

クラスタリングによる視点に不変なパターン認識の 学習

井上, 光平
九州芸術工科大学

<https://doi.org/10.11501/3168354>

出版情報：九州芸術工科大学，1999，博士（工学），課程博士
バージョン：
権利関係：

第5章

グラフスペクトル法による逐次ファジークラスタ抽出

重み付きグラフによって表される点データの集合からファジークラスタを逐次抽出する方法を提案する [50]. グラフの接点と枝は非負実数の重みをもつ. データのクラスタへのメンバシップはグラフの重み付き近接行列の第1固有ベクトルにより与えられる. クラスタ抽出後, 接点の重みを順次減少することにより抽出済みのクラスタを取り除きながら逐次にクラスタを抽出していく. 抽出されたクラスタの大きさが急な増大を示すときクラスタの抽出は終了する. 本手法を画像のセグメンテーションとカラー画像からの肌色領域の抽出に適用する.

5.1 まえがき

ファジー c -平均法 [7] や確率 c -平均法 [51] のような反復アルゴリズムによるファジークラスタリングには収束の問題があり, ノイズデータに敏感である. これら固定点反復アルゴリズムの欠点を回避する方法としてマウンテン法 [52] やサブトラクティブクラスタリング [53] のような逐次抽出法がある. また別の方法としてグラフスペクトル法がある. この方法はデータの重み付き近接行列の固有値問題を解くことによりクラスタへのメンバシップを求める [54]. すなわちグラフスペクトル法ではクラスタへのメンバシップはデータの近接行列の第1固有ベクトルにより与えられる [55]. 津田ら [56] は直線を逐次に抽出するグラフスペクトル法を提案した. その方法ではマウンテン法と同様に, 次に抽出するクラスタがすでに抽出された全てのクラスタから離れるように目的関数に順次ペナルティ項が加えられる. ペナルティ項の係数の適当な値は問題

に依存し、その決定は困難なことが多い。本章では従来法 [56] のようにある関数を順次減算していく代わりに目的関数に重み関数を順次乗じる方法を提案する。提案法では従来法に見られるペナルティ項の係数などのパラメータの調整の必要がない。画像のセグメンテーションとカラー画像からの肌色領域の抽出の例により本方法の有効性を調べる。

5.2 重み付きグラフの固有クラスタ

グラフスペクトル法ではデータの集合は無向グラフで表される。データは接点で表され、接点間の枝はデータ間の近接度に対応する重みをもつ。従来法の多く [56] では接点の重みは考慮されていない、すなわち接点の重みは全て等しく 1 に固定されている。本章では [55] と同様にどの接点も非負実数の重みをもつ。この接点の重みの導入は次節の逐次抽出法の基礎となる。

n 個のデータ d_i ($i = 1, \dots, n$) は n 個の接点をもつグラフで表される。 i 番目のデータと j 番目のデータの間の近接度は i 番目の接点と j 番目の接点の間の枝の重み h_{ij} で表される。このグラフの接点の最も大きなクラスタを見つけることを考える。提案法の出発点は接点の重みが全て 1 に等しい場合である。 x を n 次元ベクトル $x = [x_1, \dots, x_n]$ とする。 x_i は i 番目の接点がクラスタに含まれる割合を表す。各要素 x_i はそのままではメンバシップではない。後述するようにメンバシップは規格化された x_i である。 x のノルムは 1 すなわち $\|x\|^2 = x^T x = 1$ に制約される。従来法 [55, 56] に従い、接点クラスタの凝集度を $\sum_i^n \sum_j^n h_{ij} x_i x_j = x^T H x$ で測ることにする。 H はグラフの枝の重み h_{ij} の重み付き近接行列 $H = [h_{ij}]$ である。

凝集度が最大のクラスタのベクトル x は次の最適化問題の解である：

$$\begin{aligned} \max_x \quad & \sum_i^n \sum_j^n h_{ij} x_i x_j \\ \text{subj.to} \quad & \sum_i^n x_i^2 = 1 \end{aligned} \tag{5.1}$$

解はラグランジュ乗数法により直ちに得られ、解 x は行列 H の最大固有値に対応する固有ベクトルである。また最大固有値はクラスタの凝集度を与える。

固有ベクトル x はグラフの固有クラスタ (eigencluster) と呼ばれる [55]。 x の要素は全て非負である。最大の要素 x_i をもつ接点はクラスタの代表的メンバである。そこで

最大要素の変数を i^* とし: $i^* = \arg \max_i x_i$, x_i を x_{i^*} により $\hat{x}_i = x_i/x_{i^*}$ と規格化する. これは i 番目のデータのクラスタへのメンバシップを表す. このようにすると代表データ i^* が最大のメンバシップ値 1 をもつようになる.

次に各接点が実非負の重みをもつ場合を考え, 上の表現がどのように変わるかを調べる. まず $(n-1)$ 番目のデータが n 番目のデータに一致するすなわち $d_{n-1} = d_n$ となるような直観的な場合を考える. これはデータ数が $n-1$ でデータ d_{n-1} の重みが 2 である場合と等価である. $d_{n-1} = d_n$ であるから明らかに $x_{n-1} = x_n$ であり, 従って式 (5.1) は

$$\begin{aligned} \max_x \quad & \sum_i^{n-1} \sum_j^{n-1} w_i w_j h_{ij} x_i x_j \\ \text{subj.to} \quad & \sum_i^{n-1} w_i x_i^2 = 1 \end{aligned} \quad (5.2)$$

となる. ここで $w_i = 1$ ($i = 1, \dots, n-2$), $w_{n-1} = 2$ である. この導出を更に一般化し, 任意の実非負の重み w_i に対してこの式を用いる. x を $y_i = \sqrt{w_i} x_i$ とスケーリングすると式 (5.2) は式 (5.1) の形に戻り

$$\begin{aligned} \max_x \quad & \sum_i^{n-1} \sum_j^{n-1} \sqrt{w_i} \sqrt{w_j} h_{ij} y_i y_j \\ \text{subj.to} \quad & \sum_i^{n-1} y_i^2 = 1 \end{aligned} \quad (5.3)$$

となる. 従って接点の重みが w_i , 枝の重みが h_{ij} である重み付きグラフの固有クラスタは行列 $\tilde{H} = [\tilde{h}_{ij}]$; $\tilde{h}_{ij} = \sqrt{w_i} \sqrt{w_j} h_{ij}$ の第 1 固有ベクトルにより与えられる. x_i/x_{i^*} は y_i/y_{i^*} に等しいので接点のクラスタへのメンバシップは行列 \tilde{H} の規格化された固有ベクトルにより与えられる. ここで $i^* = \arg \max_i x_i = \arg \max_i y_i$ である.

5.3 ファジークラスタの逐次抽出

上の議論を基礎として, 接点の重み w_i , 枝の重み h_{ij} をもつグラフから 1 個ずつクラスタを抽出する逐次抽出法を示す. 最も凝集度の高いクラスタの抽出はすでに前節で述べた通りである. 凝集度が最大の固有クラスタは行列 $H_1 = [h_{1ij}]$; $h_{1ij} = \sqrt{w_i} \sqrt{w_j} h_{ij}$ の第 1 固有ベクトル y_1 により与えられる. 接点のクラスタへのメンバシップは規格化された固有クラスタ $m_{1i} = y_{1i}/y_{1i^*}$ により与えられる. ここで $i^* = \arg \max_i y_{1i}$ である.

次に凝集度の高い固有クラスタは、第1クラスタに含まれる全データを取り除いた後に残ったデータ内のクラスタの中で凝集度が最大のクラスタである。各データが第1クラスタに属する割合は m_{1i} であるから、第2クラスタの抽出では $1 - m_{1i}$ を乗じることにより接点の重みを減らすのが自然である。従って次に凝集度の高い固有クラスタは行列 $H_2 = [h_{2ij}]$; $h_{2ij} = \sqrt{1 - m_{1i}} \sqrt{1 - m_{1j}} h_{1ij}$ の第1固有ベクトル y_2 により与えられる。接点のこのクラスタへのメンバシップは規格化された固有クラスタ $m_{2i} = y_{2i} / y_{2i^*}$ により与えられる。ここで $i^* = \arg \max_i y_{2i}$ である。

その後同様にしてクラスタを順次抽出していく。(k+1)回めの行列は再帰的に $h_{k+1,ij} = \sqrt{1 - m_{ki}} \sqrt{1 - m_{kj}} h_{kij}$ により計算される。この行列の規格化された第1固有ベクトル $m_{k+1,i} = y_{k+1,i} / y_{k+1,i^*}$ はデータの第(k+1)クラスタへのメンバシップを与える。次にこの抽出処理を終了するための基準を考える。妥当なクラスタはクラスタの代表的メンバで最大となるつり鐘状のメンバシップをもつ。妥当なクラスタが全て抽出された後はノイズデータだけが残る、それらは広い範囲に点在する。そのようなノイズデータの固有クラスタは平坦で明瞭なピークをもたない。そこで順次抽出されるクラスタの大きさの変化を考えよう。クラスタの大きさはクラスタに含まれるデータのメンバシップの総和である。すなわち $\sum_i w_i m_{ki}$ が第kクラスタの大きさである。この大きさは妥当なクラスタが抽出される間は通常単調に減少する。しかし妥当なクラスタが全て抽出された後では、ノイズデータの固有クラスタは平坦になり、従ってその規格化されたメンバシップは多くのデータで大きい(1に近い)値となり、その総和は大きな値となるためこの大きさは急に増加する。このようにクラスタの大きさが急な跳ね返りを示すとき、妥当なクラスタは残っていないことになり、従ってそこで抽出は終了すべきである。

i 番目のデータが第kクラスタに属する度合であるメンバシップ m_{ki} の最終的に得られる集合は与えられたデータのファジークラスタを構成する。得られたメンバシップを非ファジー化しクリस्पなクラスタを得るには、まず全クラスタkに対してメンバシップ m_{ki} が予め決められたしきい値 ϵ よりも小さいデータを捨てる。そのようなデータはノイズか外れ値である。次に残ったデータについては、各データを最大のメンバシップ m_{k^*i} をもつクラスタ k^* に所属させればよい。

5.4 実験

5.4.1 提案法の検証

図5.1は630個の2次元データを示したものである。クラスタのデータ数は120, 100, 75, 45, 40(これらの場所は図5.4を参照)であり, 残りの250個はノイズデータである。この例ではデータ d_i はデータ点の2次元座標である。各クラスタは正規分布しノイズは一様分布である。枝の重みは $h_{ij} = e^{-k\|d_i - d_j\|^2}$ とし, スケール係数は $k = 100$ とする。接点の重み w_i は全データに対して等しく1とする。図5.2に第1クラスタのメンバーシップ値(左)と第5クラスタのそれ(右)を示す。他のクラスタの形状も同様であった。抽出されたクラスタの大きさを図5.3に示す。第5クラスタの抽出後, 急な跳ね返りが見られる。従ってそこで抽出は終了する。棄却しきい値 $\epsilon = 0.5$ で非ファジー化して得られたクリスピークラスタを図5.4に示す。付随している番号は抽出された順番を示す。

津田ら [56]の方法によっても η の値を 10^2 から 10^6 の間に設定すると同様な結果が得られた。この範囲はかなり広いが, この範囲外の η の値では良い結果は得られなかった。またファジー c -平均法は [56]でも指摘されているようにノイズに敏感で初期値により結果が変わり, 提案法よりも収束に時間がかかる。

