

統計的画像処理手法

栗田多喜夫

脳神経情報研究部門

産業技術総合研究所

E-mail: takio-kurita@aist.go.jp

概要

統計的手法は、画像処理でも基本的な道具として、さまざまな場面で利用されている。本稿では、基本的な統計的手法、特に、多変量データ解析手法および最近画像処理でも盛んに使われるようになってきたロバスト統計、EM アルゴリズムやモデル選択手法などの統計手法の考え方と画像処理への利用方法について、具体的な応用例を示しながら概説する。

1 はじめに

統計的手法は、画像処理でも基本的な道具として、さまざまな場面で利用されている。本稿では、基本的な統計的手法の考え方と画像処理への利用方法について概説する。特に、重回帰分析、主成分分析、判別分析などの基本的な多変量データ解析手法および最近画像処理でも盛んに使われるようになってきたロバスト統計、EM アルゴリズムやモデル選択手法などの統計手法について具体的な応用例を示しながら説明する。

2 統計量

まず、最初に、基本的な統計量について復習しておこう。

2.1 平均と分散

ある確率分布に従う標本列 $\{x_i | i = 1, \dots, N\}$ が与えられた場合、(標本) 平均 (mean) \bar{x} および (標本) 分散 (variance) σ^2 は、

$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i \\ \sigma^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2\end{aligned}\tag{1}$$

で定義される。これらは、標本列から計算される代表的な統計量であり、画像処理でも多くの応用で利用される基本的な量である。(標本) 分散の平方根 $\sigma = \sqrt{\sigma^2}$ は、標準偏差 (standard deviation) と呼ばれている。標本が多変量 $\{\mathbf{x}_i | i = 1, \dots, N\}$ の場合には、平均ベクトルおよび分散共分散行列 Σ は、それぞれ

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$$

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (2)$$

で定義される。ここで、 \mathbf{x}^T は \mathbf{x} の転置を表す。

2.2 相関係数

次に、2つの標本列 $\{x_i | i = 1, \dots, N\}$ および $\{y_i | i = 1, \dots, N\}$ が与えられた場合、これらの標本列の類似度を測る統計量として相関係数 ρ がある。それは、

$$\rho = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \quad (3)$$

で定義される。ただし、 σ_x および σ_y は、それぞれ、 $\{x_i | i = 1, \dots, N\}$ および $\{y_i | i = 1, \dots, N\}$ の標準偏差である。相関係数は、各標本列から平均を引いた値を要素とするベクトル $\tilde{\mathbf{x}} = (x_1 - \bar{x}, \dots, x_N - \bar{x})^T$ および $\tilde{\mathbf{y}} = (y_1 - \bar{y}, \dots, y_N - \bar{y})^T$ を考えて、それらのベクトルのなす角を θ とすると、

$$\rho = \cos \theta = \frac{\tilde{\mathbf{x}}^T \tilde{\mathbf{y}}}{|\tilde{\mathbf{x}}| |\tilde{\mathbf{y}}|} \quad (4)$$

となり、 -1 から 1 の範囲の値をとる。相関係数は、テンプレートマッチングやステレオやオプティカルフローの推定での対応点の探索のための評価関数として用いることができる。

2.3 自己相関関数

平均が 0 の定常時系列データ $\{x(t) | t = 0, \dots, N-1\}$ が与えられた場合、自己相関関数は、

$$r(l) = \frac{1}{N} \sum_{t=0}^{N-l-1} x(t+l)x(t) \quad (5)$$

で定義される。自己相関関数のフーリエ変換は、ピリオドグラム (パワースペクトル密度関数) と呼ばれ、

$$p(f) = \sum_{l=-(N-1)}^{N-1} \exp(-i2\pi fl) r(l) \quad (6)$$

のように定義され、時系列信号のフーリエ変換 $X(f)$ と

$$p(f) = \frac{1}{N} |X(f)|^2 \quad (7)$$

のような関係が成り立つ。

2.3.1 画像認識のための高次局所自己相関特徴

上述のように自己相関関数はパワースペクトルと密接な関係があり、定常時系列データの有効な特徴の一つである。大津等 [53] は、画像の認識や計測のために有効な基本的な画像特徴として、自己相関関数を拡張した高次局所自己相関特徴を提案し、それらの特徴を多変量データ解析手法を用いて統合して有効な特徴抽出する画像計測・認識手法を提案した。具体的な応用例については多変量データ解析手法の説明のところで述べるが、ここでは、その準備として高次局所自己相関特徴について説明しておく。

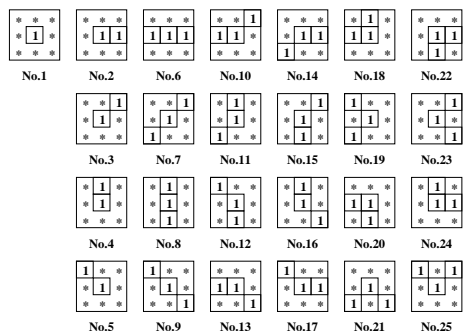


図 1: 高次局所自己相関特徴のための局所パターン

自己相関関数の高次への拡張は、高次自己相関関数と呼ばれている [45]。参照点 r での対象画像の輝度値を $I(r)$ とすと、 N 次自己相関関数は、参照点周りの N 個の変位 (a_1, \dots, a_N) に対して、

$$x(a_1, \dots, a_N) = \int I(r)I(r + a_1) \cdots I(r + a_N)dr \quad (8)$$

で定義される。高次自己相関関数は、次数 N や変位 (a_1, \dots, a_N) の取り方により無数に定義できるが、画像データでは、一般に、近くの画素間の局所的な相関の方が重要であると考えられるので、次数 N を高々 2 までとし、変位を参照点 r の周りの局所的な 3×3 画素の領域に限定する。つまり、局所的な領域内での 3 点までの相関関係を特徴とする。これにより、平行移動により等価な特徴を除くと、2 値画像に対して、特徴の数は全部で 25 個になる。図 1 に 25 個の局所パターンを示す。濃淡画像に対しても、同様に平行移動により等価な特徴を除くと 35 個の特徴が得られる。各特徴の計算は、局所パターンの対応する画素の輝度値の積を全画像に対して足し合わせればよい。こうして計算された特徴は、明らかに、対象の位置に関して不変となる。また、画像中に複数の対象があった場合、画像全体に対する特徴は、各対象の特徴の和になる。

高次局所自己相関に基づく特徴は、近傍の画素の濃淡値の積を画像全体に対して足し合わせて得られる非常に局所的な特徴であるので、高解像度の画像から抽出された特徴は顔の識別のためには細か過ぎるかもしれない。つまり、認識対象に応じた最適な解像度があると考えられる。栗田等 [34] は、認識に最適な解像度を探索する代わりに、画像ピラミッド [7] の各画像から抽出した特徴をすべて利用することを提案している。画像ピラミッドは、高解像度の画像から低解像度の画像までのいくつかの異なる解像度の画像の集合として構成されるので、画像ピラミッドの各画像から高次局所自己相関に基づく特徴を抽出すると、それらの特徴の集合には対象の詳細な情報から大まかな情報まで含まれている。しかも、対象の平行移動に関する不変性は、これらの特徴にも引き継がれる。

3 回帰分析 (最小 2 乗法)

回帰分析あるいは最小 2 乗法は、ある変数 (説明変数) とその変数に対する望みの結果 (目的変数、教師信号) が学習データとして与えられた時、説明変数から目的変数を予測するモデルを推定する統計手法であり、最も基本的で、最も広く用いられている手法である。

画像処理の分野では、ともすれば個々の課題を解決するために直観的でアドホックな手法を提案し、それらの手法の評価も人が処理結果を見て判断するなどのように主観的で不確かになりがちである。したがって、新たな課題に対して既存の手法を組み合わせることで画像処理手法を設計するためには、個々の画像処理手法に関するかなりの経験とその課題に関する十分な知識が必要となる。しかし、もし、システムが一般的

でしかも適応的な構造をしており、処理結果の量的な評価のための評価基準が与えられているなら、システム自らが課題に対して自動的に最適な処理手順を設計したり、要求された変更に関して自動的に適応したりできるようになると期待できる。回帰分析はそのような自動化のための最も基本的な手法として使うことができる。

3.1 重回帰分析と最尤推定

今、学習用データとして説明変数と目的変数に関する N 個の観測値の組 $\{ \langle x_i, y_i \rangle \mid i = 1, \dots, N \}$ が与えられているとする。この時、2つの変量間の関係を説明するためのモデルとして

$$y = f(\mathbf{x}, \boldsymbol{\theta}) + \varepsilon \quad (9)$$

を考えよう。ここで、 $\boldsymbol{\theta}$ はモデルのパラメータである。重回帰分析では、学習データに対する平均2乗誤差 (Least Mean Squared Error)

$$LMS = \frac{1}{N} \sum_{i=1}^N \varepsilon_i^2 = \frac{1}{N} \sum_{i=1}^N (y_i - f(\mathbf{x}_i, \boldsymbol{\theta}))^2 \quad (10)$$

