times [1, N_T]\}$ である．この S_{CC} の計算過程は，入力 I に対して， T (より正確には 180° 回転した T) をインパルス応答とする FIR フィルタの出力値を求めていくことと等価である．この意味でテンプレートマッチングを整合フィルタ (matched filter) と呼ぶことがある．

相互相関 S_{CC} の代わりに，正規化相互相関 (normalized cross-correlation)

$$S_{NCC}(i, j) = \frac{S_{CC}}{\sqrt{\sum_{(x,y) \in \mathcal{T}} T^2(x, y)} \sqrt{\sum_{(x,y) \in \mathcal{T}} I^2(i+x, j+y)}} \quad (2-20)$$

を用いることもある．相互相関 S_{CC} に比べ， S_{NCC} は入力画像の輝度値変化に頑強である．なお，テンプレート T のノルムをあらかじめ 1 にしておけば，式 (2-20) の分母から T に関する部分を除去できる．

相違度としては、

$$D_p(i, j) = \left(\sum_{(x,y) \in \mathcal{T}} |I(i+x, j+y) - T(x, y)|^p \right)^{1/p} \quad (2 \cdot 21)$$

のような距離尺度を用いるのが一般的である。ここで $p = 1$ の時が市街地距離 (マンハッタン距離), $p = 2$ のときがユークリッド距離である。また, $p = \infty$ のときが最大値距離

$$D_{\max}(i, j) = \max_{(x,y) \in \mathcal{T}} |I(i+x, j+y) - T(x, y)| \quad (2 \cdot 22)$$

と等価になる。

2 値画像の相互相関及び正規化相互相関を求める場合には注意が必要である。具体的には, $I(i+x, j+y)$ と $T(x, y)$ が共に 0 (黒画素) の場合も, 片方が 1 (白画素) の場合も, 積 $I(i+x, j+y)T(x, y)$ は等しく 0 になってしまう点が問題になる。すなわち, 同じ画素値で整合している前者と整合していない後者とが同じ相関値になってしまう。文献 21) では, この点に配慮した様々な類似度が比較評価されている。

(2) 用途

テンプレートマッチングは極めて基本的な画像処理であり, 画像認識・理解の分野において様々な目的に用いられている。例えば, テンプレートを検出対象の標準画像とすることで, 入力画像中の特定物体検出が実現する。また, 動画像中の物体追跡も, 各時刻のフレーム画像内でテンプレートマッチングを行えば, 単純ではあるが実現できる。

また画像対 A, B について, その間の画素対応関係をテンプレートマッチングにより求めることができる。具体的には, 画像 A 中のすべての画素 (もしくは一部の特徴点) について, その周辺の小領域に最も対応する領域を画像 B 上に検出すればよい。これは対応点検出問題のテンプレートマッチングによる解法といえる。対応点検出の困難性については, 文献 22) に平易な解説がある。

画像対 A, B が異なる視点から同一物体を撮影したものであれば, この画素対応関係により視差が求まり, ステレオ視を実現できる。また, 時間的に連続した画像対であれば, この対応関係により画像中の動き, いわゆるオプティカルフロー^{23, 24)}が求まる。動画像圧縮標準 MPEG の動き補償は一般にこの方法 (ブロックマッチングと呼ばれる) が採用されている²⁵⁾。

(3) 高速化

式 (2 \cdot 19) に従って相互相関 S_{CC} を入力画像 I 上全体で計算するには, $O(M_I N_I M_T N_T)$ の計算量を要する。この計算量は決して小さなものではない。上記 (2) で述べたように 2 画像間に画素対応関係を求めようとすれば, その計算量は更に増大する。このため, 従来より様々な高速化法が検討されてきた。以下では, いくつかの高速化法について紹介する。

(a) 並列処理

並列処理は, 考え方としては最も単純な高速化法である。すなわち, S_{CC} の計算が (i, j) ごとに独立しているため, 例えば $M_I N_I$ 個のプロセッサがあれば, それぞれで異なる (i, j) での $S_{CC}(i, j)$ を計算すればよい。

(b) 粗密探索法

粗密探索 (coarse-to-fine strategy) では、まず I, T をそれぞれ低解像度化したうえで相互相関を求め、そのピークとしてテンプレート T の位置を定める。低解像度化により画像が小さくなるので、計算量を減らせる。例えば 2^{-n} 倍に低解像度化すれば、この計算量は $O(M_I N_I M_T N_T)$ の 2^{-4n} 倍で済む。

低解像度画像上で求めたテンプレート位置は、元の解像度から見れば不正確である。そこで、その位置を中心とした小範囲において、解像度を少し上げた画像 (例えば 2^{-n+1} 倍の解像度) 上でテンプレート位置を再探索する。同様の処理を解像度を上げながら繰り返して行けば、最終的には原画像 I 上での T の位置が求まる。

(c) 残差逐次検定法 (SSDA 法)