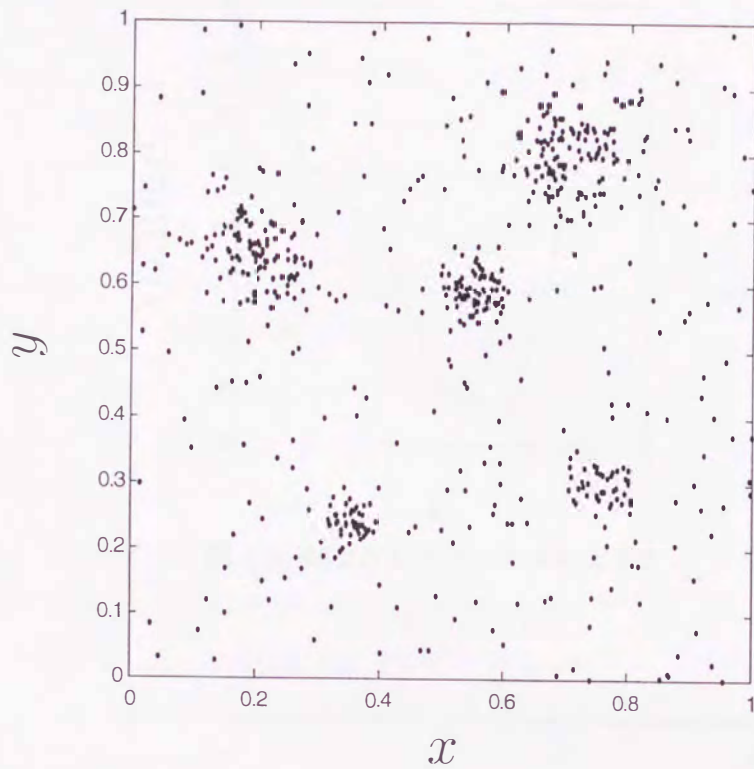


図 5.1: データの例

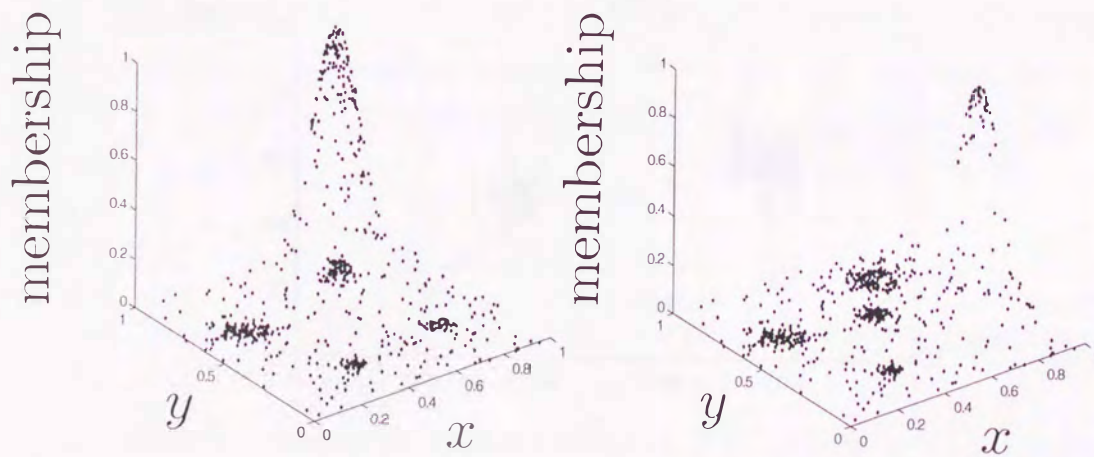


図 5.2: 第1(左), 第5(右)クラスターのメンバシップ

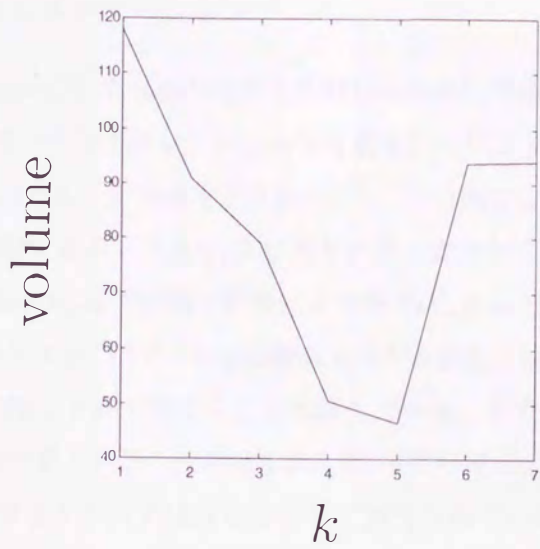


図 5.3: 抽出されたクラスタの大きさ

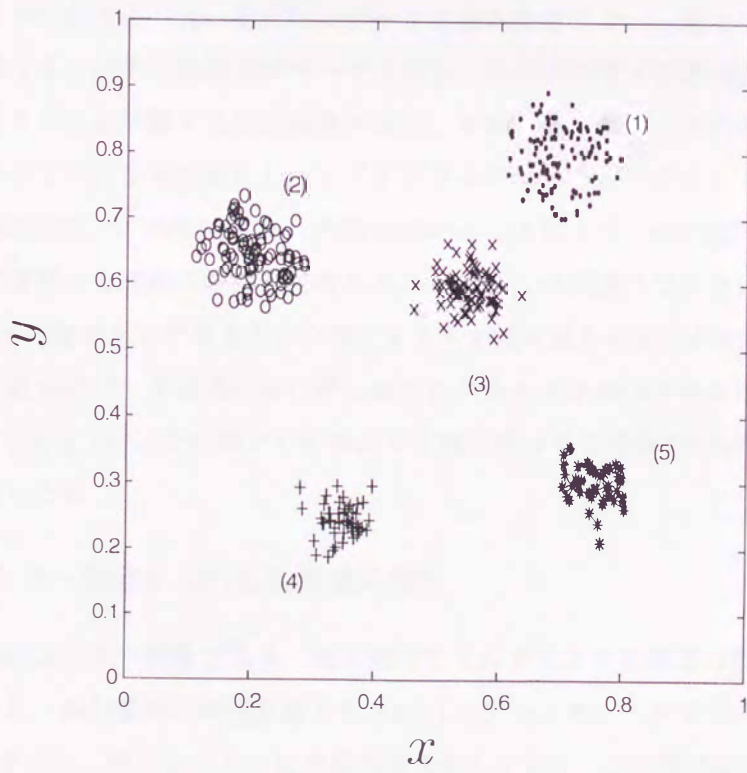


図 5.4: 非ファジー化後の抽出されたクラスタ

5.4.2 画像のセグメンテーション

次の例は図 5.5(a) に示す 166x171 サイズの Landsat 航空画像のセグメンテーションである (5つの領域へのセグメンテーションを図 5.5(b) に示す. 図はコントラストを強調している). そのヒストグラムを図 5.6 に示す. この例ではデータ d_i はビンの番号すなわち図 5.6 の横軸である. 重み w_i はヒストグラムカウントすなわちそのビンの画素数である. 枝の重み h_{ij} は上の例と同様に $e^{-k\|d_i-d_j\|^2}$, $k = 100$ である. 抽出されたクラスタの大きさを図 5.7 に示す. 大きな跳ね返りが 5 番目の抽出に見られる. この結果は第 5 クラスタで抽出を終了することを示唆している. すなわち図 5.5(a) の画像におけるセグメント数の最も顕著な候補は 5 である. 5 回の抽出後に残ったヒストグラムを図 5.8 に示す. 主要なクラスタはほとんど全て抽出されている. 図 5.7 で第 2 クラスタから第 3 クラスタにかけて大きさが増加している. これはクラスタ数のもう 1 つの候補を示している. 従って 2 がセグメント数の第 2 候補である. しかし図 5.7 から分かるように 5 回目の抽出がなされるまでは抽出が不十分である.

津田ら [56] の方法の元の定式化はデータの重みを含まない. 従ってそれをこの例に直接適用すると画像の各画素がデータとなり, 行列 H のサイズが 28386x28386 となって固有ベクトルを計算するのが困難である. それで我々はデータの重みを導入して彼らのアルゴリズムを再定式化し, ヒストグラムのビンをデータとしてこの例に適用した (付録 E 参照). この拡張された表現を用いることにより, η の値が 4×10^4 から 10^5 の範囲内で提案法と同様の結果が得られた. しかしこの範囲外では良い結果は得られなかった. η の値が小さすぎると 1 つのクラスタが繰り返し選択された. 逆に η が大きすぎると代表点がデータ領域の外に押し出されクラスタが抽出されなかった. 適当な範囲 $4 \times 10^4 \leq \eta \leq 10^5$ は上の例でのそれよりも狭く従ってこの例での η の値の設定は困難であるといえる.

5.4.3 カラー画像からの肌色領域の抽出

図 5.9(a) はカラー画像であり, 肌色抽出フィルタにより各画素の肌色度が予め与えられている. 各画素のこの肌色度を図 5.9(b) にグレイスケールで示す. 白のレベルは肌色度 0 を示し, 黒のレベルは肌色度の最大値 1 を示す. この例ではデータ d_i は i 番目の画素の 2 次元座標であり, 接点の重み w_i はその画素の肌色度である. 枝の重みは上の例と同様にスケール係数 $k = 100$ で $h_{ij} = e^{-k\|d_i-d_j\|^2}$ である. 抽出された肌色領域



(a) 原画像

(b) セグメンテーションされた画像

図 5.5: Landsat 航空画像

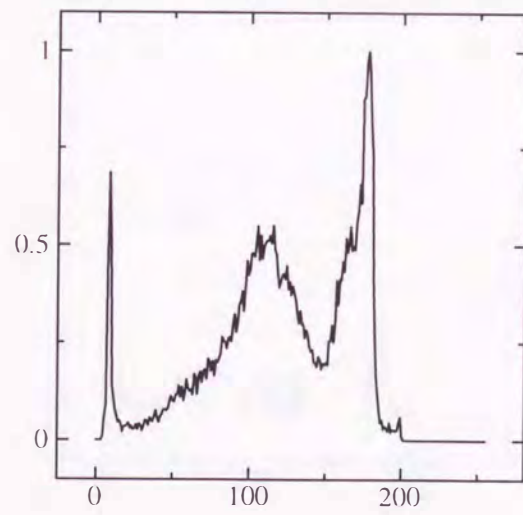


図 5.6: 図 5.5(a) の画像のヒストグラム

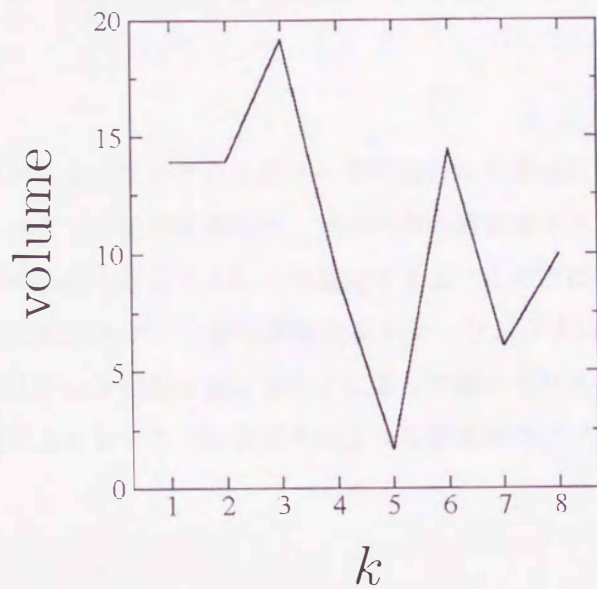


図 5.7: 抽出されたクラスタの大きさ

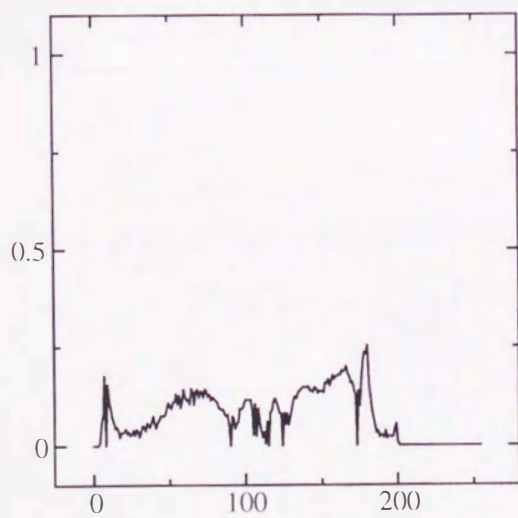


図 5.8: 5回の抽出後のヒストグラム

を図5.9(c)に示す。付随している番号は抽出された順番である。

5.5 むすび

重み付きグラフによって表される点データの集合から逐次にファジークラスタを抽出するグラフスペクトル法を提案した。データ内の適切なクラスタ数は抽出されたクラスタの大きさの変化を見ることにより推定できる。この方法で調整の必要なパラメータは枝の重みにおけるデータ間の近接度をスケールリングするパラメータ t の1つだけである。このスケール係数を変えることによって抽出されるクラスタ数を調整できる。この可変性により木ベクトル量子化のような階層的なファジークラスタ抽出が可能である。

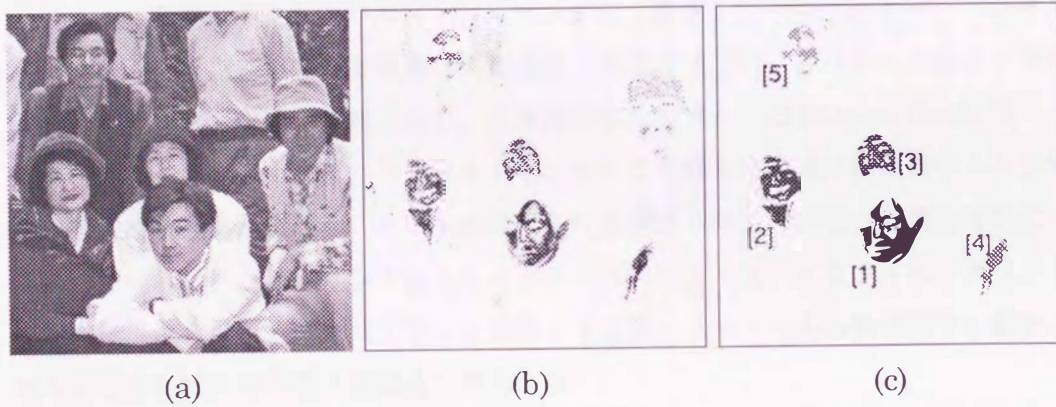


図 5.9: カラー画像からの肌色領域の抽出

第6章

点パターンマッチングに基づく平面物体の視点に不変な認識

アフィン変換に不変な点パターンマッチング法を提案し，その点パターンマッチング法に基づく視点に不変な平面物体の認識法を提案する [57]．各パターンはまず前処理によりアフィン変換を正規化され，正規化されたパターンは Shapiro-Brady 法によりマッチングされる．学習ではクラスタリングにより様々な視点の画像の中から各物体の代表的な視点が選ばれ，テスト画像は各代表画像との距離に基づく最近傍則により識別される．クラスタリング法としてグラフスペクトル法によるファジークラスタの逐次抽出法を用いる．またアフィン変換よりも狭いクラスである相似変換に制約される画像群に対する同様の認識法を提示する．

6.1 まえがき

特徴のマッチングに基づく視点に不変な物体認識を考える．特徴として特に物体のコーナーやサンプリングされた輪郭線などの点を考える．多くの方法は代数的不変量 [58, 59] を用い，ほとんどの方法で点の間の対応は既知であると仮定されている [60]．本章では点の対応は未知であり，すなわちマッチングは2つのパターンの点の間の対応を求める問題である．

2次元画像から3次元物体を認識するとき，特徴点の遮蔽は煩雑な問題を生じる．ここではそのような困難を避けるために物体のクラスを平面物体に限定する．視点に不変に物体を認識するためには透視射影に不変な点パターンマッチングが必要となる．視点の変化が小さいときには平面物体の透視射影は弱透視射影と呼ばれるアフィン変

換で近似される [33]. 従って各物体についていくつかの代表的な視点を用意することにより任意の視点の画像とそれに最も近い代表画像との関係をアフィン変換で近似できる. そうするとアフィン不変な点パターンマッチングにより測定される最近傍代表点との距離に基づいて任意の視点の画像を分類できる.