を最小とするようなパラメータを推定結果とする。このような2乗誤差を最小とするパラメータを求める推定方法は、最小2乗法と呼ばれている。特に、 $f(\mathbf{x}, \boldsymbol{\theta})$ が

$$z = f(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{w}^T \mathbf{x} + w_0 \quad (11)$$

のように線形で表される場合は、線形回帰モデルと呼ばれ、平均2乗誤差を最小とする最適なパラメータ \mathbf{w} および w_0 は、それぞれ、

$$\begin{aligned} \mathbf{w} &= \Sigma_X^{-1} \sigma_{Xy} \\ w_0 &= \bar{y} - \mathbf{w}^T \bar{\mathbf{x}} \end{aligned} \quad (12)$$

で与えられる。ここで、 Σ_X^{-1} および σ_{Xy} は、それぞれ、 x の分散共分散行列および x と y の共分散ベクトルである。

最小2乗法は、誤差の分布を正規分布と仮定した場合の最尤推定と密接な関係があることが知られている。今、誤差 ε が平均0で分散が σ の正規分布に従うと仮定すると、学習データにモデルをあてはめた時の誤差の尤度は、

$$L = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\varepsilon_i^2}{2\sigma^2}\right\} \quad (13)$$

となる。従って、その対数 (対数尤度) は、

$$l = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N \varepsilon_i^2 \quad (14)$$

となる。これを最大とするようなパラメータを求めることは、第2項の平均2乗誤差を最小とすることと等価であるので、この場合には、最尤推定と最小2乗法は同じものとなる。

3.1.1 重回帰分析による画像修復

画像の鮮鋭化・平滑化、エッジ検出、画像の修復などの画像処理の多くの問題は、もとの画像から望みの結果を出力する写像を推定する問題として定式化することができる。もし部分的にでも望みの結果があ



図 2: 画像の修復の例 . (a) 劣化画像 [左上] . (b) 学習領域 [右上] . (c) 理想画像 (学習領域の 2 値化) [右 2 番目] . (d) 学習重み (エッジ強度) [右 3 番目] . (e) 修復された学習領域 [右下] . (f) 修復画像 [左下].

らかじめ得られるなら、平均 2 乗誤差最小の意味で最適な写像を重回帰分析を用いて構成することができる。これは、最も簡単で直接的なアプローチであり、適応的な画像処理システムを設計するための有望な方法であると考えられる。大津等 [52] は、画像の修復に重回帰分析を利用する方法を提案している。

画像修復では、一般には、劣化過程が完全にわかっていることは稀であるが、重回帰分析を適用して画像を復元するためには、教師信号として元の画像の一部がわかるあるいは推測できる必要があるが、劣化画像には鋭いエッジなどの明瞭な形を推測できる部分が含まれていることが多く、そのような部分から元の画像の一部を推定することはそれほど難しくない。そのような利用可能な (部分) 画像を理想画像と呼ぶ。

多くの場合、劣化過程は局所的かつ位置不変であると見なせるので、画像修復は劣化画像の近傍での局所的な演算 (逆フィルタリング) で実現できると考えられる。したがって、画像修復の問題は、劣化画像の局所近傍特徴ベクトル x から修復画像の画素の値 z への写像を推定する問題として定式化できる。この写像を線形モデル

$$z = w^T x + w_0 \quad (15)$$

で近似する場合には、理想画像を目的変数としてパラメータ w および w_0 を重回帰分析により推定すればよい。

図 2 に、重回帰分析を用いた逆フィルタの学習による画像修復の例を示す。劣化画像は、シフトや回転を含む複雑な位置依存のボケにもかかわらず、局所的な改善がかなり見られている。

3.1.2 微分オペレータの設計

エッジ抽出やオプティカルフローの推定などでは、画像から局所的な空間微分を推定したいことがある。一般には、微分値は、予め与えられた 3×3 画素の微分オペレータと局所領域内の画素値との内積を計算して求められることが多いが、局所領域の大きさを大きくしたい場合や局所領域の中心部分により大きな重みを与えたい場合などには、微分オペレータを設計し直す必要がある。重回帰分析を用いて局所領域内で局面を当てはめると、比較的簡単に微分オペレータを設計することができる。

例えば、2 次微分まで求めたい場合には、当てはめる局面は 2 次微分可能である必要がある。また、解析が容易で、しかも、雑音などの影響を受けにくいものがよい。従って、局面の候補としては、例えば、2 次多項式

$$z(x, y) = w_0 + w_1x + w_2y + w_3x^2 + w_4xy + w_5y^2 \quad (16)$$

が考えられる。このとき、この曲線のパラメータが求めれば、微分値は、

$$\begin{aligned}
 \left. \frac{\partial z}{\partial x} \right|_{x=0, y=0} &= w_1 \\
 \left. \frac{\partial z}{\partial y} \right|_{x=0, y=0} &= w_2 \\
 \left. \frac{\partial^2 z}{\partial^2 x} \right|_{x=0, y=0} &= w_3 \\
 \left. \frac{\partial^2 z}{\partial x \partial y} \right|_{x=0, y=0} &= w_4 \\
 \left. \frac{\partial^2 z}{\partial^2 y} \right|_{x=0, y=0} &= w_5
 \end{aligned}
 \tag{17}$$

により推定できる。最小2乗法を用いて、注目点を原点とする局所領域内の座標値から画素値を推定する局面(モデル)を最小2乗法により推定し、各点の座標値とパラメータの関係を調べれば、微分オペレータが求まる。重みを考慮したい場合には、重み付き最小2乗法により曲線を当てはめればよい。

例えば、 3×3 の局所領域内で2次多項式を当てはめて設計した微分オペレータは、

$$\begin{aligned}
 w_1 &: \frac{1}{6} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \\
 w_2 &: \frac{1}{6} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix} \\
 w_3 &: \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{pmatrix} \\
 w_4 &: \frac{1}{6} \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} \\
 w_5 &: \frac{1}{6} \begin{pmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{pmatrix}
 \end{aligned}
 \tag{18}$$

のようになる。

横矢等 [82, 83] では、 5×5 の局所領域で2次多項式を当てはめることにより微分オペレータを設計し、それらの微分オペレータを利用して推定した微分値から距離画像の局所的な曲率を計算し、それを領域分割に利用している。

3.1.3 勾配法によるオプティカルフローの推定

動画画像から対象の動きのパラメータを時空間微分から推定するための方法として勾配法がある。この方法では、“物体上の点の明るさは移動後も変化しない”という仮定から時空間微分とオプティカルフローとの関係式を導出し、それを利用して対象の動きを推定する。

今、時刻 t における画像上の点 $r = (x, y)$ の輝度値を $I(x, y, t)$ とし、微小時間 Δt 後に対象が $\mathbf{u} = (\Delta x, \Delta y)^T$ だけ移動したとする。このとき、点の明るさが移動後も変わらないとすると、

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (19)$$

が成り立つ。今、右辺を Taylor 展開すると、

$$I(x, y, t) = I(x, y, t) + \Delta x \frac{\partial I}{\partial x} + \Delta y \frac{\partial I}{\partial y} + \Delta t \frac{\partial I}{\partial t} + e \quad (20)$$

となる。ここで、 e は、 $\Delta x, \Delta y, \Delta t$ に関する高次の項であるが、これを無視し、両辺を Δt で割り、 $\Delta t \rightarrow 0$ とすると、

$$(\nabla I)^T \mathbf{u} + I_t = 0 \quad (21)$$

という関係式が得られる。ただし、 $\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})^T$ および $I_t = \frac{\partial I}{\partial t}$ である。

もし、局所領域内の各点のオブティカルフロー \mathbf{u} が等しいと仮定できるなら、局所領域内で上記の関係式が最小 2 乗誤差の意味で最も良く当てはまるような \mathbf{u} を推定すればよい。つまり、局所領域 R で 2 乗誤差

$$E = \sum_{\mathbf{r} \in R} ((\nabla I)^T \mathbf{u} + I_t)^2 \quad (22)$$

を最小とする \mathbf{u} を重回帰分析を用いて求めればよい。

3.1.4 高次局所自己相関特徴と重回帰分析による画像計測

2.3.1 節で示した高次局所自己相関特徴は、画像枠内の対象の位置が変わっても変化しない特徴であり、しかも、画像内の対象に関して加法性を満たす基本的な画像特徴である。これらの性質は画像計測にとっても好ましいものである。画像から抽出した高次局所自己相関特徴と重回帰分析による例からの学習を組み合わせて使うことにより、例となる画像とそれに対する望みの計測結果を例示することにより、システムが自動的に計測課題に有効な推定方法を学習できるようになると期待できる。

今、学習に用いる対象画像の集合を $\{f_i(\mathbf{r}) | i = 1, \dots, N\}$ 、画像 f_i に対する望みの計測結果を y_i 、特徴ベクトルを x_i とすると、重回帰分析は、望みの計測結果 y_i と推定値 $z_i = A^T x_i + b$ の平均 2 乗誤差が最も小さくなるようにパラメータ A および b を決定する手法である。