残差逐次検定法 (Sequential Similarity Detection Algorithm: SSDA)^{26, 27)} は、距離計算を途中で打ち切ることによって効率化を図る方法である。具体的には、式 (2・21) で $p = 1$ とした距離 D_1 を前提としたテンプレートマッチングにおいて、ある位置 (i, j) で誤差 $|I(i+x, j+y) - T(x, y)|$ を加算していく途中、何らかの閾値 θ を超えれば、そこに T がマッチする可能性は低いとして計算を打ち切る。距離 $D_1(i, j)$ には $M_T N_T$ 回の加算が必要になるので、初期の段階で打ち切りを決定できれば、それだけ効率化が図れることになる。効率化の度合いは、閾値をどのように設定するか依存する。加算回数とは無関係に閾値一定とする方法 (固定閾値) と、加算回数に応じて徐々に閾値を大きくしていく方法 (傾斜閾値) が考えられる。

固定閾値については、これまでに計算済みの D_1 の最小値を閾値 θ とすれば、SSDA を用いない場合と全く同じ最短距離マッチング位置を保証できる。これは、ある地点での D_1 を計算する途中で、別の位置で既に得られている D_1 の最小距離を超えてしまえば、その地点での D_1 は最小距離になり得ないためである。

傾斜閾値については、それを適切に設定できれば、固定閾値よりも高い効率化を期待できる。文献 26) では、 $r (\leq M_T N_T)$ 回加算時に対する閾値 θ_r として以下が提案されている。

$$\theta_r = \lambda(r + K \sqrt{r}) \quad (2 \cdot 23)$$

ここで K は安全率であり、 λ はテンプレート T と入力画像 I 間の差異 (雑音) を指数分布でモデル化した場合のパラメータである。したがって λ の設定には T と I に関する事前知識が必要になる。文献 27) ではこれを簡略化する方法が提案されている。

(d) アクティブ探索法

式 (2・19) に従って (i, j) を変えながら類似度を求めていく場合、 (i, j) のときと次の $(i, j+1)$ のときでは、類似度を評価する領域の大部分がオーバーラップしていることが分かる。したがって、 (i, j) と $(i, j+1)$ で求まる類似度は全く無関係ではない。この性質を利用し、もし既に求めた (i, j) での類似度よりも隣の $(i, j+1)$ での類似度が小さくなることを予想できれば、後者は計算不要であり、したがって高速化が図れることになる。

アクティブ探索法²⁸⁾ は、カラーヒストグラムにより上の性質を効果的に活用した方法である。今、 (i, j) における T と I のカラーヒストグラムの類似度を考える。すなわち、 T のカラーヒストグラム H_T と、 (i, j) を基準とした I の類似度評価領域 ($M_T \times N_T$) のカラーヒストグラム $H_{I(i,j)}$ 間の類似度

$$S_H(i, j) = \sum_k \min(H_T(k), H_{I,(i,j)}(k)) \quad (2\cdot24)$$

を考える (k はヒストグラムの中のピンのインデックス). ヒストグラムが正規化されていれば, $0 \leq S_H(i, j) \leq 1$ である. このとき, 次の関係が常に成り立つ.

$$S_H(i', j') \leq \frac{\min(S_H(i, j)|B|, |A \cap B|) + |A - B|}{|A|} \quad (2\cdot25)$$

ここで A, B は (i', j') 及び (i, j) を基準とした評価領域, $A \cap B$ は A と B のオーバーラップ領域, $A - B$ は A の領域から B の領域を除いた領域であり, $|\cdot|$ は画素数を示す (以上の議論では $|A| = |B|$ であったが, $|A| \neq |B|$ でも上式は成り立つ).

上式 (2.25) は, $S_H(i, j)$ から $S_H(i', j')$ の上限値を計算できることを示している. この上限値がこれまで求まっている S_H よりも小さければ, (i', j') において最大類似度が求まる見込みはなく, $S_H(i', j')$ の計算を省略できる. なお, 以上の処理はヒストグラム間の類似度 S_H に基づくものであった. もし S_{CC} のような画像レベルでのマッチングに基づく類似度を基準としたい場合は, S_H で検出位置の候補を挙げておき, それらの位置で改めて S_{CC} を計算し, 類似度を再評価すればよい.

(e) 高速フーリエ変換による方法

よく知られているように, 空間領域での畳み込みは, 周波数領域では積となる. 式 (2.19) で見たように相互相関 S_{CC} は画像 T と I の畳み込みで求まるので, 画像 T のフーリエ変換と I のフーリエ変換の (要素ごとの) 積は S_{CC} のフーリエ変換に等しい. そして, それを逆フーリエ変換すれば, S_{CC} が求まる. いずれの処理にも高速フーリエ変換 (FFT) を適用すれば, 高速化が図れる.

(4) 拡張

(a) 位相限定相関法

位相限定相関法 (もしくは位相相関法) は, 各画像のフーリエ変換の位相成分の複素共役積を用いて逆フーリエ変換する方法である²⁹⁾. 上記の方法では, 原理的に画像平面上でのテンプレートマッチングと同じ相互相関 $S_{CC}(i, j)$ が求まる. 一方, 位相限定相関法では $S_{CC}(i, j)$ と同じものは得られないが, ピークがより鋭くなるという効果が得られる. この効果は, フーリエ変換の平行移動に対する性質から容易に説明できる³⁰⁾.