透視射影と弱透視射影については付録Fを参照されたい.

本認識システムの基礎となる処理は点パターンのアフィン不変マッチングである. これまでにグラフ理論的方法, 可変形テンプレート, 緩和法, ニューラルネットワークなど多くの点パターンマッチング法が提案されている [61]. これらの方法の多くは画像の局所的な処理に基づいているため局所最適解に捕まる危険がある. Scott と Longuet-Higgins[62] は常に大域最適解を与える固有値分解法を提案した. しかし彼らの方法はパターンの回転に弱いので, Shapiro と Brady[34] は改良法を提案し, 物体形状に基づく画像検索に応用した [63]. Sclaroff と Pentland[64] も同様の固有形状法を開発し物体認識に応用した. しかしこれらの方法はアフィン不変ではない. Shapiro-Brady 法は回転不変であるが歪み変形に対しては不安定である. Sclaroff-Pentland 法は極座標を用いることにより回転不変にできるがそれによって他の問題が生じる.

本章ではアフィン不変点パターンマッチングの2段階固有値分解法を提案する. 2次元パターンはまず1段めの固有値分解により別の2次元パターンに変換される. この変換を通して平行移動とスケール係数が較正される. 変換された2次元パターンは次に Shapiro-Brady 分解により高次元空間に写像される. Shapiro-Brady 法は回転不変であるから全体の変換はアフィン不変となる. 点パターンのマッチングはこの高次元空間で行われる. パターン間の距離はこのマッチングに基づいて評価される.

物体の認識は最近傍プロトタイプ法に基づいて行う. まず様々な視点の画像からなる学習データ集合のクラスタリングによりいくつかの代表的な視点からの画像がそれぞれの物体について選択される. このクラスタリングもパターン間の距離に基づく近接行列の固有値分解により行われる. クラスタは次々に固有値問題を解くことにより逐次に抽出される.

6.2 点パターンのアフィン不変マッチング

点パターンが画像平面上の i 番めの点の2次元位置 p_i の集合 $\{p_i\}, i = 1, \dots, n$ として与えられているとする. 本節での問題は点パターン $\{p_i\}, i = 1, \dots, n$ と $\{q_i\}, i = 1, \dots, n$

の2枚の画像の間で対応を決定することである。ここでは点パターンのアフィン不変マッチングのための2段階固有値分解法を提案する。第1段階は前処理であり、ここで元の2次元パターンは別の2次元パターンに変換される。第2段階はShapiro-Brady法であり、ここでは2次元パターンは高次元空間に写像され、そこで点の間の対応が求められる。

6.2.1 第1段階

まず平行移動 $\bar{p}_i = p_i - \sum p_i/n$ により各パターンの中心を原点に移動する。次に \bar{p}_i の内積を要素とする行列 $C = [c_{ij}]$; $c_{ij} = \bar{p}_i^T \bar{p}_j$ を構成する。付録Gに示されるようにこの行列 C のランクは2である。この行列の2個の固有ベクトル u_1, u_2 を計算する。各点はこれらの固有ベクトルにより与えられる新しい位置に写像される。これらの固有ベクトルの第 i 要素からなる2次元ベクトル $[u_{1i}, u_{2i}]$ が i 番目の点の新しい位置である。付録Gに示されるようにアフィン変換のスケール係数はこの変換を通して揃えられる。各固有ベクトルの符号には任意性があるが、後述する次の段階の結果はこの符号の取り方に不変であるためこの符号は自由に設定してよい。

図6.1の上の3枚の画像は平面物体の射影像のパターン例である。これら3個のパターンは下の3個のパターンに変換される。これらのパターンは回転によりほぼきれいに重ね合わせることができる。残った歪みは透視変換とアフィン変換の差である。

6.2.2 第2段階

次に各点はShapiro-Brady法により高次元空間に写像される。まず i 番目の点と j 番目の点の間の近接度を $e_{ij} = e^{-\alpha d_{ij}}$ で評価する。 α は正定数であり、 d_{ij} は2点間のユークリッド距離の2乗 $d_{ij} = (u_{1i} - u_{1j})^2 + (u_{2i} - u_{2j})^2$ である。次に近接行列 $E = [e_{ij}]$ ($e_{ii} = 0$)の m 個の主固有ベクトル $x_k = [x_{k1}, \dots, x_{kn}]$ ($k = 1, \dots, m$)を計算する。通常 m は n より小さく、例えば $m = n/2$ とする。 i 番目の点は m 次元空間の位置 $\tilde{x}_i = [x_{1i}, \dots, x_{mi}]$ に写像される。2つのパターン間の点のマッチングはこの高次元空間で行われる。パターン $\{p_i\}$ が $\{\tilde{x}_i\}$ に写像され、別のパターン $\{q_i\}$ が $\{\tilde{y}_i\}$ に写像されるとする。これら写像後の座標値 $\{\tilde{x}_i\}, \{\tilde{y}_i\}$ に基づいてミスマッチ行列 $Z = [z_{ij}]$; $z_{ij} = \|\tilde{x}_i - \tilde{y}_j\|$ を構成する。マッチングの決定は次のようにする。まずミスマッチ行列 Z の各行の最小要素を探し、各列でも最小値を探す。要素 z_{ij} が行と列の両方で最小となるとき、第1パターンの i 番目の点と第2パターンの j 番目の点とがマッチングするとし、それ以外ではマッ

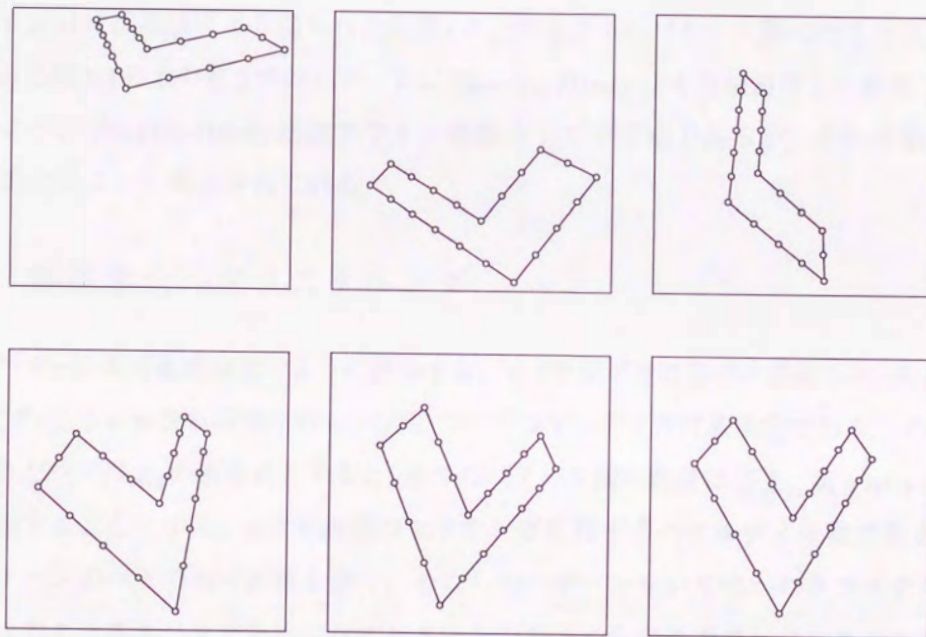


図 6.1: アフィン正規化変換

グは決定不能とする. ここでも固有ベクトル x_k の符号に任意性があることに注意する. ここでは一方のパターンを参照パターンとしてその固有ベクトル x_k の符号を固定し, 他方のパターンの固有ベクトル y_k の符号を次のようにして決めた. まず第1固有ベクトルは要素が全て同符号であるため要素の符号が揃うようにすればよい. 第 $k (\geq 2)$ 固有ベクトルについては第 $k-1$ 固有ベクトルまで符号が求まっているとすると参照パターンの座標値を $\{\tilde{x}_i^k\}$; $\tilde{x}_i^k = [x_{1i}, \dots, x_{ki}]$ として他方のパターンの第 k 固有ベクトルの符号を反転した2つの座標値 $\{\tilde{y}_i^{k+}\}$; $\tilde{y}_i^{k+} = [y_{1i}, \dots, y_{ki}]$ と $\{\tilde{y}_i^{k-}\}$; $\tilde{y}_i^{k-} = [y_{1i}, \dots, -y_{ki}]$ から mismatch 行列 Z_{k+} と Z_{k-} を構成して次節で定義されるパターン間の距離を測り, その距離が小さくなるほうの符号を選択する. このようにして第1から第 m 固有ベクトルまで順に符号を決定していく.

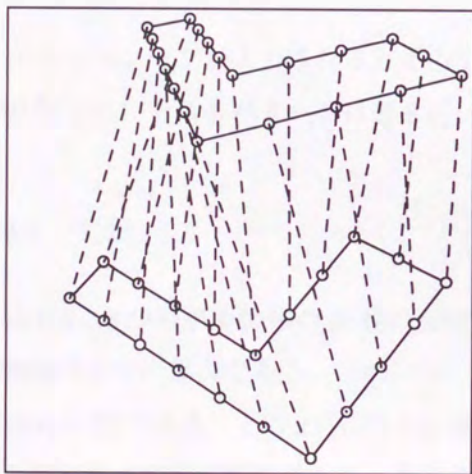
図 6.2(a) に提案法により得られた図 6.1 の上の左2つのパターン間の対応を示す. 図 6.2(b) は図 6.1 の上の左2つのパターンに Shapiro-Brady 法を直接適用した結果である. このように Shapiro-Brady 法はアフィン変換の下で不安定であるが, それが第1段階の正規化によって解消されている.

6.3 画像集合のクラスタリング

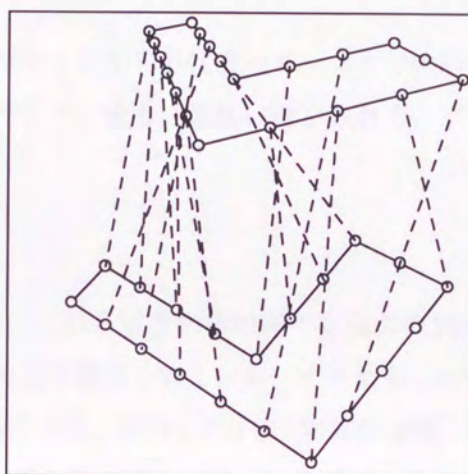
点パターン間の距離は次のように評価する. マッチングされるペアの見つかった点の数を l とする. $l < n$ ならば残りの $n-l$ 点についてはマッチングは未決定である. マッチングされたペアの z_{ij} の値を d_i とすると, 2つのパターン間の距離は $[\sum_{i=1}^l d_i + \alpha(n-l)]/n$ で評価することにする. α は未決定のマッチングに対するペナルティ係数である.

パターンのペアの間の距離を測り, それらのパターンをいくつかのクラスタに分ける. このクラスタリングを次に示すパターンの集合からの逐次ファジークラスタ抽出により行う (詳しくは参考文献 [50, 65] を参照されたい).