各計測課題に対する最適なパラメータが求めれば、システムは新たな入力画像に対する計測値の推定結果を高速に算出できるようになる。

望みの計測結果 y の与え方は、計測課題に応じていろいろ考えられる。たとえば、面積や対象の数のような計測の場合には、 y は 1 次元の数値になり、異なるふたつの対象の個数の同時計測の場合には、 z は各対象の個数を含む 2 次元のベクトルとなる。

新たに学習データを増やしたい場合や、学習データが一度に少しずつしか得られないような場合には、逐次的な学習が必要となるが、線形重回帰による学習の場合には、すべてのデータが一度に与えられた場合と結果的に同じ係数行列を計算するための逐次学習方式が存在する [31]。

具体的な応用例として、たとえば、画面内の 2 種類の直径の異なる粒子の個数を同時に計測する課題を学習させてみた。学習用データとしては無作為に生成した 2 種類の粒子を含む画像を用いた。40 枚の画像を学習データとして、重回帰分析により最適なパラメータを計算した。この場合、教師信号としては、 $y = (\text{大きい粒子の個数}, \text{小さい粒子の個数})^T$ を用いた。図 3 にテスト用の画像の例を示す。学習したパラメータを用いて粒子数を推定すると、たとえば、画像 (a) に対して $z = (4.10, 5.88)^T$ 、画像 (b) に対して $z = (3.09, 0.888)^T$ 、画像 (c) に対して $z = (1.97, 3.02)^T$ であった。量子化ノイズにもかかわらず、ある程度の推定結果が得られている。

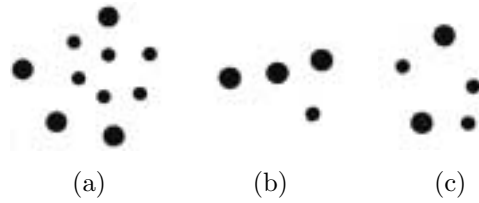


図 3: 大小 2 種類の粒子の同時計測

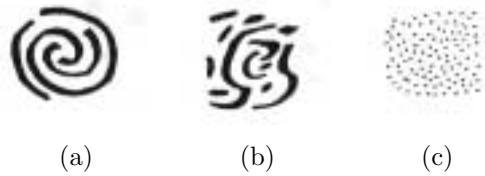


図 4: 孤立対象の個数の計測

学習用のデータを変更することにより、同様の手法を別の課題に適用することができる。そこで、次に、対象の形によらない位相的特徴の計測課題を学習させた。ここでは、画面内の分離した孤立対象の個数を推定させた (図 4)。これらの個数は対象の形に全く無関係である。さまざまな形の分離対象を含む 48 枚の画像を学習データとして、重回帰分析による学習を行った。教師信号は孤立対象の個数とした。その結果、任意に与えた画像に対して正しく分離対象の個数を推定できるようになった。

興味深いことに、結果としてシステムが学習したパラメータが、位相数学におけるオイラーの公式と関係していることが証明できた。ここで重要なのは、オイラーの公式をプログラムとしてシステムに教えたのではなく、学習例からシステムが自動的に学んだ点である。また、この場合には、平行移動だけでなく回転に対しても不変となっていて、画像を画面内でどのように置いても正しく計測できるようになった。

3.2 自己回帰モデル

自己回帰モデル (Autoregressive model) は、時系列信号の解析によく用いられる回帰分析手法であり、音声認識や制御等の信号処理のための基本的な道具の一つである。画像処理やコンピュータビジョンでもテクスチャの特徴抽出、動画像の処理、輪郭点列の処理などのようにデータに何らかの順序関係がある場合の信号処理手法として利用できる。

今、 $\{x(t)|t=0, \dots, N-1\}$ を時系列データとすると、 m 次の前向き自己回帰モデルは、 m 個前までのデータから現在の信号 $x(l)$ を線形予測するモデルであり、

$$x(l) = \sum_{i=1}^m a^m(i)x(l-i) + \epsilon_f^m(l) \quad (23)$$

のように定義される。ここで、 $\{a^m(i)\}_{i=1}^m$ および $\epsilon_f^m(l)$ は、それぞれ、前向き自己回帰係数および前向き予測誤差である。同様に、 m 次の後向き自己回帰モデルは、

$$x(l-m-1) = \sum_{i=1}^m b^m(i)x(l-i) + \epsilon_b^m(l) \quad (24)$$

のように定義される。ここで、 $\{b^m(i)\}_{i=1}^m$ および $\epsilon_b^m(l)$ は、それぞれ、後向き自己回帰係数および後向き

予測誤差である。

前向き自己回帰係数 $\{a^m(i)\}_{i=1}^m$ は、実際の信号 $x(l)$ と自己回帰モデルによる推定値 $z(l) = \sum_{i=1}^m a^m(i)x(l-i)$ との平均 2 乗誤差

$$\varepsilon^2 = \frac{1}{N} \sum_{l=m}^{N-1} (x(l) - z(l))^2 \quad (25)$$

が最小となるように求められる。具体的には、前向き自己回帰係数を要素とするベクトルを $\mathbf{a} = (a_1, \dots, a_m)^T$ とすると、最適な係数は、

$$\mathbf{a} = R^{-1}\mathbf{r} \quad (26)$$

となる。ただし、 R および \mathbf{r} は、それぞれ、自己相関関数を要素とする行列およびベクトルであり、

$$R = \begin{bmatrix} r(0), & r(1), & \dots, & r(m-1) \\ r(1), & r(0), & \dots, & r(m-2) \\ r(2), & r(1), & \dots, & r(m-3) \\ \dots & \dots & \dots & \dots \\ r(m-1), & r(m-2), & \dots, & r(0) \end{bmatrix}$$

$$\mathbf{r} = (r(1), r(2), \dots, r(m))^T \quad (27)$$

のように定義される。後向き自己回帰係数 $\{b^m(i)\}_{i=1}^m$ についても同様である。

PARCOR(偏自己相関) 係数は音響特性と密接な関係があり、声道断面積関数の予測を与えるなどの優れた利点があることが知られており、音声信号処理で良く利用されている。 m 次の PARCOR 係数 P^m は、 $m-1$ 次の前向き予測誤差 $\varepsilon_f^{m-1}(i)$ と $m-1$ 次の後向き予測誤差 $\varepsilon_b^{m-1}(i)$ の相関係数として定義される。つまり、

$$P^m = \frac{\sum_{i=m}^{t-1} \varepsilon_f^{m-1}(i)\varepsilon_b^{m-1}(i)}{\sqrt{\sum_{i=m}^{t-1} \{\varepsilon_f^{m-1}(i)\}^2 \{\varepsilon_b^{m-1}(i)\}^2}}. \quad (28)$$

である。したがって、PARCOR 係数は線形予測係数と違ってモデルの次数には依存しない量となる。

PARCOR 係数は、自己回帰係数と密接な関係があり、 m 次の PARCOR 係数は、実は、 m 次の自己回帰モデルを当てはめたときの m 次の自己回帰係数 $a^m(m)$ あるいは $b^m(m)$ に一致する。

自己回帰係数や PARCOR 係数は、 R の逆行列を計算することにより求めることができるが、1 次のモデルからはじめて次数を次第に大きくする再帰的なアルゴリズム (Levinson-Durbin のアルゴリズム)[14] を用いることにより高速に計算することができる。このアルゴリズムでは、1 次から m 次までのすべての自己回帰係数と PARCOR 係数を $O(m^2)$ の計算量で求めることができる。

3.2.1 ジェスチャー認識のための PARCOR 画像

栗田等 [39, 40, 41] は、動画画像から動きの情報を抽出するために自己回帰モデルを利用する方法を提案し、それをジェスチャー認識に利用した。

画像は、画素値の集合 $\{I_{(x,y)} | 0 \leq x < W, 0 \leq y < H\}$ として表現することができる。動画画像は画像の時系列であるが、これは、画素値の時系列の集合 $\{I_{(x,y)}(t) | 0 \leq x < W, 0 \leq y < H, 0 \leq t < L\}$ と見做すことができる。そこで、各画素値の時系列に対して自己回帰モデルを当てはめ、それから計算される PARCOR 係数を画像として再構成すると、それは動画画像から動きに関する情報を抽出した画像となる。この画像を PARCOR 画像と呼び、その画像から高次局所自己相関特徴を抽出し、ジェスチャー認識のための特徴を構成した。背景などの時間的変化のない部分では、PARCOR 係数は 0 となるため、背景に影響を受けない特徴が得られる。また、自己相関関数を過去の推定値 $r(l)^{(t)}$ から、忘却率 $0 < \alpha \leq 1$ で

$$r(l)^{(t+1)} \leftarrow (1 - \alpha)r(l)^{(t)} + \alpha x(t+1)x(t+1-l). \quad (29)$$



(a) 動画像.



(b) 1 次の PARCOR 画像

図 5: ジェスチャー動画像と 1 次の PARCOR 画像.