(b) サブピクセルマッチング

一般のテンプレートマッチングでは i や j を整数値とする. これはそもそもの画素位置が整数値で表現されるためである. これに対し, 4.5 や 12.5 のように例えば 0.5 刻み (すなわち 1/2 画素刻み) にすることで, より精密な検出を目指す場合がある. これはサブピクセルマッチングといわれる. 動画画像圧縮標準 H.264/AVC では 1/4 画素レベルのマッチングに基づく動き補償が規定されている. 非整数値の位置には元々画素が存在しないため, マッチングの計算に際しては一般に何らかの補間挿入処理が必要となる. なお, 前出の位相限定相関法でも, 周波数領域でのピークモデルフィッティングにより, サブピクセルマッチングが可能である²⁹⁾.

(c) 弾性マッチング

2パターンをマッチングする際、一方に微小な歪み(ずれ)があると整合性が急に悪くなる。このため、マッチングの際、一方のパターンをゴムのように非線形伸縮させ、最も整合した時点をマッチング結果とすることがある。要するに画像をゴム膜のように変形させながらマッチングを図る技術であり、これは弾性マッチングと呼ばれる。弾性マッチングの実体は画素対応関係を表現する線形もしくは非線形写像の最適化であり、したがって様々な最適化手法がこれに利用されている^{31, 32, 33, 34}。

(5) 特徴点を用いたマッチング

以上のテンプレートマッチングはあくまで画像と画像の照合に基づくものであった。これに対し、両方の画像から複数の特徴点を抽出しておき、それら特徴点における特徴量に基づいてマッチングを求める方法がある。すなわち画像 I と T それぞれにおいて 2-1-4 節で述べた SIFT などによる特徴点を多数抽出し、次に T の特徴点それぞれについて、それと最も類似した特徴量をもつ特徴点を I 上に探索することで、 I と T の間に点对応を求める。その後、その点对応を用いて、最終的な位置合せ、すなわち T に対応する I の領域を求める。

図 2-7 は SIFT 特徴点・特徴量を用いた 2 画像 I, T のマッチング結果である。ここではテンプレート T として I の一部を回転させたものを用いている。回転による誤差で、 T と I の間には差異があることに注意されたい。同図右が上述の方法で点对応を求めた結果である。ただし、その類似度が閾値以下のものは省略している。単純な方法ながら、非常に高い精度でのマッチング実現されている。

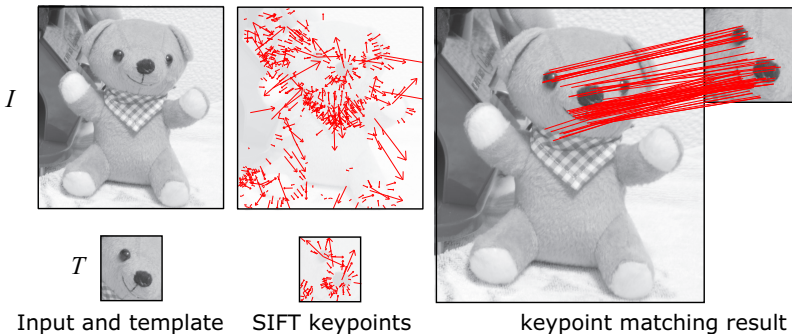


図 2-7 特徴点を用いたマッチング

位置合せを行うためには、この対応結果を用いて、 T の移動量や回転量の推定を行うことになる。より一般的に、いずれかの画像が射影変換を受けている場合は、射影変換パラメータの推定することになる。最小二乗法によりパラメータ推定を行うことも考えられる。しかし、点对応関係は不安定になることもあり、大きく誤った点对応が、全体的な推定結果に悪影響を与えることもあり得る。このため、RANSAC³⁵⁾のようなロバスト推定法が利用されることも多い。

2-2-2 DP マッチング

(1) 概要

パターン照合(マッチング)により音声やジェスチャなどの時系列パターンを認識する場合, 一般にパターンに発生した非線形時間伸縮や系列長の差異を補償する必要がある. すなわち, 1次元の弾性マッチング(2-2-1(c)参照)を行う必要がある. 動的計画法(Dynamic Programming, DP)を用いた弾性マッチングは, その代表的な手法であり, 一般にDPマッチング, もしくはDTW(Dynamic Time Warping)と呼ばれる.

DPマッチングの結果として得られるのは, パターン間距離と2パターン間の最適対応関係(図2・8(a))の二つである. 前者は, 俗にDP距離ともいわれ, 非線形伸縮に対して不変なマッチング距離であり, したがって変形を伴うパターンの認識に有効である. 長さが違う2パターンにおいても計算可能な距離である点も重要である. また対応関係は一方のパターンを基準としてもう一方の変形状況を表現したものであり, 構造解析の手法や統計的解析手法との接点となり得る.