パターンの認識は最近傍プロトタイプ法で行う. そこで各物体についていろいろな視点のパターンをクラスタリングして代表パターンを求める. クラスタリングはグラフスペクトル法に基づく逐次ファジークラスタ抽出により行う. 学習パターンの数を N , I 番目のパターンと J 番目のパターンとの距離を D_{IJ} とする. 6.2.2 節と同様に近接行列を $H = [h_{IJ}]$; $h_{II} = 0, h_{IJ} = e^{-bD_{IJ}}$ として構成する. b は正定数である. H の最大固有値に対応する固有ベクトルすなわち H の第1固有ベクトル s を求める. 第1クラスタの代表パターンは s の要素 s_I ($I = 1, \dots, N$) の中で最大値 s_{I^*} を取る I^* である. 各



(a)



(b)

図 6.2: 提案法 (a) と Shapiro-Brady 法 (b) による対応

パターンの第1クラスタへのメンバシップは $t_{1I} = s_I/s_I$ である。第2クラスタは行列 $H_2 = [h_{2IJ}]$; $h_{2IJ} = \sqrt{1-t_{1I}}\sqrt{1-t_{1J}}h_{IJ}$ の第1固有ベクトル s_2 を計算することにより抽出される。第2クラスタの代表パターンは最大の s_{2I} をもつ I である。同様に第 k クラスタは行列 $H_k = [h_{kIJ}]$; $h_{kIJ} = \prod_{l=1}^{k-1} \sqrt{1-t_{lI}}\sqrt{1-t_{lJ}}h_{IJ}$ の第1固有ベクトルを計算することにより与えられる。このようにして抽出された代表パターンを次節の物体認識で用いる。

6.4 平面物体の視点に不変な認識

テストパターンが入力されるとそのテストパターンから各代表パターンまでの距離が計算され、テストパターンは最も近い代表パターンをもつ物体に識別される。

6.4.1 実験

図6.3に3つの平面物体の30枚の画像を示す。これらは各物体の様々な視点の10枚の画像の3つの集合である。どのパターンも20点で構成されている。すなわちこの例では $n = 20$ である。上の2行は第1物体の画像であり、中の2行は第2物体の画像、下の2行は第3物体の画像である。各物体に対して3個の代表パターンを選ぶことにする。代表として選ばれたパターンは第1物体について5番め、6番め、9番めのパターン、第2物体について6番め、7番め、9番めのパターン、第3物体について5番め、6番め、9番めのパターンである。パラメータ値は $m = n/2 = 10$, $a = 15$, $b = 15$, $\alpha = 2$ とした。図6.4(a)は第1物体の10個のパターンから9個の代表パターンまでの距離である。横軸は代表パターンの番号である(最初の3個は第1物体の代表パターン、中の3個は第2物体の代表パターン、最後の3個は第3物体の代表パターンである)。縦軸は距離である。各画像のデータは線で結ばれている。従ってこのグラフには10本の線がプロットされている。10個全ての画像について左の3個の代表パターンのうちの1つとの距離が最小になっている。従って最近傍識別則によりどの画像も第1物体として正しく識別される。同様に図6.4(b)は第2物体の10個のパターンから9個の代表パターンまでの距離を示す。また図6.4(c)は第3物体についての距離である。これらの結果はどれも正しい識別を示す。

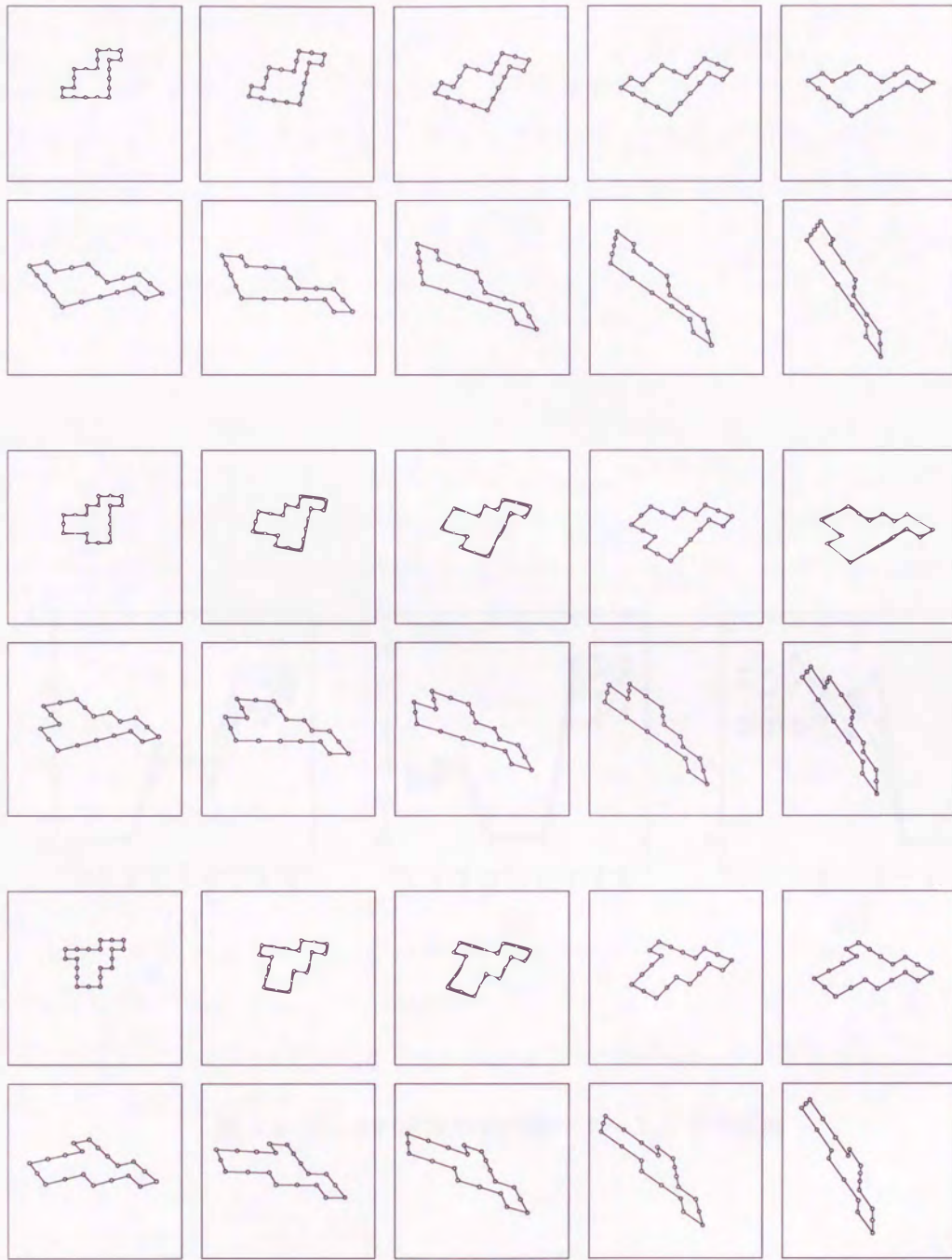
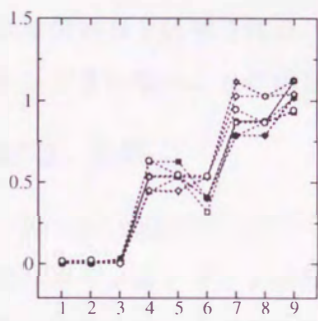
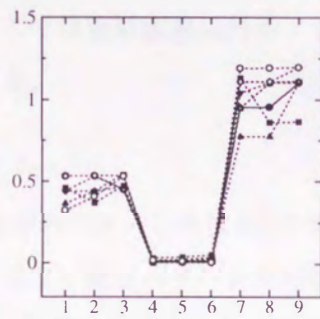


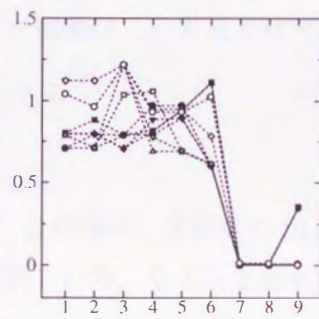
図 6.3: 3つの平面物体の様々な視点からの画像



(a)



(b)



(c)

図 6.4: 図 6.3 の画像から代表パターンまでの距離

6.5 相似不変マッチングによる認識

前節では任意の視点からの画像を考えたが、ここでは物体の向きが固定されているような状況を考える。視線方向からの物体平面の向きが一定ならばその画像は相似変換で関係付けられる。この場合、一般のアフィン変換で関係付けられる画像は異なる物体を表す。従って上述のアフィン不変法はこのような状況では物体を正しく識別できない。この場合は相似変換だけに不変なマッチングが必要である。

6.5.1 相似不変マッチング

まずアフィン変換の場合と同様に各パターンを中心を $\bar{p}_i = p_i - \sum p_i/n$ により原点に平行移動する。次にこれらの座標値を $p'_i = \bar{p}_i / \sqrt{\sum \bar{p}_i^T \bar{p}_i}$ と正規化する。これが点パターンの相似不変マッチングの第1段階である。第2段階は Shapiro-Brady 法である。ここでは正規化されたパターン $\{p'_i\}$ と $\{q'_i\}$ がマッチングされる。この方法の相似不変性は付録 H で証明される。そこでは相似変換以外のアフィン変換はこの方法ではマッチングされないことが示される。

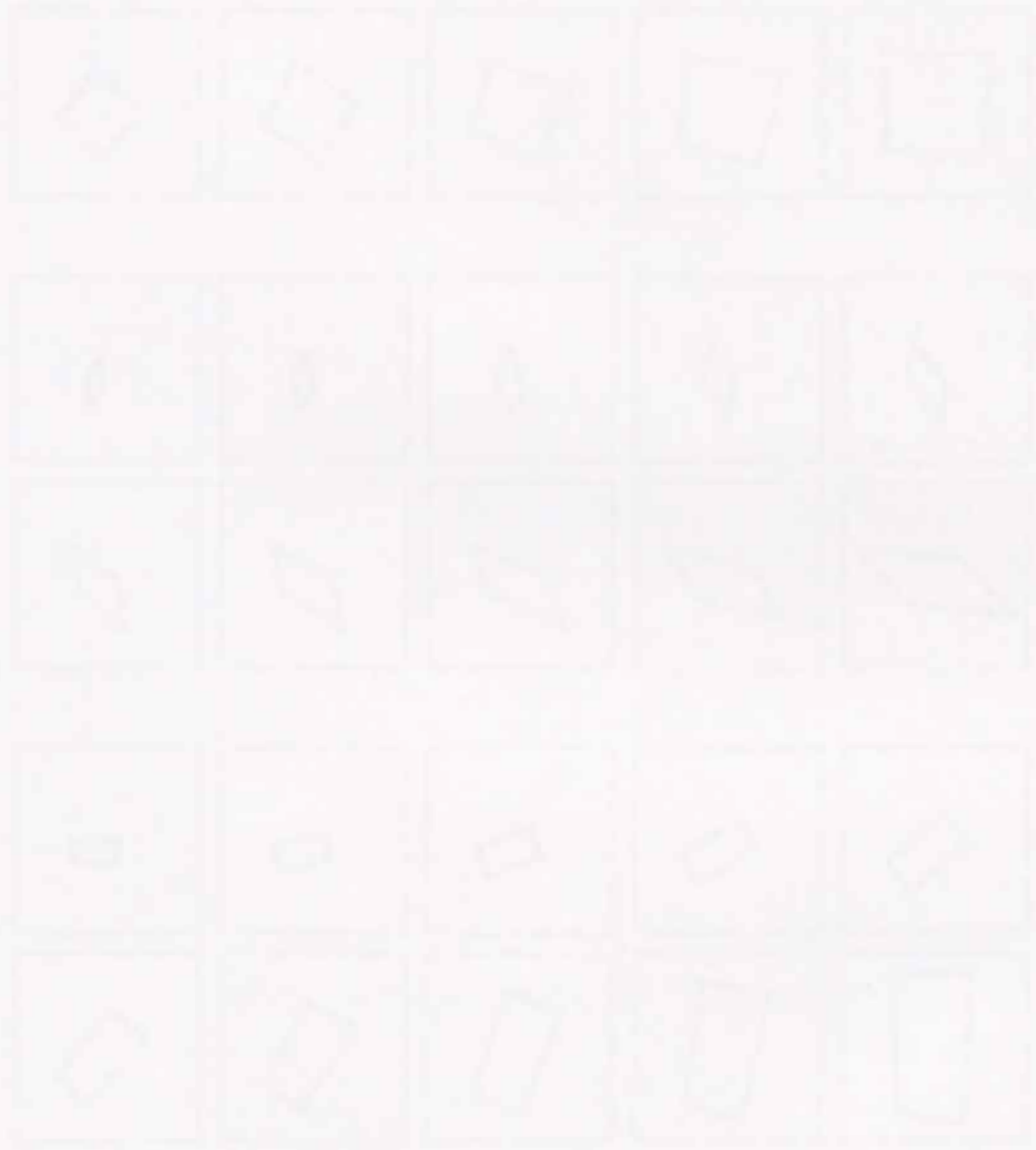
6.5.2 実験

図 6.3 と同様に 3 つの平面物体のパターンを図 6.5 に示す。この場合、各物体の 10 個のパターンはアフィン変換ではなく標準パターンを相似変換した後、小さな変形を加えて作成した。上の 2 行は第 1 物体の画像であり、中の 2 行は第 2 物体の画像、下の 2 行は第 3 物体の画像である。これら 3 つの物体はアフィン変換により関係付けられる。従って 6.2 節の方法では、これら 30 個のパターンを区別することはできない。6.5.1 節の方法を用いると図 6.6 に示すようにこれら 3 つの物体を正しく識別できる。■ 6.6 の見方は図 6.4 と同じである。

6.6 むすび

点パターンマッチングに基づき視点に不変に平面物体を認識する方法を提案した。マッチング法はアフィン変換により関係付けられたパターンに対するものと相似変換により関係付けられたパターンに対するものとを提示した。提案法は点パターンマッチングのサブシステムとして Shapiro-Brady 法を用いる。しかし Shapiro-Brady 法はノイズによる変形に敏感なので認識性能を改善するには別のより安定な方法が望まれ

る。本章では特徴点の遮蔽の問題を避けるために物体のクラスを平面物体に限定した。アスペクト検出のステップを付加することにより一般の剛体の認識への応用が可能になると思われる。提案法のような特徴点のマッチングに基づく方法においては物体の画像からの安定な特徴点抽出も重要である。特徴抽出法の検討は今後の課題である。