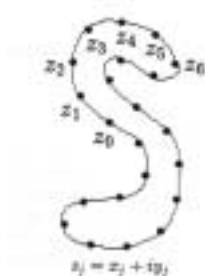


図 6: 輪郭列の複素数表現

のように再帰的に更新し、PARCOR 係数の計算のための高速アルゴリズムを利用することにより、実時間での PARCOR 画像の推定が可能となる。

図 5 にジェスチャー動画像と 1 次の PARCOR 画像の例をに示す。

3.2.2 形の計測 (複素自己回帰モデル)

平面図形 (2 値画像) の認識は、パターン認識における最も基本的な課題の一つである。「形 (shape)」は、本来、パターンの相似変換に不変な概念であり、形の本質的な情報を担う外形 (輪郭線) を相似変換に不変な形で記述することは、認識や分類のために重要である。関田等 [66, 67, 33, 37] は、輪郭点の座標値を複素数で表現し、輪郭点列を複素数値列とみなして複素自己回帰モデルを当てはめ、相似変換に不変な特徴量を抽出する方法を提案し、平面図形の認識や分類に適用した。

平面図形の輪郭線 (平面曲線) を追跡して得られる点列を (x_j, y_j) ($j = 0, 1, \dots, N - 1$) とし、その複素表現を $z(j) = x_j + iy_j$ とする。このとき、 m 次の複素自己回帰モデルは、実数値列に対する自己回帰モデルと同様に、

$$z(l) = \sum_{i=1}^m a^m(k) z(l-i) + \epsilon_f^m(l) \quad (30)$$

のように輪郭点 m 個前までの輪郭点の線形結合で近似するモデルとして定義される。ここで、複素自己回帰係数 $\{a^m(k)\}_{k=1}^m$ は、やはり、平均 2 乗予測誤差が最小となるように決められる。

輪郭点列が複素表現されている場合、原点まわりの回転は各輪郭点を $e^{i\theta}$ 倍することに対応する。回転された輪郭点列 $\{e^{i\theta} z(j)\}_{j=0}^{N-1}$ に対して複素自己回帰モデルを当てはめる場合には、(30) 式の両辺が単に

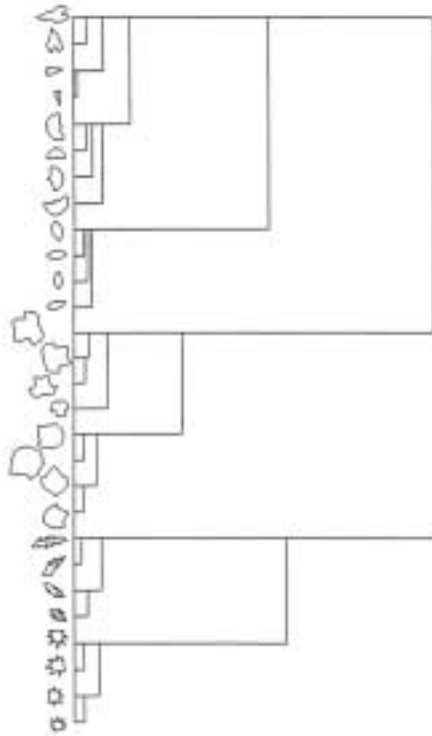


図 7: 相似変換に不変な形の自動分類

$e^{i\theta}$ 倍されるだけで、複素自己回帰係数には影響が及ばない。従って、複素自己回帰係数は原点まわりの回転に不変となる。また、複素自己回帰係数の計算は、輪郭を追跡する際の始点位置の選び方にも依存しない。これらの性質は、形の記述、表現という意味で好ましいものである。さらに、興味深い性質として輪郭線の追跡方向 (時計回りか反時計回り) を変えると、得られる複素自己回帰係数が複素共役となる。複素 PARCOR 係数は、 m 次の複素自己回帰モデルの m 次の複素自己回帰係数 $a^m(m)$ と等しいので、複素自己回帰係数について成立する性質は同時に複素 PARCOR 係数でも成り立つ。

これをさらに、平行移動に不変にするためには、重心を原点にとり輪郭点列を表現する方法や隣り合う輪郭点の差分に対して複素自己回帰モデルをあてはめる方法が考えられる。また、大小伸縮に不変にするためには、輪郭の全周長を N 等分する区間に分割し、各区間内のデータ点を平均値で代表させ、輪郭点 $z(j)$ とすればよい (図 6)。大きさの違う輪郭点列は $\{\rho z(j)\}$ と表現されるので、これらも、式 (30) の両辺で打ち消しあい、結局、複素自己回帰係数は大小伸縮に対しても不変となる。

形の認識のための特徴としては、複素自己回帰係数や複素 PARCOR 係数などを用いればよいが、形の自動分類のためには、形間の距離を定義する必要がある。複素自己回帰モデルを利用すると、輪郭点列 (形) 間の相似変換に不変な距離を定義することができる [33, 37]。図 7 に、複素パワーケプストラムに基づく形状間の距離を用いて、木の葉を自動分類した結果を示す。

4 主成分分析

主成分分析 (Principal Component Analysis) は、多変量の計測値から変量間の相関を無くし、しかもより低次元の変量によって元の計測値の特性を記述するための多変量データ解析手法であり、画像圧縮等の

情報圧縮やパターン認識のための特徴抽出等の多くの応用で利用されている。

4.1 主成分分析と主成分スコア

今、多変量の計測値を $\{x_i | i = 1, \dots, N\}$ とし、各データから平均ベクトル \bar{x} を引いたベクトルを \tilde{x} で表すとす。このとき、主成分分析の特徴ベクトル (主成分スコア) y は、

$$y = U^T \tilde{x} = U^T (x - \bar{x}) \quad (31)$$

のように直交行列 U で定義される直交変換により計算される。したがって、 \tilde{x} の U の各列を正規直交基底とする部分空間への射影は、

$$\hat{x}_i = UU^T \tilde{x}_i \quad (32)$$

となり、元の計測値ベクトル \tilde{x}_i の近似となる。主成分スコア y によって、平均 2 乗誤差の意味で元の計測値の情報 (分散) を最大限抽出するような直交行列 U を求めるには、元の \tilde{x}_i とその近似 \hat{x}_i との平均 2 乗誤差

$$\varepsilon^2(U) = \frac{1}{N} \sum_{i=1}^N |\tilde{x}_i - \hat{x}_i|^2 \quad (33)$$

を最小とする正規直交基底 $U = [u_1, \dots, u_L]$ を求めれば良い。最適な直交行列 U は、 X の分散共分散行列 $\Sigma_X = \frac{1}{N} \sum_{i=1}^N \tilde{x}_i \tilde{x}_i^T$ の固有値問題

$$\Sigma_X U = U \Lambda, \quad (U^T U = I) \quad (34)$$

の解として求まる [51, 13]。ただし、 Λ は固有値行列である。また、 U としては、固有値の大きさの順番に対応する固有ベクトルを L 個まで取るものとする。

この時、主成分スコア y の平均および分散共分散行列は、それぞれ、

$$\begin{aligned} \bar{y} &= \mathbf{0} \\ \Sigma_Y &= U^T \Sigma_X U = U^T U \Lambda = \Lambda \end{aligned} \quad (35)$$

となり、変量間の相関が無くなるのがわかる。

4.1.1 直線の当てはめ

実際の画像処理や画像理解においては、画像から必要な情報を取り出すために、エッジ抽出等によって得られた平面上の点の集合に直線や曲線を当てはめる等の方法が取られることが多い。

今、平面上のデータ点の集合を $S = \{(x_i, y_i)\}_{i=1}^N$ とする。ここで、各点に付けられた添字の順番に意味はないものとする。

多変量データ解析では、このようなデータ点集合に直線 $y = ax + b$ を当てはめる場合、重回帰分析を用いるのが普通である。つまり、データ点から直線への y 軸に平行な誤差の 2 乗の平均、すなわち、

$$\frac{1}{N} \sum_{i=1}^N (y_i - ax_i + b)^2 \quad (36)$$

が最も小さくなるように直線を当てはめる (図 8(a))。しかし、この方法は、変量 x の線形式で変量 y を予測することを意図している。したがって、予測の方向を入れかえて y から x を予測する場合の回帰直線は、 x から y を予測する場合と同じにはならない。しかし、画像処理への応用では、 x と y とは、対等に

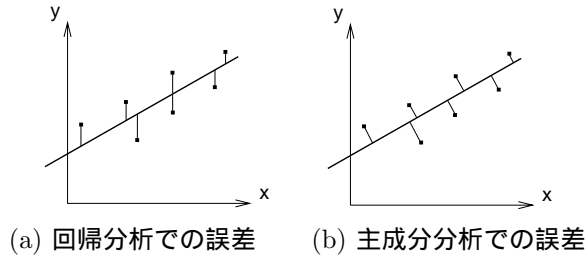


図 8: 誤差の測り方

扱わなければならないことが多く、こうした方向性は好ましくないものと考えられる。主成分分析を用いると x と y を対等に扱い、図 8(b) のように誤差をデータ点から直線までの最短距離として評価することができる。