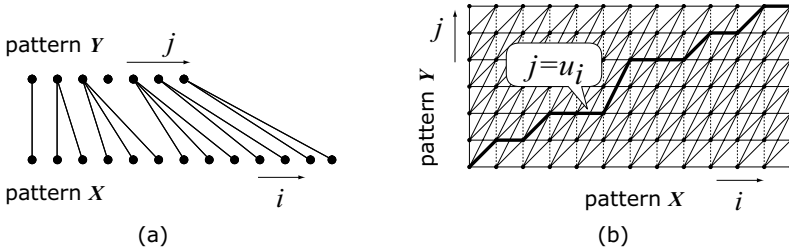


図 2・8 DP マッチング

DP マッチングは 1960 年代末から 1970 年代に音声認識^{36, 37, 38)}や文字認識向けに開発された. 理論的明快さ, アルゴリズム実装の容易さ, そして非常に少ない計算量ながら最適マッチングが求まる(長さ I と J の時系列パターンであれば, $O(IJ)$ の計算量で済む)という, 非常に優れた性質をもつ. このため時系列パターン認識における基本的道具となっている^{39, 40)}. また, 時系列ではないが, 同じく系列パターンである DNA 塩基配列の類似性解析においても, 様々に拡張されながら多用されている^{43, 44)}. なお, 最適化法としての DP の一般的な性質については, DP の創始者 R.Bellman による文献⁴¹⁾や, 鍋島⁴²⁾を参照されたい.

(2) 原理

二つの系列パターン $X = x_1, x_2, \dots, x_i, \dots, x_I$, $Y = y_1, y_2, \dots, y_j, \dots, y_J$ 間の弾性マッチングを考える. ここで x_i と y_j は特徴ベクトルである. 一般に, 弾性マッチングとは, X の第 i 要素 x_i と Y の第 j 要素 y_j との対応づけ $j = u_i$ ($i = 1, \dots, I$) を最適化する問題である.

この要素間の対応づけのコスト(局所距離)を

$$d_i(u_i) = \|x_i - y_{u_i}\| \quad (2\cdot26)$$

とすると, 弾性マッチングを求めるための最適化問題は次のように定式化される.

$$\left. \begin{array}{ll} \text{minimize} & F = \sum_{i=1}^I d_i(u_i), \\ \text{with respect to} & u_1, \dots, u_I, \\ \text{subject to} & 0 \leq u_i - u_{i-1} \leq 2, u_1 = 1, u_I = J \end{array} \right\} \quad (2 \cdot 27)$$

第一の制約条件は単調連続性制約と呼ばれ、(i) u_i は u_{i-1} より小さくなることなく (単調性)、(ii) u_i は u_{i-1} より大きくなり過ぎることはない (連続性)、という条件である。連続性条件は対応づけ $j = u_i$ を関数 (歪み関数と呼ばれる) としてみたときの傾き du_i/di を 2 以下にするための条件であり、2 倍以上の速度変化を許さない条件となっている。第二、第三は始端・終端条件である。

この最適化問題の解法には単純な総当りを含めいくつか考えられるが、以下では DP の考えに従った解法を導く。まず問題式 (2・27) を次のように表現する。

$$\begin{aligned} \min F &= \min_{\substack{u_1, \dots, u_I \\ 0 \leq u_i - u_{i-1} \leq 2}} \sum_{i=1}^I d_i(u_i) \\ &= \min_{\substack{u_2, \dots, u_I \\ 0 \leq u_i - u_{i-1} \leq 2}} \left[\sum_{i=3}^I d_i(u_i) + \left\{ d_2(u_2) + \min_{\substack{u_1 \\ 0 \leq u_2 - u_1 \leq 2}} d_1(u_1) \right\} \right] \end{aligned} \quad (2 \cdot 28)$$

ここで、右辺 {} 内を

$$g_2(u_2) = d_2(u_2) + \min_{\substack{u_1 \\ 0 \leq u_2 - u_1 \leq 2}} d_1(u_1) \quad (2 \cdot 29)$$

と置くと、

$$\min F = \min_{\substack{u_2, \dots, u_I \\ 0 \leq u_i - u_{i-1} \leq 2}} \left[\sum_{i=3}^I d_i(u_i) + g_2(u_2) \right] \quad (2 \cdot 30)$$

となり、見かけ上、制御変数 u_1 を消去できる (見かけ上というのは、式 (2・29) で min を与える u_1 がまだ確定できていないためである)。上式を更に変形して、

$$\min F = \min_{\substack{u_3, \dots, u_I \\ 0 \leq u_i - u_{i-1} \leq 2}} \left[\sum_{i=4}^I d_i(u_i) + \left\{ d_3(u_3) + \min_{\substack{u_2 \\ 0 \leq u_3 - u_2 \leq 2}} g_2(u_2) \right\} \right] \quad (2 \cdot 31)$$

として、右辺 {} 内を

$$g_3(u_3) = d_3(u_3) + \min_{\substack{u_2 \\ 0 \leq u_3 - u_2 \leq 2}} g_2(u_2) \quad (2 \cdot 32)$$