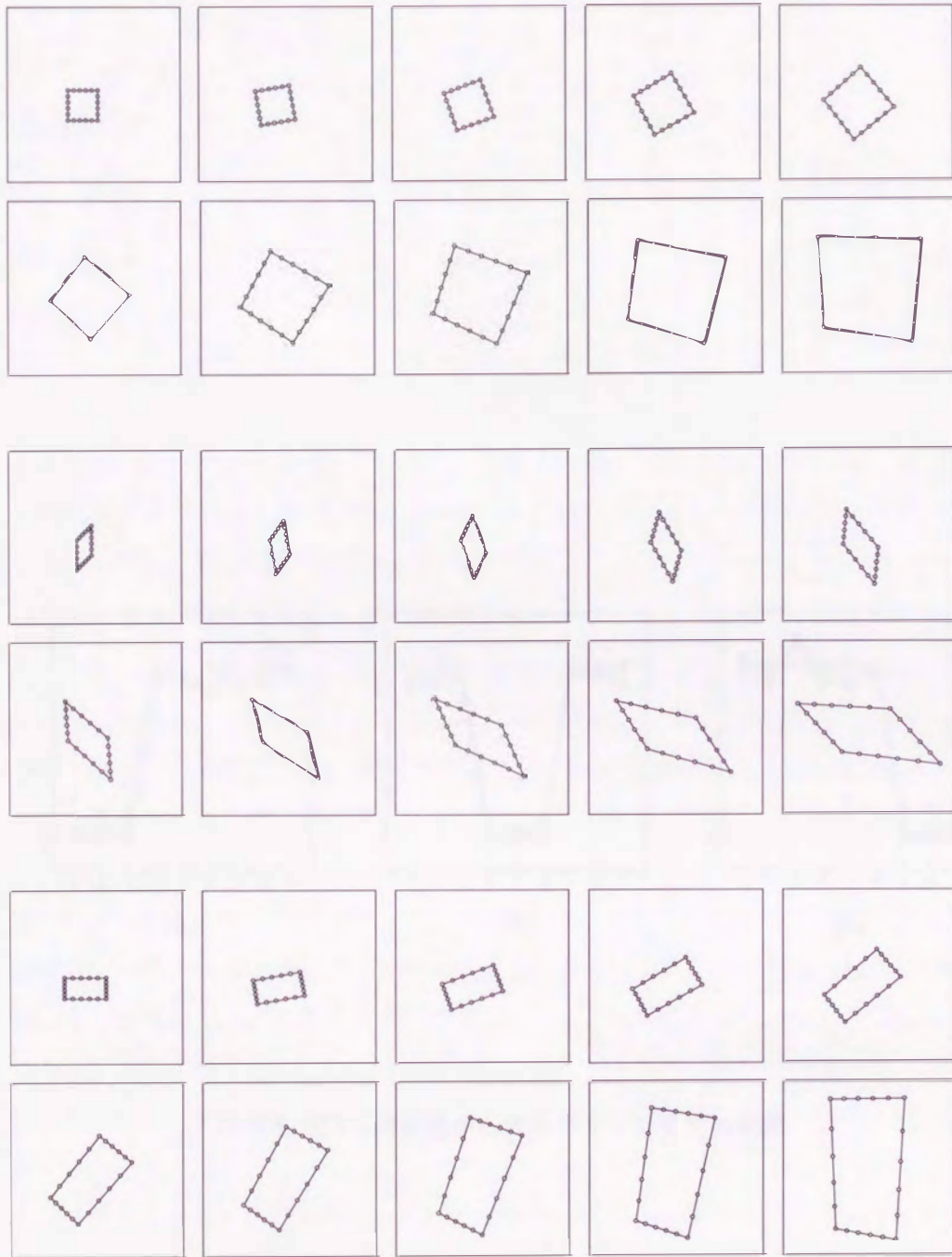
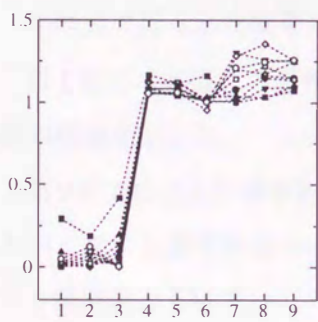
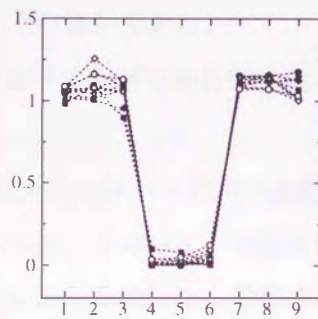


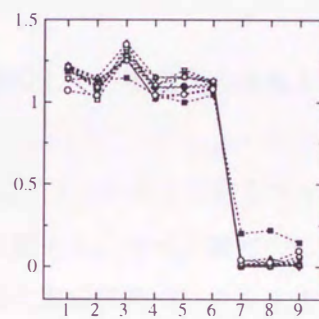
図 6.5: 小さな歪みを含む相似変換の下での3つの平面物体



(a)



(b)



(c)

図 6.6: 図 6.5 の画像から代表パターンまでの距離

第7章

結論

本研究では生体の認識機能を解明するという立場から特に視覚情報に基づくパターン認識を考えて、視点に不変な物体認識のモデル化を試みた。本論文はこの物体認識における視点不変性のモデル化に関する研究をまとめたものである。

本研究で得られた成果は以下の通りである。

第1章では本研究の背景を述べて本研究の目的を明らかにして、本論文の構成と各章の概要を示した。

第2章では文脈情報を取り入れたパターン認識の基礎としてメンバシップ値をフィードバックする機構を持つマルチモーダルパターン認識器を提案し、ベイズ識別則に基づく識別法とEMアルゴリズムによる教師なし学習法と最尤推定に基づくマルチモーダルパターンの再構成法を提案した。またマガーク効果を例に取り上げてその性質を調べた。マガーク効果はノイズのある環境下でのみ生じるという性質を再現し、心理学において報告されているいくつかのモードの入力パターンから別のモードパターンへの知覚誘導と生理学において観測されているモード情報が複数の感覚野からのフィードバックパスを通るトップダウン信号により誘導されるというニューロンの活動を説明した。

第3章では第2章のメンバシップ値をフィードバックするパターン認識器を基礎として時間的あるいは空間的な文脈を取り入れたパターン認識器を提案した。すなわち文脈情報はパターン認識器にトップダウンで与えられ、入力パターンからのボトムアップ情報を修飾するため本章で提案したモデルはトップダウン情報とボトムアップ情報を統合するマルチモーダルパターン認識器として捉えられる。時間的な文脈については1時刻前の識別結果が次の時刻にフィードバックされる例を考え、まず簡単な2次元

データを用いて、それらがパターンに類似度すなわちここでは2次元平面上でのユークリッド距離ではなくパターンの提示時刻の近接性によってクラスタリングされることを示した。次に位置不変なパターン認識の簡単な例として縦棒と横棒の認識を行い、棒の画像上での提示位置に関係なく縦棒であるか横棒であるかを識別できることを示した。空間的な文脈については空間的に1つ隣りにあるニューロンへメンバシップ値をフィードバックする例を考え、まず簡単な1次元データを用いて空間文脈を伝搬することによって空間データの緩和整合化が行われることを示した。すなわちデータに含まれるノイズは平滑化され多重データによるあいまいさは低減されデータの無い部分は文脈情報によって値が充填される。またデータの欠落と多重性を伴う空間データの例としてランダムドットステレオグラムを取り上げて視差の計算を行い、ほぼ良好な結果が得られた。空間文脈についてはBaumの増大変換[38]との関連性からその反復計算の収束性について調べた。ここでは時間的な文脈については1時刻前、空間的な文脈については1つ隣りだけを考えたが、文脈の影響範囲を広げてその効果を調べるのが今後の課題である。

第4章では第3章で提案した時間的な文脈に基づくパターン認識器を視点に不変なパターン認識に応用した。ここではPoggioら[18]の提案した2次元画像による3次元物体認識のモデルであるRBFネットとの比較を行いながら時間的な文脈に基づく視点に不変なパターン認識器の導出を行った。顔画像による個人識別を例として、視点が時間的に変化する時系列画像データを提示する教師なし学習によって視点に不変なパターン認識器が構成されることを示した。高位のWTAニューロンが生理学において観測されている顔の向きによらずに顔に応答するニューロンと似た応答をして、下位のRBFニューロンが顔の向きに選択的に応答するニューロンと似た応答を示した。またRBFニューロンの応答にロバストなクラスタ分布を用いてその効果を調べた。縦棒と横棒の画像に外れ値として斜めの棒を混入したデータで学習と識別を行い、ロバスト化によってこの外れ値の影響を取り除くことができることを示した。この結果はガウス分布による識別が全画素一致による識別であるのに対してロバスト分布による識別が多数決による識別であることから説明される。またロバスト化によって画像のセグメンテーションを陽に行うことなく注視に似た処理が行えることを示し、例として3次元物体の画像からの注視領域の抽出を行った。ここで提案したニューラルネットモデルには位相がなく、RBFニューロンの順序には意味はないが、生理実験

では視点の変化につれて活動する場所も連続的に移動することが観測されており，空間的なラテラル結合などを付け加えて位相保存性を持たせるのが今後の課題である。

第5章ではグラフスペクトル法においてグラフの接点の重みを考慮したファジークラスタの逐次抽出法を提案した。データは重み付きグラフで表され枝の重みはデータ間の類似度であり，接点の重みはファジーなデータの個数すなわち残存率である。グラフはデータ間の類似度を要素とする隣接行列で表され，各データのクラスタへのメンバシップ値が隣接行列の第1固有ベクトルとして求まる。データの残存率を隣接行列の対応する要素に乗じることにより抽出済みのクラスタを取り除きながら順にクラスタを抽出していく。適切なクラスタ数は抽出したクラスタの大きさの変化に基づいて決定される。まず簡単な2次元データで提案法の検証を行った。次にLandsat画像のセグメンテーションを例として津田ら[32]の方法との比較を行い有効性を確認した。グレイスケール画像の階調値のみによるセグメンテーションでは重みによる表現を用いることによってデータ数を画素数から階調数に変換することができ画像のサイズによらない処理が可能になる。また応用例としてカラー画像からの肌色領域の抽出を行った。本方法で調整の必要なパラメータはスケールパラメータの1つだけであり，これを調整することにより階層的なクラスタリングが可能である。

第6章では第5章の逐次クラスタ抽出法を用いて平面物体の視点に不変な認識を行った。ここでは平面物体は2次元平面上に分布する点の集合として表され，その認識は点パターンのマッチングに基づいている。マッチング結果から点パターン間の距離を測り上記のクラスタ抽出法により多数の透視射影像から代表的な視点の画像を少数個選択し，その代表画像によって1つの平面物体を表現する。入力パターンは代表画像との最近傍識別により識別される。点パターンマッチング法として2回の固有値分解を用いるアフィン不変な点パターンマッチング法を提案した。まず点パターンの重心を原点に一致させて平行移動を揃え，点パターン内の点同士の内積を要素とする行列の非零固有値の固有ベクトルを求める。この固有ベクトルは2個あり，これらがアフィン変換のスケール係数を規格化した点パターンの2次元座標値になっている。その後Shapiro-Brady法によって点パターンの反転と回転を揃える。このような2段階の処理を経てアフィン不変なマッチングが得られる。相似変換についても同様なマッチング法を提案した。ここでは特徴点の自己遮蔽の問題を避けるために対象を平面物体に限定したが，アスペクト検出のステップを付加することによって一般の剛体の認識への

応用が可能になると思われる。また提案法のような特徴点のマッチングに基づく方法においては画像からの安定な特徴抽出が重要になる。特徴抽出法の検討が今後の課題である。

第7章では本研究で得られた成果をまとめて今後の課題を述べた。

付録 A

ベイズ識別におけるメンバシップ

ベイズ識別ではデータ d の所属するクラス i ($i = 1, \dots, m$) は

$$\arg \max_i p(i|d) \quad (\text{A.1})$$

で決定される。ここで $p(i|d)$ はクラス i の事後確率である。この識別により誤識別率は最小になる [10]。ベイズ識別は次のような誤識別率の最小化問題として表される:

$$\begin{aligned} \min_q \quad & \sum_{i=1}^m q_i [1 - p(i|d)] \\ \text{subj.to} \quad & \sum_{i=1}^m q_i = 1 \\ & q_i \in \{0, 1\} \quad (i = 1, \dots, m) \end{aligned} \quad (\text{A.2})$$

ここで q_i はデータ d のクラス i へのメンバシップである。式 (A.2) は整数条件をエントロピーで表す [66] と

$$\begin{aligned} \min_q \quad & \sum_{i=1}^m q_i [1 - p(i|d)] \\ \text{subj.to} \quad & \sum_{i=1}^m q_i = 1 \\ & \sum_{i=1}^m q_i \ln q_i = 0 \end{aligned} \quad (\text{A.3})$$

となる。この最小化問題のラグランジュ関数は

$$L = \sum_{i=1}^m q_i [1 - p(i|d)] + \lambda (\sum_{i=1}^m q_i - 1) + \frac{1}{\alpha} \sum_{i=1}^m q_i \ln q_i \quad (\text{A.4})$$

となる。ここで $\lambda, 1/\alpha$ はラグランジュ乗数である。式 (A.4) を q_i と λ で偏微分して 0 とおくとそれぞれ

$$\frac{\partial L}{\partial q_i} = 1 - p(i|d) + \lambda + \frac{1}{\alpha} (\ln q_i + 1) = 0 \quad (\text{A.5})$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^m q_i - 1 = 0 \quad (\text{A.6})$$

となる. 式(A.5)と(A.6)から λ を消去すると, メンバシップ

$$q_i = \frac{e^{\alpha p(i|d)}}{\sum_{j=1}^m e^{\alpha p(j|d)}} \quad (\text{A.7})$$

が得られる. ベイズの公式

$$p(i|d) = \frac{p(d|i)p(i)}{p(d)} \quad (\text{A.8})$$

を用いると式(A.7)は

$$q_i = \frac{e^{\beta p(d|i)p(i)}}{\sum_{j=1}^m e^{\beta p(d|j)p(j)}} \quad (\text{A.9})$$

となる. ここで $p(i)$ はクラス i の事前確率, $p(d|i)$ はクラス i でのデータ d の確率密度, $p(d)$ はデータ d の確率密度である. また $\beta = \alpha/p(d)$ とおいた. 更に $p(i) = 1/m$ ($i = 1, \dots, m$)のときは式(A.9)は

$$q_i = \frac{e^{\gamma p(d|i)}}{\sum_{j=1}^m e^{\gamma p(d|j)}} \quad (\text{A.10})$$

となる. ここで $\gamma = \beta/m$ とおいた.