一般に、平面上の直線は、方向ベクトル $\mathbf{a} = (D_x, D_y)^T$ (ただし、 $\mathbf{a}^T \mathbf{a} = 1$) を用いて、

$$\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{a}t \quad (37)$$

のように表せる。ここで、 $\mathbf{r}(t) = (x(t), y(t))^T$ および $\mathbf{r}_0 = (x_0, y_0)^T$ である。

このとき、データ点 $\mathbf{r}_i = (x_i, y_i)^T$ からこの直線上の点 $\mathbf{r}(t)$ までの 2 乗距離は

$$d_i^2(t; \mathbf{a}, \mathbf{r}_0) = t^2 + 2\mathbf{a}^T(\mathbf{r}_0 - \mathbf{r}_i)t + (\mathbf{r}_0 - \mathbf{r}_i)^2 \quad (38)$$

で与えられる。直線を固定すると、この距離が最小となるのは、データ点から直線に垂直な直線を下ろした場合であり、データ点からこの直線までの最短距離の 2 乗は、

$$d_i^2(\mathbf{a}, \mathbf{r}_0) = (\mathbf{r}_i - \mathbf{r}_0)^2 - \mathbf{a}^T(\mathbf{r}_i - \mathbf{r}_0)(\mathbf{r}_i - \mathbf{r}_0)^T \mathbf{a} \quad (39)$$

となる。したがって、このようにして測ったデータ点の集合から直線への平均 2 乗誤差は、

$$\varepsilon^2(\mathbf{a}, \mathbf{r}_0) = \frac{1}{N} \sum_{i=1}^N d_i^2(\mathbf{a}, \mathbf{r}_0) \quad (40)$$

と表される。方向ベクトルに関する制約条件のもとで、この平均 2 乗誤差を最小とする直線のパラメータ \mathbf{a} および \mathbf{r}_0 を求める問題は、ラグランジュ乗数を λ として、評価関数

$$J(\mathbf{a}, \mathbf{r}_0) = \varepsilon^2(\mathbf{a}, \mathbf{r}_0) + \lambda(\mathbf{a}^T \mathbf{a} - 1) \quad (41)$$

を最小化する問題に帰着される。最適なパラメータ \mathbf{a} および \mathbf{r}_0 は、

$$\begin{aligned} \mathbf{r}_0 &= \boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^N \mathbf{r}_i \\ C\mathbf{a} &= \lambda\mathbf{a} \end{aligned} \quad (42)$$

となる。ただし、

$$C = \frac{1}{N} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu})(\mathbf{r}_i - \boldsymbol{\mu})^T \quad (43)$$

はデータ点集合の分散共分散行列である。つまり、式 (42) は、主成分分析の固有値問題と全く同じとなる。

このとき達成される最小 2 乗誤差は、

$$\varepsilon^2 = \text{tr } C - \lambda \quad (44)$$

となる。

4.1.2 固有顔による顔画像の認識

顔画像の認識でまず考えられるのは、目・鼻・口などの顔を構成する部品の形状やそれらの配置の個人差に着目し、これらから特徴点を抽出して認識に利用する方法である。しかし、顔画像からこれらの部品を精度よく抽出することはかなり難しい。また、各部品がうまく抽出できたとしても、類似した形状の差を認識に利用することはそれほど容易ではない。そこで、このような顔に特有の知識を用いるのではなく、顔画像そのものをパターンとして扱い、統計的パターン認識手法を適用する方向の研究が活発に行われている。

最も簡単なパターン認識手法は、パターン間のマッチングに基づく方法であるが、画像そのものをパターンとして扱った場合には、パターンの次元が膨大になってしまう。そのため、パターンを情報圧縮した後でマッチングを行う方法がいくつか提案されている。パターンを情報圧縮すれば、入力条件の変動に対して頑健な認識結果が得られると期待できる。Turk 等 [77, 78] が提案した固有顔 (eigenface) による方法では、主成分分析によりパターンを情報圧縮し、顔画像の識別に利用している。手法の単純さと固有顔という名前の付け方の上手さから、顔画像の認識において最も有名な手法のひとつとなっている。

今、 N 枚の学習用の顔画像のうちの i 番目の画像を、各画素の値をならべた M 次元のベクトル x_i として表現する。また、 N 枚の画像の平均ベクトルを $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ とし、各画像から平均ベクトルを引いたベクトルを $\tilde{x}_i = x_i - \bar{x}$ で表し、各画像から平均ベクトルを引いた画像の集合を行列 $\tilde{X} = [\tilde{x}_1, \dots, \tilde{x}_N]$ で表す。

学習用の画像集合を平均 2 乗誤差の意味で最適に近似する正規直交基底 U は、主成分分析を用いて構成することができる。このとき、 U の各固有ベクトル u_i を、固有顔 (eigenface) と呼んでいる。また、ある画像 x に対する主成分スコア y の各成分は、画像 x を表現するための各固有顔の貢献度を表していると解釈できる。固有ベクトルの次元 L を小さくすることにより、固有顔 U とスコア y を用いて、もとの画像を

$$\hat{x} = UU^T \tilde{x} = Uy \quad (45)$$

のように低次元で近似して表現することが可能となる。

また、2 枚の画像 x_1 と x_2 が与えられた場合、各画像の主成分スコア間の距離は、

$$|y_1 - y_2|^2 = |U^T(\tilde{x}_1 - \tilde{x}_2)|^2 = |\hat{x}_1 - \hat{x}_2|^2 \quad (46)$$

のように、もとの画像の U の列の張る空間への射影 (もとの画像の平均 2 乗近似) 間の距離と同じになる。したがって、顔画像の識別に画像の主成分スコア間の距離を用いることは、もとの画像間の距離を U の列の張る空間で近似的に計算することに対応する。

Kirby 等 [30] は、100 枚の顔画像に対して主成分分析を行い固有空間の次元を 50 次まで取れば 95% 以上の情報が復元できることを示している。また、Turk 等 [77, 78] は、16 人の被験者から撮影した 2500 枚の顔画像のデータベースを用いた認識実験を行い、向きや大きさが一定なら 96% の認識率を得ている。Pentland 等 [55] は、3000 人から取った 7562 枚の顔画像のデータベースを用い、128 枚の代表的な顔画像から主成分分析により構成した 20 次元の固有空間を用いた識別で、95% の認識率を得ている。さらに、目・鼻・口などの部品に対しても固有空間を構成し、それらと顔全体の固有空間を組み合わせると認識により、98% の認識率を得ている。

5 判別分析

主成分分析は、情報の圧縮という観点では有効な手法であるが、各データがどのクラスに属しているかの情報は全く利用しておらず、識別のための有効な特徴を構成する手法としては必ずしもよい手法ではない。

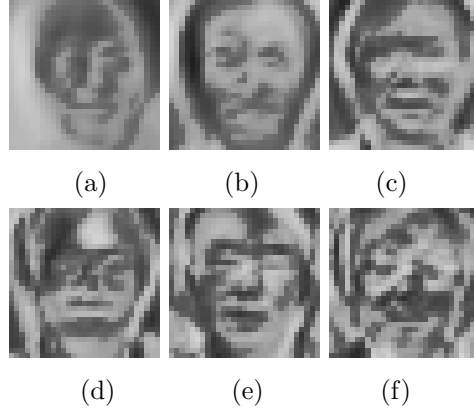


図 9: 固有顔の例 ((a) 第 1 固有ベクトル, (b) 第 6 固有ベクトル, (c) 第 11 固有ベクトル, (d) 第 16 固有ベクトル, (e) 第 21 固有ベクトル, (f) 第 26 固有ベクトル)

学習データが予め分類されており、各データがどのクラスに属しているかの情報が得られる場合には、同一クラス内のデータはなるべく近くなり、逆にクラス間のデータはなるべく離れるような線形写像を構成する必要がある。そのために、判別分析 (Discriminant Analysis) が有効である。

5.1 判別分析と判別基準

今、学習用の計測値データ (ベクトル) の集合 $X = \{x_1, \dots, x_N\}$ は、予め何らかの方法で K 個のクラスに分類されており、各計測ベクトルには K 個のクラス $C_k = \{x\}$ ($k = 1, \dots, K$) のどのクラスに属しているかの情報が与えられているとする。

このとき、計測ベクトル x のクラス内共分散行列 Σ_W およびクラス間共分散行列 Σ_B は、それぞれ、

$$\begin{aligned}\Sigma_W &= \sum_{k=1}^K \omega_k \Sigma_k \\ \Sigma_B &= \sum_{k=1}^K \omega_k (\bar{x}_k - \bar{x}_T)(\bar{x}_k - \bar{x}_T)^T\end{aligned}\quad (47)$$

となる。ただし、 ω_k , \bar{x}_k , \bar{x}_T , および Σ_k は、それぞれ、クラス C_k の先験確率、クラス C_k の平均ベクトル、全平均ベクトル、および、クラス C_k の共分散行列である。