と置くと、

$$\min F = \min_{\substack{u_3, \dots, u_I \\ 0 \leq u_i - u_{i-1} \leq 2}} \left[\sum_{i=4}^I d_i(u_i) + g_3(u_3) \right] \quad (2 \cdot 33)$$

となり，今度は制御変数 u_2 を消去できる．このように

$$g_i(u_i) = d_i(u_i) + \min_{\substack{u_{i-1} \\ 0 \leq u_i - u_{i-1} \leq 2}} g_{i-1}(u_{i-1}) \quad (2 \cdot 34)$$

を繰り返し計算することで次々に変数を消去でき，結局

$$\min F = \min_{u_I} g_I(u_I) \quad (2 \cdot 35)$$

として目的関数 F の最小値が得られる．境界条件 $u_I = J$ があるため，結局 $\min F = g_I(J)$ となる．これが前述の DP 距離であり，非線形伸縮に対して不変なマッチング距離となる．また，いわゆるバックトラック処理（後述）により，DP 距離を与える最適な対応づけ $\bar{u}_1, \dots, \bar{u}_i, \dots, \bar{u}_I$ が求まる．

式 (2・34) は DP 漸化式と呼ばれ，これが DP マッチングの基本式となっている．この DP 漸化式を， $i = 1$ から I まで順にすべての j において計算すること最適解が得られる．このことは，図 2・8(b) のような $I \times J$ 個の格子点をもつトレリス（DP 平面と呼ばれる）上のすべての点において，式 (2・34) 中の最小値選択に従った局所的な経路選択を順次行っていることに相当する．

Input:
 $\mathbf{X} = x_1, x_2, \dots, x_i, \dots, x_I, \mathbf{Y} = y_1, y_2, \dots, y_j, \dots, y_J$

Output:
 $\min F$: Minimum elastic matching distance between \mathbf{X} and \mathbf{Y} .
 $\bar{u}_1, \dots, \bar{u}_i, \dots, \bar{u}_I$: Optimal elastic matching between \mathbf{X} and \mathbf{Y} .

Initialization:
 1: $g_i(j) = \infty$ and $b_i(j) = \text{nil}$ for $\forall i, \forall j$

DP Recursion:
 2: $g_1(1) = d_1(1) = \|x_1 - y_1\|$
 3: for $i = 2$ to I do
 4: for $j = 1$ to J do
 5: $d_i(j) = \|x_i - y_j\|$
 6: $g_i(j) = d_i(j) + \min_{\max(j-2, 1) \leq j' \leq j} g_{i-1}(j')$
 7: $b_i(j) = \operatorname{argmin}_{\max(j-2, 1) \leq j' \leq j} g_{i-1}(j')$

Termination:
 8: $\min F = g_I(J)$
 9: $\bar{u}_I = J$
 10: for $i = I - 1$ downto 1 do $\bar{u}_i = b_{i+1}(\bar{u}_{i+1})$

図 2・9 DP マッチングの擬似コード

この DP 漸化式を含めたアルゴリズムの全体を擬似コードとして図 2・9 に示す．同図第 6 行が DP 漸化式に相当する．アルゴリズム自体は，このようにたかだか 10 行程度で表せるほど簡潔である．第 10 行はバックトラック処理であり，これにより DP 距離 $\min F$ を与えた最適対応づけが， $\bar{u}_I, \dots, \bar{u}_i, \dots, \bar{u}_1$ の順に求まる．DP 距離だけがが必要な場合，バックトラッ

ク処理は不要である。ループ構造からも明らかのように、このアルゴリズムは探索幅 J 、深さ I の幅優先探索であり、計算量は $O(IJ)$ となる。

(3) 高速化

DP マッチングは効率的な手法であるが、それでも問題が大規模化すれば計算量が問題になる場合がある。例えば、認識問題など、多数回 DP マッチングを繰り返す場合がこれに当たる。このため、様々な高速化法が提案されている。代表的なものとして、ビームサーチ（枝刈り）と整合窓が挙げられる。ビームサーチは、各 i において、最適経路になる見込みのない $g_i(u_i)$ について DP 漸化式を計算せずにスキップする方法である。局所探索の導入に相当し、最適解の保証はなくなるが、実用上十分な解が得られる場合が多い。見込みの有無の判定は、一般に $g_i(u_i)$ に対するしきい値処理に基づく。整合窓はマッチングの範囲を単純に制約 ($u_i \sim i$) するものであり、図 2・9 の第 4 行の j に関するループを「for $j = i - \epsilon$ to $i + \epsilon$ do」と書き直すことで実装される ($\epsilon \ll J$) は正定数)。ほかに、粗密 DP⁴⁵⁾ や解析的 DP⁴⁶⁾、また DNA 塩基配列解析用開発された FASTA や BLAST^{43, 44)} などが提案されている。

(4) マルチプルアライメント

以上では二つのパターン間のマッチングを求める場合について議論したが、DNA 塩基配列解析などにおいては、三つ以上のパターン間について同時にマッチングを求める場合がある。これはマルチプルアライメント⁴⁴⁾ と呼ばれる問題である。図 2・10 に示すように、DP の探索空間を組合せ的に拡張することで、最適解を得ることもできる（多次元動的計画法と呼ばれる）。しかし同図より明らかなように、パターン数が増えたとこの探索空間は爆発的に大きくなってしまふ。このため、ツリーベース法やセンタースター法など、2 パターン間のマッチング結果を順次拡張していくことで近似解を求める方法が提案されている⁴⁴⁾。