付録 B

EM アルゴリズムからの導出

2.2.3節では対数尤度の偏微分を0とおいて反復公式を得たが、ここではEMアルゴリズムから同じ反復公式を導く。

EMアルゴリズムでは各学習データがどのクラスタに所属するかが分からないものとして、各データ d_j の対数尤度 $\ln p(d_j, k) = \ln p(d_j|k)p(k)$ の事後確率 $p(k|d_j)$ による期待値

$$\sum_{j=1}^m \sum_{k=1}^n p(k|d_j) \ln p(d_j|k)p(k) \quad (\text{B.1})$$

を最大化することにより間接的にデータの対数尤度 $\sum_{j=1}^m \ln p(d_j)$ の最大化を図る。反復 ξ 回での事後確率 $p(k|d_j, r^{(\xi)}) = \prod_{i'=1}^l p(d_{i'j}|k, r^{(\xi)}) / \sum_{k'=1}^n \prod_{i'=1}^l p(d_{i'j}|k', r^{(\xi)})$ による、 $\xi+1$ 回目の反復における対数尤度の期待値は

$$\sum_{j=1}^m \sum_{k=1}^n \frac{\prod_{i'=1}^l p(d_{i'j}|k, r^{(\xi)})}{\sum_{k'=1}^n \prod_{i'=1}^l p(d_{i'j}|k', r^{(\xi)})} \sum_{i=1}^l \ln p(d_{ij}|k, r^{(\xi+1)}) \quad (\text{B.2})$$

となる。これを $r_{ik}^{(\xi+1)}$ で偏微分すると

$$\sum_{j=1}^m \frac{\prod_{i'=1}^l p(d_{i'j}|k, r^{(\xi)})}{\sum_{k'=1}^n \prod_{i'=1}^l p(d_{i'j}|k', r^{(\xi)})} \cdot \frac{1}{p(d_{ij}|k, r^{(\xi+1)})} e^{-a_i \|d_{ij} - r_{ik}^{(\xi+1)}\|^2} (d_{ij} - r_{ik}^{(\xi+1)}) \quad (\text{B.3})$$

となる。ここで $p(d_{ij}|k, r^{(\xi+1)}) \approx p(d_{ij}|k, r^{(\xi)})$ と近似したものを0とおくと、反復公式

$$r_{ik}^{(\xi+1)} = \frac{\sum_{j=1}^m \frac{\prod_{i' \neq i}^l p(d_{i'j}|k, r^{(\xi)})}{\sum_{k'=1}^n \prod_{i'=1}^l p(d_{i'j}|k', r^{(\xi)})} d_{ij} e^{-a_i \|d_{ij} - r_{ik}^{(\xi)}\|^2}}{\sum_{j=1}^m \frac{\prod_{i' \neq i}^l p(d_{i'j}|k, r^{(\xi)})}{\sum_{k'=1}^n \prod_{i'=1}^l p(d_{i'j}|k', r^{(\xi)})} e^{-a_i \|d_{ij} - r_{ik}^{(\xi)}\|^2}} \quad (\text{B.4})$$

が得られる。

付録 C

EM アルゴリズムによる混合密度推定とファジークラスタリング

データ d の確率密度を混合密度

$$p(d) = \sum_{j=1}^n p(d|j)p(j) \quad (\text{C.1})$$

で表す. n はクラスタの個数であり, $p(j)$ は第 j クラスタの事前確率であり, $p(d|j)$ は第 j クラスタでのデータ d の確率密度である. m 個の学習データ $\{d_i\} (i = 1, \dots, m)$ が与えられるとき

$$\max \sum_{i=1}^m \sum_{j=1}^n p(j|d_i) \ln p(d_i|j)p(j) \quad (\text{C.2})$$

によって学習するのが EM アルゴリズムである. $p(j|d)$ は第 j クラスタの事後確率である. ここで $p(d|j) \propto e^{-a\|d-r_j\|^2}$, $p(j) = 1/n$ の場合を考える. a は正定数, r_j は第 j クラスタの代表点である. このとき式 (C.2) は

$$\min \sum_{i=1}^m \sum_{j=1}^n p(j|d_i) \|d_i - r_j\|^2 \quad (\text{C.3})$$

とできる. これは $p(j|d_i)$ をデータ d_i の第 j クラスタへのメンバシップとして捉えるとファジークラスタリングの評価関数になっている. 従って混合密度推定における EM アルゴリズムは特別な場合としてファジークラスタリングを含む, より一般的な学習法であるといえる.

付録 D

アニーリングの性質

$p(d|i)$ の具体例 $e^{-a\|d-r_i\|^2}$ について、式(4.6)の \sum_i と対数を取り除いた

$$\max_r \sum_{i=1}^m p(i) e^{-a\|d-r_i\|^2} \quad (\text{D.1})$$

は

$$\min_{x,r} \sum_{i=1}^m p(i) [x_i \|d-r_i\|^2 + \frac{1}{a} x_i (\ln x_i - 1)] \quad (\text{D.2})$$

と等価である (max と min の違いに注意). なぜなら式 (D.2) を x_i で偏微分して 0 とおけば x_i が $x_i = e^{-a\|d-r_i\|^2}$ と求まるからこれを式 (D.2) に代入すると式 (D.2) は式 (D.1) となるからである. 式 (D.2) はクラスタリングの評価関数であり, メンバシップ x_i が大きい (すなわち r_i に所属する) データ d の量子化誤差 $\|d-r_i\|^2$ を最小にすることを表す. 第2項はファジー化のためのエントロピーである. 従って4章の学習はファジークラスタリングと等価であり, ファジークラスタリングでのアニーリング [66] と同じく,

[性質] a が増大するに従い全量子化誤差 $\sum_i \sum_{l=1}^m p(i) p(d^{(l)}|i)$ は単調に減少する.

が成り立つ. 従って r_i は第 i クラスターの重心に大域的に収束することが期待される.

付録 E

津田らの方法の重み付きデータへの拡張

津田らの方法 [32, 56] では, 2個のクラスタ l, k 間の重なりが第 l, k 固有クラスタの内積 $x_l^T x_k = \sum_i^n x_{li} x_{ki}$ で与えられる. 抽出済みの $(k-1)$ 個の固有クラスタを x_l ($l = 1, \dots, k-1$) とするとき第 k 固有クラスタ x_k は

$$\begin{aligned} \max_{x_k} \quad & \sum_i^n \sum_j^n h_{ij} x_{ki} x_{kj} - \frac{\eta}{k-1} \sum_l^{k-1} \lambda_l \sum_i^n x_{li} x_{ki} \\ \text{subj.to} \quad & \sum_i^n x_{ki}^2 = 1 \end{aligned} \quad (\text{E.1})$$

の解として求まる. h_{ij} はデータ d_i, d_j 間の近接度である. 式 (E.1) は新たに抽出する固有クラスタ x_k と抽出済みの x_l ($l = 1, \dots, k-1$) との重なりを重み付き平均をペナルティ項として式 (5.1) に加えた形になっている. η はペナルティ項 (式 (E.1) 第2項) の影響を調整するパラメータである. $H_k = [h_{kij}]$; $h_{kij} = h_{ij}$, $h_{kii} = -\eta / (k-1) \sum_l^{k-1} \lambda_l x_{li}$ とおくと式 (E.1) の解 x_k は H_k の第1固有ベクトルとなる. λ_l は H_l の最大固有値である.

以上が津田らの方法であり接点の重みは特に考慮されていない. すなわち全接点の重みは1になっている. そこで5章と同様に i 番目の接点の重みを w_i とすると式 (E.1) は

$$\begin{aligned} \max_{x_k} \quad & \sum_i^n \sum_j^n w_i w_j h_{ij} x_{ki} x_{kj} - \frac{\eta}{k-1} \sum_l^{k-1} \lambda_l \sum_i^n w_i x_{li} x_{ki} \\ \text{subj.to} \quad & \sum_i^n w_i x_{ki}^2 = 1 \end{aligned} \quad (\text{E.2})$$

となる. ここで $y_{ki} = \sqrt{w_i} x_{ki}$ とおくと式 (E.2) は

$$\begin{aligned} \max_{x_k} \quad & \sum_i^n \sum_j^n \sqrt{w_i} \sqrt{w_j} h_{ij} y_{ki} y_{kj} - \frac{\eta}{k-1} \sum_l^{k-1} \tilde{\lambda}_l \sum_i^n \frac{y_{li} y_{ki}}{w_i} \\ \text{subj.to} \quad & \sum_i^n y_{ki}^2 = 1 \end{aligned} \quad (\text{E.3})$$

となる. $\tilde{H}_k = [\tilde{h}_{kij}]$; $\tilde{h}_{kij} = \sqrt{w_i}\sqrt{w_j}h_{ij}$, $\tilde{h}_{kii} = -\eta/(k-1)\sum_l^{k-1} \tilde{\lambda}_l y_{li}/w_i$ とおくと式 (E.3) の解は \tilde{H}_k の第1固有ベクトルとなる. $\tilde{\lambda}_l$ は \tilde{H}_l の最大固有値である.

付録 F

透視射影の線形近似

3次元空間において視点の位置を原点として正規直交基底 $x = (1, 0, 0)^T, y = (0, 1, 0)^T, z = (0, 0, 1)^T$ を定めて視線の向きを z として n 個の点の集合 $P = \{p_i\}, p_i = (x_i, y_i, z_i)^T$ ($i = 1, \dots, n$) を視線上の点 $(0, 0, f)^T$ を通り視線に垂直な平面に透視射影する. 射影された点の平面上での座標を $(u_i, v_i)^T$ ($i = 1, \dots, n$) とすると

$$u_i = f \frac{x_i}{z_i} \quad (\text{F.1})$$

$$v_i = f \frac{y_i}{z_i} \quad (\text{F.2})$$

となる. これの最も簡単な近似は次の直交射影である:

$$u_i^a = x_i \quad (\text{F.3})$$

$$v_i^a = y_i \quad (\text{F.4})$$

直交射影では各点 p_i の x, y 座標がそのまま平面上での座標になるため奥行き情報が失われる. z 軸方向の距離に基づくスケールリングを考慮したものが次の弱透視射影 (weak perspective projection) である:

$$u_i^b = f \frac{x_i}{z_0} \quad (\text{F.5})$$

$$v_i^b = f \frac{y_i}{z_0} \quad (\text{F.6})$$

ここで

$$z_0 = \frac{1}{n} \sum_{i=1}^n z_i \quad (\text{F.7})$$

である。更に点集合 P の視線からのずれによる射影像の変形を考慮したのが次の擬似透視射影 (paraperspective projection)[67] である:

$$u_i^c = \frac{f}{z_0} \left(x_i - \frac{x_0}{z_0} z_i + x_0 \right) \quad (\text{F.8})$$

$$v_i^c = \frac{f}{z_0} \left(y_i - \frac{y_0}{z_0} z_i + y_0 \right) \quad (\text{F.9})$$

ここで

$$x_0 = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{F.10})$$

$$y_0 = \frac{1}{n} \sum_{i=1}^n y_i \quad (\text{F.11})$$

である。 $g = (x_0, y_0, z_0)^T$ は点集合 P の重心であり参照点 (reference point)[68] と呼ばれる。ここで式 (F.1) を参照点 g の周りでテーラー展開して1次の項までを見てみる:

$$\begin{aligned} u_i &\approx f \frac{x_0}{z_0} + f \left[\frac{\partial u_i}{\partial x} (x_i - x_0) + \frac{\partial u_i}{\partial z} (z_i - z_0) \right] \\ &= f \frac{x_0}{z_0} + f \left[\frac{1}{z_0} (x_i - x_0) - \frac{x_0}{z_0^2} (z_i - z_0) \right] \\ &= \frac{f}{z_0} \left(x_i - \frac{x_0}{z_0} z_i + x_0 \right) \end{aligned} \quad (\text{F.12})$$

これから擬似透視射影が透視射影の1次近似であることが分かる。また $g = (0, 0, z_0)^T$ のとき擬似透視射影は弱透視射影に一致する。弱透視射影と擬似透視射影は線形変換でありそれぞれ次のようにアフィン変換で表現される。式 (F.5) と (F.6) を行列形式で表すと

$$\begin{bmatrix} u_i^b \\ v_i^b \end{bmatrix} = \frac{f}{z_0} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (\text{F.13})$$

となる。また式 (F.8) と (F.9) を行列形式で表すと

$$\begin{bmatrix} u_i^c \\ v_i^c \end{bmatrix} = \frac{f}{z_0} \begin{bmatrix} 1 & 0 & -x_0/z_0 & x_0 \\ 0 & 1 & -y_0/z_0 & y_0 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (\text{F.14})$$

となる。

付録 G

6.2節のマッチングのアフィン不変性

互いにアフィン変換で関係付けられる2つのパターン $\{p_i\}$ と $\{q_i\}$ を考える. P と Q を $2 \times n$ 行列 $P = [\bar{p}_1, \dots, \bar{p}_n], Q = [\bar{q}_1, \dots, \bar{q}_n]$ とする. ここで $\bar{p}_i = p_i - \sum p_i/n, \bar{q}_i = q_i - \sum q_i/n$ である. 6.2.1節の第1段階でパターン $\{p_i\}$ は行列 $C = P^T P$ の固有ベクトル u_1, u_2 に写像され, パターン $\{q_i\}$ は行列 $D = Q^T Q$ の固有ベクトル v_1, v_2 に写像される. この写像は次の性質を満たす:

命題 1 $\{p_i\}$ と $\{q_i\}$ がアフィン変換によって関係付けられるならばパターン $U = [u_1, u_2]^T$ と $V = [v_1, v_2]^T$ は回転と反転によって関係付けられる.