判別分析では、計測ベクトル x に対するスコア (固有空間での表現) を、

$$y = A^T \tilde{x}\quad (48)$$

により構成する。判別スコアの空間でのクラス内共分散行列 $\hat{\Sigma}_W$ およびクラス間共分散行列 $\hat{\Sigma}_B$ は、それぞれ、

$$\begin{aligned}\hat{\Sigma}_W &= A^T \Sigma_W A \\ \hat{\Sigma}_B &= A^T \Sigma_B A\end{aligned}\quad (49)$$

となる。

係数行列 $A = [a_1, \dots, a_L]$ を決定するための判別空間の良さ (判別力) の評価には、判別基準

$$J = \text{tr}(\hat{\Sigma}_W^{-1} \hat{\Sigma}_B)\quad (50)$$

使われる。判別基準は、クラス内の分散が小さくなり、各クラスの平均ベクトルが互いに離れるような場合に大きくなる。判別基準 J を最大とする最適な係数行列 A は、 Λ を Lagrange 乗数行列とする固有値問題

$$\Sigma_B A = \Sigma_W A \Lambda, \quad A^T \Sigma_W A = I_L \quad (51)$$

の解として求まる。ここで、 $\Lambda = \text{diag}(\lambda_1 \geq \lambda_2 \geq \dots \lambda_L > 0)$ は、固有値を要素とする対角行列である。また、判別空間の次元 L は、 $L \leq \min(K-1, M)$ でおさえられる。

主成分分析の場合には、2つの計測ベクトル x_1 と x_2 に対する主成分スコア間の距離は、元のベクトル間の距離と関係があったが、判別分析の場合には、2つの計測ベクトル x_1 と x_2 に対する主成分スコア間の距離 $|y_1 - y_2|^2$ は、多クラスの分布間の平均マハラノビス汎距離 $(\tilde{x}_1 - \tilde{x}_2)^T \Sigma_W^{-1} (\tilde{x}_1 - \tilde{x}_2)$ と密接に関係していることが知られている [51]。つまり、判別分析の場合には、平均クラス内分散の逆 Σ_W^{-1} で重み付けたベクトル間の距離を近似的に計算していることに対応する。

判別分析は、主成分分析と同様に、固有顔を構成する場合にも利用でき、係数行列の各列 a_i は、主成分分析の場合と同様に、固有顔として解釈できる。Swets 等 [73] は、判別分析を用いた固有顔による顔画像認識手法を主成分分析を用いた場合と比較し、判別分析を用いた方法の優位性を示している。

5.1.1 判別基準に基づく画像の二値化

濃淡画像を対象領域と背景に分離する2値化は、画像処理あるいはパターン認識の最も基本的な手続きの一つである。

濃淡画像の画素の集合を $\{x\}$ とし、各画素の濃淡値を g とすると、しきい値 k による二値化は、各画素 x が二つのクラス C_1 と C_2 のどちらに属するかを

$$g(x) \leq k \rightarrow x \in C_1, \quad g(x) > k \rightarrow x \in C_2 \quad (52)$$

のように決定する手続きである。ここで、 C_1 および C_2 は、それぞれ背景と対象領域である。対象を背景から正確に分離するためには、適切なしきい値を選ぶ必要がある。

大津は、このようなしきい値選定の問題を教師なしの決定問題として捉え、判別分析の立場から非常に簡単で、しかも汎用性をもった自動しきい値選定法 [49, 50] を提案した。なお判別分析そのものは教師ありの (supervised な) 場合のノンパラメトリックな統計手法であるが、判別基準そのものは、教師なしの (unsupervised な) 場合へも有効に利用できる。

今、画像は L 階調の輝度レベル $S = \{1, 2, \dots, L\}$ で表現されているものとする。レベル i の輝度を持つ画素数を n_i 、全画素数を $N = n_1 + n_2 + \dots + n_L$ とすると、輝度の正規化ヒストグラムは、

$$p_i = \frac{n_i}{N} \quad \left(i \in S, p_i \geq 0, \sum_{i=1}^L p_i = 1 \right) \quad (53)$$

で与えられる。このとき、画像の平均輝度レベルおよび分散は、それぞれ

$$\begin{aligned} \mu_T &= \sum_{i=1}^L i p_i \\ \sigma_T^2 &= \sum_{i=1}^L (i - \mu_T)^2 p_i \end{aligned} \quad (54)$$

で与えられる。

今、 k をしきい値として、区間 $S_1 = [1, \dots, k]$ および $S_2 = [k+1, \dots, L]$ に属する画素をそれぞれ2クラス C_1 および C_2 に分類するものとする。また、正規化ヒストグラムに対して、次の二つの累計量

$$\begin{aligned}\omega(k) &= \sum_{i=1}^k p_i & (\omega(L) = 1) \\ \mu(k) &= \sum_{i=1}^k ip_i & (\mu(L) = \mu_T)\end{aligned}\quad (55)$$

を定義しておく。このとき、 k をしきい値としたときの各クラスの生起確率は、この累計量を用いて、

$$\begin{aligned}\omega_1 &= \sum_{i \in S_1} p_i = \omega(k) \\ \omega_2 &= \sum_{i \in S_2} p_i = 1 - \omega(k)\end{aligned}\quad (56)$$

で与えられる。また、各クラスの平均輝度は

$$\begin{aligned}\mu_1 &= \frac{\mu(k)}{\omega(k)} \\ \mu_2 &= \frac{\mu_T - \mu(k)}{1 - \omega(k)}\end{aligned}\quad (57)$$

で与えられる。これらの量の間には、しきい値 k によらず常に、

$$\omega_1 \mu_1 + \omega_2 \mu_2 = \mu_T, \quad \omega_1 + \omega_2 = 1 \quad (58)$$

のような関係が成立する。また、各クラスの輝度値の分散は、

$$\begin{aligned}\sigma_1^2 &= \sum_{i \in S_1} \frac{(i - \mu_1)^2 p_i}{\omega_1} \\ \sigma_2^2 &= \sum_{i \in S_2} \frac{(i - \mu_2)^2 p_i}{\omega_2}\end{aligned}\quad (59)$$

で与えられる。

画像の輝度値をしきい値 k により二つのクラスに分けたとき、クラスの間隔がよければそのしきい値はよいしきい値であることが期待できる。そこで、しきい値 k のよさを評価するために、判別基準

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad \kappa = \frac{\sigma_T^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (60)$$

を用いる。ここで、

$$\sigma_W^2 = \omega_1 \sigma_1^2 + \omega_2 \sigma_2^2 \quad (61)$$

$$\sigma_B^2 = \omega_1 \omega_2 (\mu_1 - \mu_2)^2 \quad (62)$$

は、それぞれ、クラス内分散およびクラス間分散である。クラス内分散、クラス間分散、および全分散の間には、

$$\sigma_W^2 + \sigma_B^2 = \sigma_T^2 \quad (63)$$

の基本的関係が常に成立する。3種類の判別基準 λ 、 κ 、および η は、すべて同値となるが、 η は、1次の統計量のみから計算できるので、最適なしきい値 k を探索するには最も簡単な評価基準となる。これは、 σ_B^2 を最大とする k を最適しきい値とすることと同値となる。

5.1.2 高次局所自己相関特徴と判別分析を用いた顔画像認識

画像集合を直接主成分分析や判別分析して固有顔を構成して、顔画像の認識に利用する方法を示したが、このような方法では、画像中の顔の位置がちょっとでも変化するとパターンが大きく変化するので、画像の位置や大きさなどを予め正規化しておく必要がある。大津等 [53, 34] は、これを改善するために、画像から対象の位置や大きさの変化に頑健な特徴 (高次局所自己相関特徴) を抽出し、その特徴の基づいて識別する方法を提案している。高次局所自己相関特徴に基づく方法は、対象の位置の変化に頑健で、しかも、非常に簡単な計算で特徴を計算できるので、実時間で顔画像認識が必要なヒューマンインタフェースなどにも利用されている [26]。

栗田等 [34] は、高次局所自己相関特徴と判別分析を組み合わせた方法を顔画像の識別に適用し、50 人から撮影した約 1000 枚の顔画像の識別実験において 92.2% の認識率を得ている。また、画像ピラミッドの各画像から特徴を抽出することにより、認識率が向上することを示している。Goudail 等 [22] は、画像ピラミッドの各解像度の画像から抽出した高次局所自己相関特徴で独立に判別分析を行い、さらにそれらを統合して識別するなどの工夫により、116 人から撮られた 11600 枚の顔画像の識別において、99.9% の認識率を達成している。大きさの変化に対しても頑健にする試みとしては、画像ピラミッドに基づく方法 [32] や抽出された特徴の組み合わせ方を工夫することにより大きさに不変な特徴を構成する方法 [1] などがある。また、栗田等 [38] は、能動視覚システムを利用した認識を目指して、Log-Polar 画像から高次局所自己相関特徴を抽出することにより、大きさの変化に強くすることを試みている。