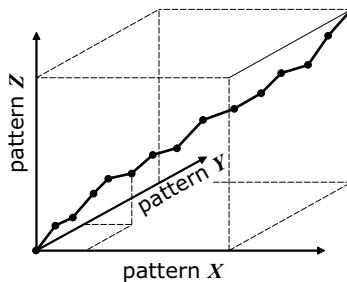


図 2・10 多次元動的計画法によるマルチプルアライメント

2-2-3 グラフマッチング

画像から得られた特徴を頂点とし、特徴間の関係を辺とするグラフをつくり、グラフ間の対応づけを行うことにより特徴集合間を照合するのが、ここでいうグラフマッチングである。無向グラフ、特に二部グラフにおいて、辺で結ばれた頂点を二つずつペアにすることもグラ

フのマッチングというが、それとは異なる。

過去 30 年間、グラフマッチングは 2 次元及び 3 次元画像解析、文字認識、物体認識、顔認識、指紋認識、インデクシングと検索、動画画像解析など、幅広い分野に応用されている¹⁴⁾。

今、二つのグラフを $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$, とする。ここで V_i は頂点の集合、 E_i は頂点の組で表した辺の集合である。例えば $u, v \in V_1$ は u, v が G_1 の頂点であることを表し、 $(u, v) \in E_1$ はその 2 頂点間に辺があることを表す。パターン認識で使うグラフには、例えば辺に正の数値をもたせて頂点で表される特徴間の類似度を表すなど、頂点や辺に何らかの属性を結び付けている場合が多い。

(1) 厳密マッチング

グラフのマッチングは何らかのかたちで二つのグラフ間に対応がつくことを意味するが、パターン認識で使われる条件にも様々な種類が存在する。まず、何らかの条件を厳密に満たすことを要求する場合を述べる。

最も完全な対応は、グラフ同型、すなわちグラフとして全く同じであるというものである。つまり、1 対 1 写像

$$f: V_1 \rightarrow V_2 \quad (2\cdot36)$$

が存在して、

$$(u, v) \in E_1 \implies (f(u), f(v)) \in E_2 \quad (2\cdot37)$$

$$(u, v) \in E_1 \longleftarrow (f(u), f(v)) \in E_2 \quad (2\cdot38)$$

が成り立つことを要求する。

これより弱い条件は部分グラフ同型で、 f が 1 対 1 であるという条件を落として、単に単射である（つまり $u \neq v$ ならば常に $f(u) \neq f(v)$ ）という条件にしたものである。この場合も式 (2·37)(2·38) の条件は要求する。これは、一方のグラフと他方のグラフの部分グラフの間にグラフ同型が成立するという意味する。

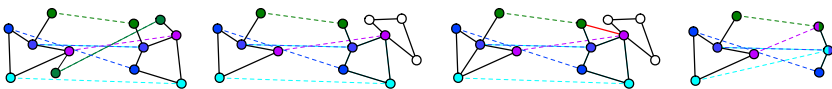


図 2-11 左からグラフ同型、部分グラフ同型、単射準同型、準同型

更に弱い条件は、単射準同型で、式 (2·38) の条件を落として式 (2·37) を満たす単射 f が存在することのみを要求する。最も弱い条件は、グラフ準同型で、 f は単射である必要もなく、式 (2·37) の条件を満たす写像 f が存在することのみを要求する。

またほかのマッチングの条件として、2 グラフ間に同型な最大の部分グラフ (maximum common subgraph, MCS) を見つけるというものがある。この問題は 2 グラフの連合グラフと呼ばれるものをつくり、その中に極大クリークを見つめる問題に帰結できることが広く知られている²⁰⁾。

以上のマッチングを見つける問題には多項式時間のアルゴリズムは知られていないが、パターン認識への応用では、グラフの種類が限られていることや頂点や辺の属性を利用して、一般で純粋なグラフアルゴリズムよりも高速化する手法が多数提案されている。

パターン認識においては、パターンそのものの多様性やノイズなどの影響で、認識したい対象からつくられるグラフに揺らぎが存在することが多い。そのため、グラフ同型が条件として使われることはほとんどない。部分グラフ同型あるいは準同型の方が、より高コストな問題であるにもかかわらずよく使われる。また MCS 発見問題も、知られている方法はどれもあまり大きなグラフでは使用不能であるにもかかわらず、注目を集めている。

厳密マッチングのためのアルゴリズムのほとんどが、何らかの木探索に基づいている。基本的には、マッチする頂点のペアを両グラフから見つけて、これを部分的なマッチングに繰り返し加えて拡大していき、行き詰った場合にはバックトラックしてやり直すというものである。このようなアルゴリズムは、頂点や辺の属性を利用してマッチングを規制することが容易であり、特にパターン認識用途に向いている。この種のアルゴリズムの最初で代表的なのは J.R. Ullman によるもの¹⁵⁾で、現在でもおそらく最もよく使われる。