(証明) 2つのパターン $\{p_i\}$ と $\{q_i\}$ はアフィン変換によって一致するので

$$Q = APT \quad (\text{G.1})$$

となるような 2×2 行列 A と $n \times n$ 置換行列 T が存在する. マッチングはこの置換行列 T を決定する問題である. 明らかに $C = P^T P$ のランクは P のそれと同じく 2 である. 従って行列 C は 2 つの非零固有値 λ_1, λ_2 をもつ. C の固有値分解は

$$CU^T = U^T \Lambda \quad (\text{G.2})$$

と書ける. ここで $U = [u_1, u_2]^T, \Lambda = \text{diag}(\lambda_1, \lambda_2)$ である. $C = P^T P$ と式 (G.2) から

$$P = (UP^T)^{-1} \Lambda U \quad (\text{G.3})$$

を得る. 式 (G.3) を式 (G.1) に代入すると

$$Q = A(UP^T)^{-1} \Lambda UT = BUT \quad (\text{G.4})$$

となる. ここで B は $B = A(UP^T)^{-1}\Lambda$ と定義する. 式 (G.4) から

$$D = Q^T Q = T^T U^T B^T B U T \quad (\text{G.5})$$

を得る. $B^T B$ の固有値分解を

$$B^T B = R_B^T \Lambda_B R_B \quad (\text{G.6})$$

とする. 式 (G.6) を式 (G.5) に代入すると

$$D = T^T U^T R_B^T \Lambda_B R_B U T \quad (\text{G.7})$$

となり, これから行列 D の固有ベクトル $V = [v_1, v_2]^T$ が

$$V = R_B U T \quad (\text{G.8})$$

と表されることが分かる. これは U と V が回転と反転で一致することを示す. (証明終)

次に第2段階では U と V は Shapiro-Brady 法により高次元空間に写像される. ここでパターンの近接行列 E は点の間の距離からなる. 従って E はパターンの回転と反転には不変である. それゆえパターン U の近接行列 E_U とパターン V の近接行列 E_V は $E_V = T^T E_U T$ で関係付けられ, それらの固有ベクトルをならべた行列 X と Y は $Y = XT$ で関係付けられる. 従って X と Y を直接マッチングすることにより元のパターン $\{p_i\}$ と $\{q_i\}$ の間の対応が得られる. 結果としてこのマッチングは元のパターンに対してアフィン不変である.

付録 H

6.5節のマッチングの相似不変性

6.5.1節ではパターン $P = [\tilde{p}_1, \dots, \tilde{p}_n]$ は $P' = P/\sqrt{\text{tr}(P^T P)}$ に変換される. 同様に $Q = [\tilde{q}_1, \dots, \tilde{q}_n]$ は $Q' = Q/\sqrt{\text{tr}(Q^T Q)}$ に変換される. この変換は次の性質を満たす:

命題 2 $\{p_i\}$ と $\{q_i\}$ が相似変換によって関係付けられるならばパターン P' と Q' は回転と反転で関係付けられる.

(証明) $\text{tr}(P^T P) = \lambda_1 + \lambda_2$ であるから $P' = P/\sqrt{\text{tr}(P^T P)}$ は $P' = P/\sqrt{\lambda_1 + \lambda_2}$ と書き直せる. 式 (G.3) をこの表現に代入すると

$$P' = (UP^T)^{-1}\Lambda U/\sqrt{\lambda_1 + \lambda_2} \quad (\text{H.1})$$

を得る. 2つのパターン $\{p_i\}$ と $\{q_i\}$ は相似変換を通して一致するので

$$Q = sRPT' \quad (\text{H.2})$$

となるようなスカラー s と 2×2 回転行列 R と $n \times n$ 置換行列 T が存在する. 式 (G.3) を式 (H.2) に代入すると

$$Q = sR(UP^T)^{-1}\Lambda UT \quad (\text{H.3})$$

となる. 式 (H.2) から $Q^T Q$ の固有値は $s^2\lambda_1$ と $s^2\lambda_2$ であり, 従って Q は $Q' = Q/(s\sqrt{\lambda_1 + \lambda_2})$ に写像される. この式に式 (H.3) を代入すると

$$Q' = R(UP^T)^{-1}\Lambda UT/\sqrt{\lambda_1 + \lambda_2} = RP'T' \quad (\text{H.4})$$

を得る. この関係は P' と Q' が回転と反転で関係付けられることを示す. (証明終)

次のように命題 2 の逆もまた成り立つ:

命題 3 パターン P' と Q' が回転と反転によって関係付けられるならば $\{p_i\}$ と $\{q_i\}$ は相

似変換で関係付けられる。

(証明) $P^T P$ の固有値を $\lambda_{p1}, \lambda_{p2}$ とし, $Q^T Q$ のそれを $\lambda_{q1}, \lambda_{q2}$ とすると, $P' = P/\sqrt{\lambda_{p1} + \lambda_{p2}}$, $Q' = Q/\sqrt{\lambda_{q1} + \lambda_{q2}}$ である. P' と Q' は回転と反転で関係付けられるので $Q' = RP'T$ となるような直交行列 R と置換行列 T が存在する. 上の P' と Q' の表現をこの式に代入すると

$$Q = \sqrt{\frac{\lambda_{q1} + \lambda_{q2}}{\lambda_{p1} + \lambda_{p2}}} R P T \quad (\text{H.5})$$

を得る. これは $s = \sqrt{\lambda_{q1} + \lambda_{q2}}/\sqrt{\lambda_{p1} + \lambda_{p2}}$ とすると式(H.2)になる. これは $\{p_i\}$ と $\{q_i\}$ が相似変換で関係付けられることを示す. (証明終)

命題2と3を組み合わせて次の系を得る:

系 1 パターン $\{p_i\}$ と $\{q_i\}$ が相似変換で関係付けられるための必要十分条件は P' と Q' が回転と反転で関係付けられることである.

この系は相似変換以外のアフィン変換で関係付けられるパターンは6.5.1節の方法ではマッチングされないことを保証する. Shapiro-Brady法の回転不変性はすでに付録Gで示されている.

謝 辞

本研究を遂行し、論文にまとめるにあたり、御指導、御鞭撻を頂いた九州芸術工科大学 画像設計学科 浦浜喜一教授、瀧山龍三教授に深く感謝します。浦浜喜一教授には毎日の議論を通して多くのことを学び、個々の問題について深く考える機会を与えて頂きました。また私がパターン認識の分野に関心を持ち研究を続けることができたのは、瀧山龍三教授をはじめ、画像工学研究室の先生方の御指導によるものであります。心から感謝します。

適切な御意見と御助言を頂いた九州芸術工科大学大学院 芸術工学研究科 福島重廣教授、九州芸術工科大学 画像設計学科 長島健次教授に深く感謝します。

有益な御助言と励ましを頂いた九州芸術工科大学 画像設計学科 小野直樹助教授、坂本博康講師、九州芸術工科大学大学院 芸術工学研究科 吉永幸靖助手、長崎総合科学大学 機械工学科 金子照之講師に深く感謝します。

有益なアドバイスと議論をして頂いた株式会社富士通九州システムエンジニアリング 松永浩之氏、九州職業能力開発大学校 情報技術科 岡田正之講師、日本電信電話株式会社 NTT コミュニケーション科学研究所 山田辰美氏、九州日本電気ソフトウェア株式会社 相良哲生氏、九州芸術工科大学大学院 博士後期課程 武市義弘氏、堀田政二氏、同大学院 博士前期課程 原泰久氏、村尾晶弘氏、そして画像工学研究室の皆様へ深く感謝します。

参考文献

- [1] 小川英光(編著), “パターン認識・理解の新たな展開 —挑戦すべき課題—,” 社団法人 電子情報通信学会, 1994.
- [2] 上坂吉則, “ニューロコンピューティングの数学的基礎,” 近代科学社, 1993.
- [3] S. Theodoridis and K. Koutroumbas, “Pattern recognition,” Academic Press, 1999.
- [4] 宮本定明, “ファジィ理論のクラスタリングへの応用,” システム制御情報学会, 1992.
- [5] J.T. Tou and R.C. Gonzalez, “Pattern recognition principles,” Addison-Wesley Publishing Company, 1974.
- [6] 坂和正敏, “ファジィ理論の基礎と応用,” 森北出版, 1989.
- [7] J.C. Bezdek, “Pattern recognition with fuzzy objective function algorithms,” Prentice Hall, New York, 1981.
- [8] R.N. Davé and R. Krishnapuram, “Robust clustering methods: A unified view,” IEEE trans. Patt. Anal. Mach. Intelli., 5, 2, pp.270-293, 1997.
- [9] C.J. Alpert, A.B. Kahng and S. Yao, “Spectra of graphs,” Academic Press, N.Y., 1979.
- [10] R.O. Duda and P.E. Hart, “Pattern classification and scene analysis,” John Wiley & Sons, Inc., 1973.
- [11] C.M. Bishop, “Neural networks for pattern recognition,” Oxford University Press Inc., 1995.

- [12] 乾敏郎 (編著), “認知と学習,” 丸善, 1993.
- [13] 乾敏郎, 安藤広志 (訳), “ビジョン —視覚の計算理論と脳内表現—,” 産業図書, 1987.(D. Marr, “Vision,” W.H. Freeman and Company, N.Y.,1982.)
- [14] 佐藤淳, “コンピュータビジョン —視覚の幾何学—,” コロナ社, 1999.
- [15] 徐剛, 辻三郎, “3次元ビジョン,” 共立出版, 1998.
- [16] S. Ullman and R. Basri, “Recognition by linear combination of models,” *IEEE Trans. Patt. Anal. Mach. Intelli.*, 13, 10, pp.992-1006, 1991.
- [17] 村瀬洋, S.K. Nayar, “2次元照合による3次元物体認識 —パラメトリック固有空間法—,” *信学論*, J77-D-II, 11, pp.2179-2187, 1994.
- [18] T. Poggio and S. Edelman, “A network that learns to recognize three dimensional objects,” *Nature*, 343, pp.263-266, 1990.
- [19] D.I. Perret and M.W. Oram, “The neurophysiology of shape processing,” *Vision Image Comput.*, 11, 6, pp.317-333, 1993.
- [20] P. Foldiak, “Learning invariance from transformation sequences,” *Neural Comp.*, 3, pp.194-200, 1991.
- [21] E.T. Rolls, “Learning mechanisms in the temporal lobe visual cortex,” *Behav. Brain Res.*, 66, pp.177-185, 1995.
- [22] M. Bartlett, M. Stewart and T. Sejnowski, “Unsupervised learning of invariant representations of faces through temporal association,” in “*Computational Neuroscience: International Review of Neurobiology Suppl.*,” J.M. Bower, Ed., pp.317-322, 1996.
- [23] S. Becker, “Implicit learning in 3D object recognition: The importance of temporal context,” *Neural Comp.*, 11, pp.347-374, 1999.
- [24] K. Sakai and Y. Miyashita, “Neural organization for the long-term memory of paired associates,” *Nature*, 354, pp.152-155, 1991.

- [25] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature*, 264, pp.746-748, 1976.
- [26] D.W. Massaro and D. Friedman, "Models of integration given multiple sources of information," *Psychol. Rev.*, 97, pp.225-252, 1990.
- [27] J. Robert-Ribes, J.L. Schwartz and P. Escudier, "A comparison of models for fusion of the auditory and visual sensors in speech perception," *Artif. Intelli. Rev.*, 9, pp.323-346, 1995.
- [28] V. de Sa, "Minimizing disagreement for self-supervised classification," 1993 Connectionist Models Summer School, pp.300-307, 1993.
- [29] K. Yamaguchi, M. Oota and N. Ishii, "A self-supervised learning system for category detection by sensory integration," *Prog. Neural Inf. Process. (ICONIP'96)*, pp.103-108, 1996.
- [30] S. Akaho, S. Hayamizu, O. Hasegawa, T. Yoshimura and H. Asoh, "Concept acquisition from multiple information sources by the EM algorithm," *IEICE Trans. A*, J80-A, 9, pp.1546-1553, 1997.
- [31] 松永浩之, 浦浜喜一, "マルチモーダルパターン識別器の教師なし学習," *信学論*, J81-D-II, 3, pp.574-582, 1998.
- [32] 津田宏治, 仙田修司, 美濃導彦, 池田克夫, "逐次的ファジークラスタ抽出法とそのノイズに対するロバスト性," *信学論*, J80-D-II, 1, pp.190-197, 1997.
- [33] S. Ullman, "High level vision," MIT Press, 1996.
- [34] L.S. Shapiro and J.M. Brady, "Feature-based correspondence: An eigenvector approach," *Image Vision Comput.*, 10, pp.283-288, 1992.
- [35] K. Inoue and K. Urahama, "Multimodal pattern classifiers with feedback of class memberships," *IEICE Trans., Inf. & Syst.*, E82-D, 3, pp.712-716, 1999.