6 正準相関分析

主成分分析では、1つの計測ベクトルから変量間の相関を無くし、しかも元の計測ベクトルの持つ情報 (分散) をなるべく抽出するような低次元の変量が構成される。しかし、2組の計測ベクトルが与えられ、それらの間の相関関係を分析したいような場合には、正準相関分析 (Canonical Correlation Analysis) を利用する必要がある。正準相関分析では、2つのベクトル間の相関をもっともよく表すような新しい変量に変換し、その変量によって2つのベクトル間の相関について理解しようとする。

6.1 正順相関分析と新変量の統計量

今、 p 個の変量を含む計測ベクトル $\mathbf{x} = (x_1, \dots, x_p)^T$ と q 個の変量を含む計測ベクトル $\mathbf{y} = (y_1, \dots, y_q)^T$ があるとする。正準相関分析では、学習用のデータ集合 $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$ に対して、線形変換

$$\begin{aligned} \mathbf{s}_i &= A^T(\mathbf{x}_i - \bar{\mathbf{x}}) = A^T \tilde{\mathbf{x}}_i \\ \mathbf{t}_i &= B^T(\mathbf{y}_i - \bar{\mathbf{y}}) = B^T \tilde{\mathbf{y}}_i \end{aligned} \quad (64)$$

で写された2つの新変量群間の相関行列のトレースの絶対値が最大となるような係数行列 A および B を求める。最適な係数行列は固有値問題

$$\Sigma_{XY} \Sigma_Y^{-1} \Sigma_{YX} A = \Sigma_X A \Lambda^2 \quad (A^T \Sigma_X A = I_L) \quad (65)$$

あるいは、

$$\Sigma_{YX} \Sigma_X^{-1} \Sigma_{XY} B = \Sigma_Y B \Lambda^2 \quad (B^T \Sigma_Y B = I_L) \quad (66)$$

の解として求まる。ただし、 Λ^2 は、固有値を対角要素とする対角行列であり、

$$\Sigma_X = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$$

$$\begin{aligned}
\Sigma_Y &= \frac{1}{N} \sum_{i=1}^N (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})^T \\
\Sigma_{XY} &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{y}_i - \bar{\mathbf{y}})^T \\
\Sigma_{YX} &= \Sigma_{XY}^T
\end{aligned} \tag{67}$$

である。このとき、係数行列 A と B の間には、

$$\Sigma_{XY}B = \Sigma_X A\Lambda, \quad \Sigma_{YX}A = \Sigma_Y B\Lambda \tag{68}$$

という関係が成り立つ。また、 L は、正準主成分 s_i および t_i の次元で、係数行列として固有値の大きいものから順に L 個取られる。

このとき、変量 s および t の統計量は、

$$\begin{aligned}
\bar{s} &= \mathbf{0} \\
\bar{t} &= \mathbf{0} \\
\Sigma_S &= \frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})(s_i - \bar{s})^T = A^T \Sigma_X A = I_L \\
\Sigma_T &= \frac{1}{N} \sum_{i=1}^N (t_i - \bar{t})(t_i - \bar{t})^T = B^T \Sigma_Y B = I_L \\
\Sigma_{ST} &= \frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})(t_i - \bar{t})^T = A^T \Sigma_{XY} B = \Lambda \\
\Sigma_{TS} &= \Sigma_{ST}^T = \Lambda
\end{aligned} \tag{69}$$

となる。したがって、新しい変量 s から t へ、あるいは t から s への線形回帰式は、

$$\tilde{t} = \Lambda s, \quad \tilde{s} = \Lambda t \tag{70}$$

で与えられる。

6.1.1 印象語からの絵画データベースの検索

「ロマンチックで暖かい」というような視覚的印象は、利用者の好みや文化的背景などの違いにより個人ごとに異なっていると考えられる。したがって、そうした要求からの検索を実現するためには、システムはその利用者にとって「ロマンチックで暖かい」とはどういうことなのか評価できなければならない。ただし、データベースに蓄えられるすべての絵画に対して、利用者ごとの主観的属性をあらかじめ登録することは困難である。そこで、栗田等 [35] は、各利用者に学習用の絵画に対して印象語を付けてもらい、その結果から正準相関分析により絵画の色彩特徴と印象語の組の相関関係を学習し、その利用者のための検索空間を構成する方法を提案した。

栗田等の実験用の絵画データベース (ART MUSEUM) には、印象派の風景画を中心に約 200 枚の絵画が蓄えられており、概略画からの検索と絵画から受ける印象からの検索を可能にしている。

ART MUSEUM では、絵画の色彩に関する画像特徴として、局所自己相関に基づく特徴を用いている。今、カラー画像を

$$\{f(i, j) = (r(i, j), g(i, j), b(i, j))^T \mid i \in I, j \in J\} \tag{71}$$

とする。ただし、 $r(i, j)$, $g(i, j)$, $b(i, j)$ は、それぞれ、画素 (i, j) における赤、緑、青成分を表す。このとき、カラー画像に対する 0 次の自己相関は、

$$\begin{pmatrix} \bar{r} \\ \bar{g} \\ \bar{b} \end{pmatrix} = \frac{1}{IJ} \sum_{i \in I, j \in J} \mathbf{f}(i, j) \quad (72)$$

で定義される。これらは画像の赤、緑、青成分の平均値である。これにより 3 個の特徴が得られる。また、1 次の自己相関特徴は、変移方向を $\mathbf{a} = (a_i, a_j)$ とすると、

$$\begin{bmatrix} \sigma_r & \sigma_{rg} & \sigma_{rb} \\ \sigma_{rg} & \sigma_g & \sigma_{gb} \\ \sigma_{rb} & \sigma_{gb} & \sigma_b \end{bmatrix} = \frac{1}{IJ} \sum_{i \in I, j \in J} \mathbf{f}(i, j) \mathbf{f}(i + a_i, j + a_j)^T \quad (73)$$

で定義される。変移方向 \mathbf{a} の取り方により、無数の可能性が考えられるが、ここでも、簡単のため、それを 3×3 の局所領域に限定する。局所領域のシフトにより同じになる変移方向を除くと、可能な変移方向のパターンは、中心点を参照点として 5 種類に限定される。各パターンに対して、6 個の独立な特徴が計算されるので、合計 30 個の 1 次の局所自己相関特徴が得られる。したがって、0 次と 1 次の特徴を合わせると、各画像から全体として 33 個の特徴が得られる。

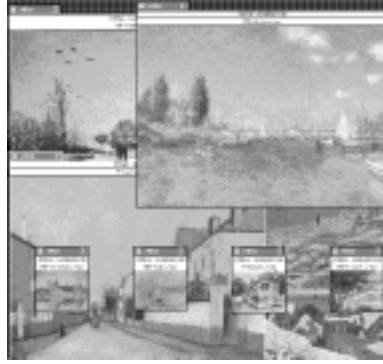
一方、色彩に対する印象を表現するには、言葉 (印象語) を用いるのが最も基本的かつ簡単である。効率的な印象語からの検索を可能にするためには、印象語として色彩に対してもつ印象をうまく表現するような形容詞を選ばなければならない。ART MUSEUM では、色彩と色彩に対する印象語の関係について書かれているいくつかの文献 [24, 70] において、色に対してよく使われている言葉を参考にして、表 1 にあげる 30 個の形容詞を印象語として用いている。

表 1: 絵画検索のための印象語

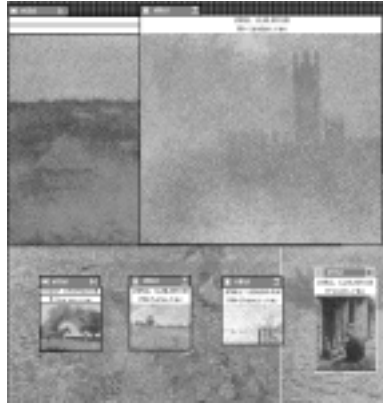
すがすがしい (clear)	華麗な (gogeous)
若々しい (fresh)	エスニックな (ethnic)
洗練された (refined)	躍動的な (dynamic & active)
モダンな (modern)	柔らかな (soft)
格調のある (authentic)	重々しい (heavy)
ダンディーな (dandy)	暖かい (warm)
クラシックな (classic)	クールな (cool)
ナチュラルな (natural)	地味な/落ち着いた (sober)
シックな (chic)	明るい/輝かしい (bright)
清雅な (elegant/tasteful)	静かな (quiet)
優雅な (elegant/relaxed)	固い (hard)
甘美な (romantic)	清潔な/新鮮な (clean)
かわいい (pretty)	スポーティー (sporty)
なごやかな (casual/relaxed)	日本的 (Japanesqe)
楽しい (casual/pleasant)	田舎風の (country)

利用者は、これらの形容詞の中から検索したい絵画の視覚的印象を表すのに適当と思われるものをいくつか選択することによって視覚的印象を表現する。それは、システムの内部では、対応する形容詞が選択されたかどうかによって 0 か 1 の値をもつ 30 次元のベクトルとして表現される。さらに、このベクトルを 1 の個数で正規化し、印象の表現ベクトル x とする。