(2) 非厳密マッチング

パターン認識においては、上述のような理由で、同じ対象からつくられたグラフでも一定からは程遠い。そのため、式(2.37)の条件を厳密に要求せず、満たさない場合にペナルティを課すことにし、その和であるマッチングコストを最小にするマッチングを探すアルゴリズムがよく使われる。このようなアルゴリズムは、最適解を要求すれば厳密マッチングよりも高コストだが、近似解を求める場合には厳密マッチングより高速化できる場合も多い。

マッチングコストの定義としては、具体的なエラーすなわち厳密なマッチからの違いについてそれぞれコストを定義する方法がある。これらの方法はエラー訂正 (error-correcting) 法またはエラートレラント (error-tolerant) 法と呼ばれる。また、頂点挿入や頂点削除などの一連のグラフ編集操作を定義してコストを決め、一方のグラフから他方のグラフをつくるために必要な編集の最低コスト (グラフ編集コスト) をマッチングコストとする方法もある。グラフ編集コストが距離の公理を満たすとき、グラフ編集距離と呼ぶが、H. Bunke は適当な編集操作コストを決めることで、MCS 発見問題をグラフ編集距離の計算の特別な場合と考えることができることを示した¹⁷⁾。

厳密マッチングの場合同様、木探索に基づいたアルゴリズムが多数提案されている。最も最初のもは W.H. Tsai と K.S. Fu による¹⁶⁾。

一方、離散問題であるこの問題を連続変数の非線形最適化問題に翻訳して解こうとする一連の方法もある。その一つは M. Fischler と R. Elschlager¹⁸⁾にはじまる弛緩ラベリング (relaxation labeling) 法である。この方法では、一方のグラフの各頂点に他方のグラフの頂点を表すラベルを与え、各頂点にはラベルの確率分布を維持し、マッチングコストに基づいて繰り返し確率を再計算する。

また、スペクトル法と呼ばれる方法では、グラフの隣接行列の固有値と固有ベクトルが頂点の置換によって不変であることから、これらを比較することでグラフをマッチングする。この方法の先駆的な研究には梅山伸二¹⁹⁾の方法がある。

参考文献

- 1) S.M. Smith and J.M. Brady, "SUSAN - a new approach to low level image processing," *Int'l J Comput. Vision*, vol.23, no.1, pp.45-78, 1997.
- 2) J. Shi and C. Tomasi, "Good Features to Track," *Proc. IEEE Comp. Soc. Conf. Comput. Vision and Patt. Recogn.*, pp.593-600, 1994.
- 3) C. Harris and M. Stephens, "A combined corner and edge detector," *Proc. 4th Alvey Vision Conf.*, pp.147-151, 1988.
- 4) J. Matas and O. Chum, M. Urba and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," *Proc. British Machine Vision Conf.*, pp.384-396, 2002.
- 5) D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int'l J Comput. Vision*, vol.60, no.2, pp.91-100, 2004.
- 6) Lindeberg, T., "Scale-space theory: A basic tool for analysing structures at different scales," *J Appl. Stat.*, vol.21, no.2, pp.224-270, 1994.
- 7) A. Witkin, "Scale-space filtering," *Proc. Int'l Joint Conf. on Artif. Intell.*, pp.1019-1022, 1983.
- 8) A. Grossman and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shapes," *SIAM J. Math. Anal.*, vol.15, no.4, pp.723-736, 1984.
- 9) I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Comm. Pure Appl. Math.*, vol.41, pp.909-996, 1988.
- 10) S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$," *Trans. Amer. Math. Soc.*, vol.315, pp.69-87, 1989.
- 11) S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.11, no.7, pp.674-693, 1989.
- 12) D. Gabor, "Theory of Communication," *J. IEE*, vol.93, no.26, pp.429-457, 1946.
- 13) G. Strang and T. Nguyen, "Wavelets and Filter Banks," Cambridge Univ. Press, 1996. (高橋進一, 池原雅章 訳, "ウェーブレット解析とフィルタバンク I (入門編), II (応用編)," 培風館, 1999.)
- 14) D. Conte, P. Foggia, C. Sansone and M. Vento, "Thirty Years of Graph Matching in Pattern Recognition," *Int'l J Patt. Recogn.*, vol.18, no.3, pp.265-298, 2004.
- 15) J.R. Ullman, "An algorithm for subgraph isomorphism," *J. Assoc. Comput. Mach.*, vol.23, pp.31-42, 1976.
- 16) W.H. Tsai and K.S. Fu, "Error-correcting isomorphisms of attributed relational graphs for pattern analysis," *IEEE Trans. Syst. Man Cybern.*, vol.9, pp.757-768, 1979.
- 17) H. Bunke, "On a relation between graph edit distance and maximum common subgraph," *Patt. Recogn. Lett.*, vol.18, pp.689-694, 1997.
- 18) M. Fischler and R. Elschlager, "The representation and matching of pictorial structures," *IEEE Trans. Comput.*, vol.22, pp.67-92, 1973.
- 19) S. Umeyama, "An eigendecomposition approach to weighted graph matching problems," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.10, pp.695-703, 1988.
- 20) A.P. Ambler, H.G. Barrow, C.M. Brown, R.M. Burstall and R.J. Popplestone, A versatile computer-controlled assembly system, *Proc. Int'l Joint Conf. on Artif. Intell.*, pp.298-307, 1973.
- 21) J.D. Tubbs, "A Note on Binary Template Matching," *Patt. Recogn.*, vol.22, no.4, pp.359-365, 1989.
- 22) 熊澤逸夫, "コンピュータビジョンの基礎となる対応点問題をめぐって," *映情学誌*, vol.60, no.3, pp.313-320, 2006.
- 23) S.S. Beauchemin and J.L. Barron, "The Computation of Optical Flow," *ACM Comput. Surv.*, vol.27, no.3, pp.433-467, 1995.
- 24) 井宮 淳, "動画理解の数理," *情処学 CVIM 研報*, 2006-CVIM-152, 2006.
- 25) 安田 浩, 渡辺 裕, "デジタル画像圧縮の基礎," *日経 BP 出版センター*, 1996.