- [36] K. Sekiyama, "Differences in auditory-visual speech perception between Japanese and Americans: McGurk effect as a function of incompatibility," *J. Acoust. Soc. Japan*, 15, pp.143-158, 1994.
- [37] V.R. de Sa and D. Ballard, "Perceptual learning from crossmodal feedback," in "Psychology of Learning and Motivation," R.L. Goldstone, P.G. Schyns and D.L. Medin (eds), 36, pp.309-351, 1997.
- [38] L.E. Baum and G.R. Sell, "Growth transformation for functions on manifolds," *Pac. J. Math.*, 27, pp.211-227, 1968.
- [39] K. Urahama, "Equivalence between some dynamical systems for optimization," *Neural process. Lett.*, 1, 2, pp.14-17, 1994.
- [40] K. Sekiyama and Y. Tohkura, "Role of lip-read information in speech perception," *ITEJ Tech. Rep.*, 13, 44, pp.31-36, 1989.
- [41] A.M. Sillito, K.L. Grieve, H.E. Jones, J. Cudeiro and J. Davis, "Visual cortical mechanisms detecting focal orientation discontinuities," *Nature*, 378, pp.492-496, 1995.
- [42] 井上光平, 浦浜喜一, "メンバシップフィードバックによる文脈伝搬," *信学論*, J82-D-II, 3, pp.494-500, 1999.
- [43] S. Becker, "Learning temporally persistent hierarchical representations," *Adv. Neural Inf. Process. Syst. (NIPS)*, 9, pp.824-830, 1996.
- [44] R.A. Hummel and S.W. Zucker, "On the foundations of relaxation labeling processes," *IEEE Trans. Patt. Anal. mach. Intelli.*, 5, 3, pp.267-287, 1983.
- [45] 井上光平, 浦浜喜一, "時間的な文脈に基づく視点に不変なパターン認識器の学習," *信学技報*, NC98-43, pp.1-8, 1998.
- [46] K. Inoue and K. Urahama, "Learning of view-invariant pattern recognizer with temporal context," *Patt. Recog.*, (印刷中), 1999.

- [47] N.K. Logothetis, J. Pauls, H.H. Bulthoff and T. Poggio, "View-dependent object recognition by monkeys," *Curr. Biol.*, 4, pp.401-414, 1994.
- [48] D. Weinshall and S. Edelman, "A self-organizing multiple view representation of 3D objects," *Biol. Cybern.*, 64, 3, pp.209-219, 1991.
- [49] S.M. Kosslyn, N.M. Alpert, W.L. Thompson, V. Maljkovic, S.B. Weise, C.F. Chabris, S.E. Hamilton, S.L. Rauch and F.S. Buonanno, "Visual mental imagery activates topographically organized visual cortex: PET investigations," *J. Cog. Neurosci.*, 5, pp.263-287, 1993.
- [50] K. Inoue and K. Urahama, "Sequential fuzzy cluster extraction by a graph spectral method," *Patt. Recog. Lett.*, 20, 7, pp.699-705, 1999.
- [51] R. Krishnapuram, J.M. Keller, "A possibilistic approach to clustering," *IEEE Trans. Fuzzy Syst.*, 1, pp.98-110, 1993.
- [52] R.R. Yager, D.P. Filev, "Approximate clustering via the mountain method," *IEEE Trans. Syst. Man Cybernet.*, 24, 8, pp.1279-1284, 1994.
- [53] S. Chiu, "Fuzzy model identification based on cluster estimation," *J. Intell. Fuzzy Syst.*, 2, 3, pp.267-278, 1994.
- [54] D.M. Cvetkovic, M. Doob, H. Sachs, "Spectra of graphs," Academic Press, New York, 1980.
- [55] S. Sarkar, K.L. Boyer, "Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors," *Comput. Vision Image Understanding*, 71, 1, pp.110-136, 1998.
- [56] K. Tsuda, M. Minoh, K. Ikeda, "Extracting straight lines by sequential fuzzy clustering," *Patt. Recog. Lett.*, 17, pp.643-649, 1996.
- [57] 井上光平, 相良哲生, 浦浜喜一, 瀧山龍三, "点パターンマッチングに基づく平面物体の視点に不変な認識," *信学技報*, PRMU99-80, pp.19-24, 1999.

- [58] A. Zisserman, D. Forsyth, J. Mundy, C. Rothwell, J. Liu and N. Pillow, "3D object recognition using invariance," *Artif. Intelli.*, 78, pp.239-288, 1995.
- [59] S. Startchik, R. Milaness, C. Rauber and T. Pun, "Planar shape databases with affine invariant search," in "Image databases and Multi-media Search," A. W. M. Smeulders and R. Jain (eds), pp.285-292, World Scientific, 1997.
- [60] M. Werman and D. Weinshall, "Similarity and affine invariant distances between 2D point sets," *IEEE Trans. Patt. Anal. Mach. Intelli.*, 17, 8, pp.810-814, 1995.
- [61] A.D.J. Cross and E.R. Hancock, "Graph matching with a dual-step EM algorithm," *IEEE Trans. Patt. Anal. Mach. Intelli.*, 20, 11, pp.1236-1253, 1998.
- [62] G.L. Scott and H.C. Longuet-Higgins, "An algorithm for associating the features of two images," *Proc. Roy. Soc. London*, 244, pp.21-26, 1991.
- [63] B. Günsel and A.M. Tekalp, "Shape similarity matching for query-by-example," *Patt. Recog.*, 31, 7, pp.931-944, 1998.
- [64] S. Sclaroff and A. Pentland, "Modal matching for correspondence and recognition," *IEEE Trans. Patt. Anal. Mach. Intelli.*, 17, 6, pp.545-561, 1995.
- [65] 堀田政二, 井上光平, 浦浜喜一, "ファジークラスタリングによる画像ナビゲーション," *信学技報*, NC98-48, pp.41-46, 1998.
- [66] 浦浜喜一, "組合せ最適化のニューラル解法の数理計画論的基礎," *信学論*, J78-A, 1, pp.47-53, 1995.
- [67] Y. Ohta, K. Maenobu and T. Sakai, "Obtaining surface orientation from texels under perspective projection," *Proc. 7th Int. Joint Conf. Artif. Intelli.*, pp.746-751, 1981.
- [68] R. Basri, "Paraperspective \equiv affine," *Int. J. Comput. Vision*, 19, 2, pp.169-179, 1996.

目次

2.1	2モードパターン認識のためのニューラルネットワーク	15
2.2	入力データ(●)と代表点(○)	19
2.3	出力 q_k	20
2.4	図2.3と逆の入力に対する出力 q_k	21
2.5	聴覚-視覚空間における音の配置	23
2.6	p_{ik} の初期値	23
2.7	p_{ik} の収束値	24
2.8	マガーク入力からの再構成値	24
2.9	d_1 からの再構成	25
2.10	d_2 からの再構成	25
3.1	時間文脈伝搬ネットの例	29
3.2	クラスタリングの例	32
3.3	時間文脈クラスタリング	32
3.4	第1グループの代表点の収束の様子	34
3.5	学習用画像データ	34
3.6	代表画像の初期値	35
3.7	代表画像の学習結果	35
3.8	空間文脈伝搬ネットの例	38
3.9	空間データ例	40
3.10	収束値	40
3.11	ランダムドットステレオグラム	42
3.12	データの欠落部と多重部	42

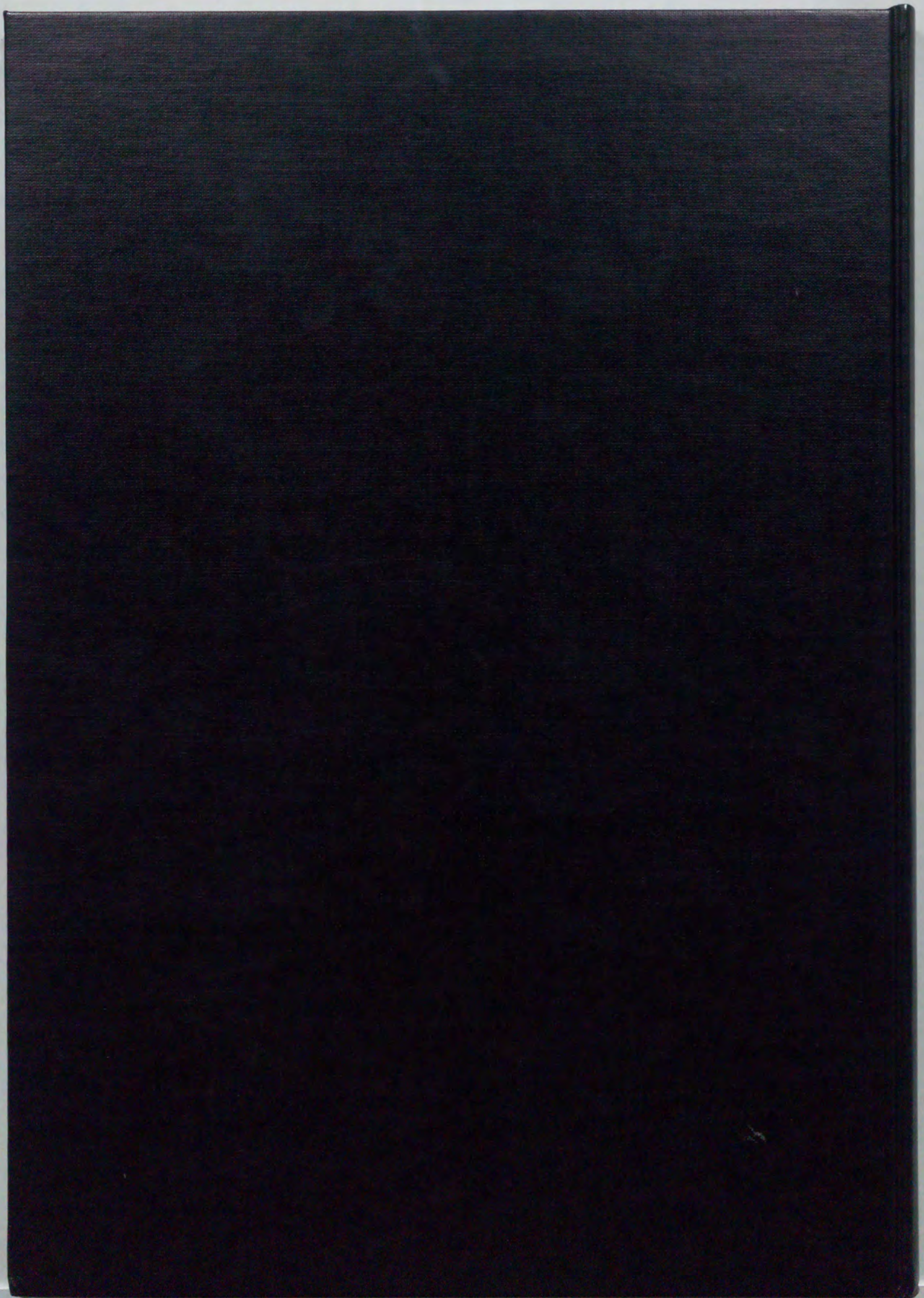
3.13 得られた視差	43
4.1 RBF ネット	45
4.2 顔ニューロンの応答	47
4.3 時間文脈伝搬ネット	49
4.4 顔画像の例	52
4.5 代表画像の初期値	53
4.6 代表画像	54
4.7 文脈なし学習の代表画像	55
4.8 RニューロンとWニューロンの応答	57
4.9 外れ値を含む画像データ例	59
4.10 代表画像	60
4.11 テスト画像	60
4.12 3次元物体の画像例	62
4.13 テスト画像	62
4.14 図4.12 下列の画像の背景	65
4.15 図4.13での注視領域	65
5.1 データの例	71
5.2 第1(左), 第5(右) クラスターのメンバシップ	71
5.3 抽出されたクラスターの大きさ	72
5.4 非ファジー化後の抽出されたクラスター	72
5.5 Landsat 航空画像	74
5.6 図5.5(a)の画像のヒストグラム	74
5.7 抽出されたクラスターの大きさ	75
5.8 5回の抽出後のヒストグラム	75
5.9 カラー画像からの肌色領域の抽出	77
6.1 アフィン正規化変換	81
6.2 提案法(a)とShapiro-Brady法(b)による対応	83
6.3 3つの平面物体の様々な視点からの画像	85
6.4 図6.3の画像から代表パターンまでの距離	86

6.5	小さな歪みを含む相似変換の下での3つの平面物体	89
6.6	図6.5の画像から代表パターンまでの距離	90

表目次

表目次

2.1 マガーク錯覚の例	21
------------------------	----



Inches 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

© Kodak, 2007 TM: Kodak



Kodak Gray Scale



© Kodak, 2007 TM: Kodak

A 1 2 3 4 5 6 **M** 8 9 10 11 12 13 14 15 **B** 17 18 19