- 26) D.I. Barnea and H.F. Silverman, "A Class of Algorithms for Fast Digital Image Registration," IEEE Trans. Comput., vol.C-21, no.2, pp.179-186, 1972.
- 27) 尾上守夫, 前田紀彦, 斎藤 優, "残差逐次検定法による画像の重ね合わせ," 情報処理, vol.C-21, no.2, pp.179-186, 1972.
- 28) 村瀬 洋, V.V. Vinod, "局所色情報を用いた高速物体探索 — アクティブ探索法 —," 信学論 (D), vol.J81-D-II, no.9, pp.2035-2042, 1998.
- 29) 青木孝文, 伊藤康一, 本間尚文, "位相情報に基づく画像マッチング技術とその応用展開: 3D ビジョンからバイオメトリクスまで," 信学誌, vol.90, no.8, pp.680-685, 2007.
- 30) L.G. Brown, "A Survey of Image Registration Techniques," ACM Comput. Surv., vol.24, no.4, pp.325-376, 1992. (白井良明 訳, "画像の位置合せ手法の概観," bit 別冊, pp.78-119, 1994 年 11 月号)
- 31) A.K. Jain, Y. Zhong and M. -P. Dubuisson-Jolly, "Deformable Template Models: A Review," Signal Process., vol.71, no.2, pp.109-129, 1998.
- 32) C.A. Gralbey and K.V. Mardia, "A Review of Image-Warping Methods," J Appl. Stat., vol.25, no.2, pp.155-171, 1998.
- 33) H. Lester and S.R. Arridge, "A survey of hierarchical non-linear medical image registration," Patt. Recogn., vol.32, no.1, pp.129-149, 1999.
- 34) S. Uchida and H. Sakoe, "A survey of elastic matching techniques for handwritten character recognition," IEICE Trans. Inf. & Syst., vol.E88-D, no.8, pp.1781-1790, 2005.
- 35) M.A. Fischler and R.C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM, vol.24, no.6, pp.381-395, 1981.
- 36) 迫江博昭, 千葉成美, "動的計画法を利用した音声の時間正規化に基づく連続音声認識," 音響誌, vol.27, no.9, pp.483-490, 1971.
- 37) H. Sakoe and S. Chiba, "A Dynamic Programming Algorithm Optimization for Spoken Word Recognition," IEEE Trans. Acoust. Speech and Signal Proc., vol.ASSP-26, no.1, pp.43-49, 1978.
- 38) H. Sakoe, "Two-Level DP-Matching Algorithm—A Dynamic Programming Based Pattern Matching Algorithm for Continuous Speech Recognition," IEEE Trans. Acoust. Speech and Signal Proc., vol.ASSP-27, no.6, pp.588-595, 1979.
- 39) 大田友一 and 山田博三, "動的計画法によるパターンマッチング," 情報処理, vol.30, no.9, pp.1058-1066, 1989.
- 40) 内田誠一, "DP マッチング概説," 信学技報, vol.PRMU2006-166, 2006.
- 41) R. Bellman and S. Dreyfus, "Applied Dynamic Programming," Princeton Univ. Press, 1962.
- 42) 鍋島一郎, "動的計画法," 森北出版, 1968. (POD 版, 2005)
- 43) R. Durbin, S. Eddy, A. Krogh and G. Mitchison, "Biological Sequence Analysis," Cambridge Univ. Press, 1998.
- 44) 阿久津達也, "バイオインフォマティクスの数理とアルゴリズム," 共立出版, 2007.
- 45) C. Raphael, "Coarse-to-fine dynamic programming," IEEE Trans. Patt. Anal. Mach. Intell., vol.23, no.2, pp.1379-1390, 2001.
- 46) 内田誠一, 迫江博昭, "解析的 DP マッチング," 信学論 (D), vol.J90-D, no.8, pp.2137-2146, 2007.