何枚かの絵画に対する印象を利用者に答えてもらったデータに対して、正準相関分析を適用すると、その利用者の印象ベクトルと画像特徴ベクトルとの相関が最も大きくなるような写像が構成される。これにより、そのままでは比較することができなかった印象ベクトルと画像特徴ベクトルとの比較が可能となる。



(a) 「クール, すがすがしい, 清潔な/新鮮な」



「暖かい, 甘美, 柔らかい」

図 10: 印象語からの検索結果

もちろん、有効な検索のためには、正準相関分析のための学習用のデータはデータベースに蓄えられる絵画の特性を充分反映するように選ばれていなければならない。また、絵画の検索を高速に行うには、検索空間の次元 L は小さい方がよい。正準相関分析を用いると、印象ベクトルと画像特徴ベクトルの相関の大きい(固有値の大きい)成分から順番に正準主成分が構成されるので、検索空間の次元を減らすことができる。検索空間の次元は検索時間に直接的に関係するので、この次元を減らすことは検索時間の短縮につながる。

一度、正準相関分析により各利用者の印象と絵画の色彩特徴との相関関係を学習しておく、利用者の主観的印象を表す印象ベクトル x に対する検索空間での表現ベクトル $\tilde{t} = \Lambda A^T(x - \bar{x})$ とデータベースに蓄えられている絵画に対する検索空間での表現ベクトル $t_k = B^T(y_k - \bar{y})$ の距離に近い絵画を検索することができる。

200 枚の画像データに対し、三つの印象語をキーとした検索した結果を図 10 に示す。図 10(a) は、「クールな」、「すがすがしい」、「清潔な/新鮮な」という印象語で検索した絵画である。上段の左側が第 1 候補として検索された絵画で、右側が第 2 候補である。下段は、左から第 3 候補、第 4 候補の順番で表示してある。一方、図 10(b) は、「暖かい」、「甘美な」、「柔らかい」という印象語で検索した結果である。この場合には、暖色系の色彩の絵画が検索されている。

7 ロバスト統計

重回帰分析 (最小 2 乗法) では、モデルとデータとの誤差が平均 0 の正規分布に従う場合には、推定されたモデルは最適なものとなるが、データに例外値が含まれているような場合には、推定結果は例外値の影響を大きく受けて、必ずしもよいモデルが推定できるとは限らない。

画像処理やコンピュータビジョンでは、データに例外値が含まれていることを前提にしてデータからモデルのパラメータを推定できれば便利なが多い。例えば、複数の動きを含んだデータから主要な動きのみを推定したい場合には、異なる動きのデータを同時に処理し、データの主要な部分のみから運動のパラメータを推定できれば便利である。

ロバスト統計を用いると、例外値をある程度含むようなデータからでも比較的安定にモデルのパラメータを推定できる。以前は、ロバスト統計に基づく手法は、メディアンフィルタなどの画像の平滑化以外ではあまり利用されていなかったが、1990 年代に入って、様々な用途で使われるようになり、最近では、画像処理やコンピュータビジョンの分野でもモデル推定のための標準的な手法とみなされるようになってきている。

7.1 ロバスト推定手法

ロバスト推定のための代表的な手法は、M-estimators と呼ばれている方法と Least Median of Squares (LMedS) 基準に基づく方法である [27, 25, 58, 59]。

7.1.1 M-estimator

M-estimator は、最も良く利用されるロバスト推定手法のひとつであり、最小 2 乗法で用いられる最小 2 乗 (LMS) 基準

$$LMS = \min \sum_i \varepsilon_i^2 \quad (74)$$

の代わりに、それをちょっとだけ変形した評価基準を用いる。ここで、 ε_i はデータと推定値との誤差である。最小 2 乗基準はすべての誤差を均等な重みで扱っているため、ひとつの大きな例外値によって大きな影響を受けてしまう。そこで、M-estimator では、これを例外値には小さな重みを与えるように変形し、

$$M = \min \sum_i \rho(\varepsilon_i) \quad (75)$$

を評価基準とする。ここで、 $\rho(x)$ は、 $x = 0$ で唯一の最小値を持つ正定値偶関数である。この基準は、もし $\rho(x) = x^2$ ならば、平均 2 乗誤差基準と同じになり、その意味で最小 2 乗法の拡張になっている。

関数 ρ によってモデルからずれたデータに対してどれくらいの重みを与えられるかを見るためには、 ρ の微分 $\Psi(x) = \frac{\partial \rho(x)}{\partial x}$ を取ってみればよい。 Ψ は、influence function と呼ばれている。平均 2 乗誤差基準に対する influence function は、図 11 (b) のようになり、データがモデルから離れるにつれてその影響が線形に増大する。一方、図 11(c) のような Geman と McClure の ρ 関数では、データがモデルからある程度離れるとその影響はほとんどなり、0 に近づく (図 11(d))。

M-estimator による推定のためのアルゴリズムは、式 (75) を最小化する最適化問題として定式化することができ、重み付き最小 2 乗問題となる。ただし、このアルゴリズムは、 ρ 関数によっては必ずしも最適解に収束することは保証されないので、良い初期値から出発する必要がある。

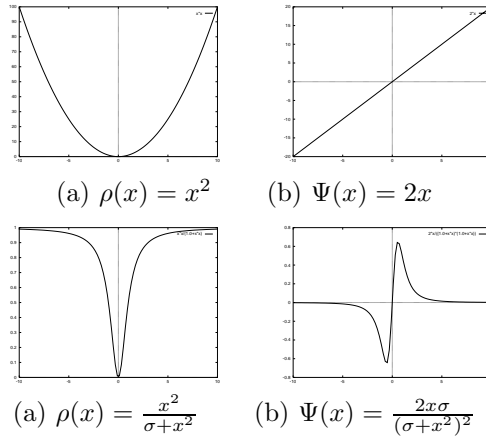


図 11: Examples of ρ and Ψ functions.

7.1.2 LMedS 推定

Rousseeuw が提案した LMedS 基準は、

$$LMedS = \min \text{med } \varepsilon_i^2 \quad (76)$$

で定義される。ここで med は、中央値を取ることを表す。こうした評価基準が例外値に対してどれくらい強いかが（例外値に対するロバストネス）を評価するためには、breakdown point と呼ばれる指標が用いられている [27, 59]。それは、例外値が無い場合の結果と例外値を含む場合の結果が非常に大きくずれることなく、全データの内の何割までのデータを非常に大きな例外値に置き換えることができるかを測るものである。最小 2 乗基準は、ひとつの大きな例外値によって大きな影響を受けるので、その breakdown point は 0 であるが、LMedS 基準の breakdown point は 0.5 であり [58]、全データのうちの 50% が例外値でも結果が大きくずれることがない。これは、考えられる最大の breakdown point を達成する基準である。ただし、漸近有効性（asymptotic efficiency）の観点からは、LMedS 基準はそれほど良いとは言えない。

LMedS 基準によるパラメータ推定は非線形の最適化問題となる。1 次元のデータに対して位置を推定する場合には比較的効率的なアルゴリズムが知られているが [59]、多次元の場合には最適解を見つけるのはかなり難しい。そこで、準最適解をランダムサンプリングによって探索するアルゴリズムが開発されている。例えば、データに F 個のパラメータ $\omega_j, j = 0, \dots, F-1$ を持つ線形モデル

$$z_i = \sum_{j=1}^{F-1} \omega_j x_j(i) + \omega_0 \quad (77)$$

を当てはめる場合を考えよう。この時、LMedS 基準は

$$LMedS = \min \text{med } (y_i - z_i)^2 \quad (78)$$

となる。この場合のランダムサンプリングを用いた推定アルゴリズムは以下のようになる。

1. 全データから F 個のデータをランダムに選ぶ。
2. F 個のデータを用いてモデルのパラメータを推定する。
3. LMedS 基準によりそのパラメータのモデルを評価する。

以上を繰り返して、最良のパラメータを選択する。

このアルゴリズムに必要な繰り返し回数は、 q 回のランダムサンプリングで少なくとも 1 個のサンプルには例外値が含まれ無い確率を考えることにより決定できる。全データ中の例外値の割合を ϵ とすると、この確率は、

$$P = 1 - \{1 - (1 - \epsilon)^F\}^q \quad (79)$$

となる。例えば、 $\epsilon = 0.3$, $F = 3$ の時、この確率が $P = 0.01$ になるためのサンプリングの回数は $q = 11$ である。つまり、この場合には、 $q = 11$ 回以上のサンプリングが必要であることを意味している。

このアルゴリズムは、単純なランダムサンプリングによるアルゴリズムであるが、これを遺伝的アルゴリズムを使って高速化する試みも報告されている [60]。

ロバスト推定によって得られたモデルとデータとの誤差 ε_i から、誤差の標準偏差は、

$$\hat{\sigma} = C \left\{ 1 + \frac{5}{n - F} \right\} \text{med} \sqrt{\varepsilon_i^2} \quad (80)$$

のように推定できる。ここで、 $C = 1.4826$ は、誤差が正規分布の場合に一致させるための係数であり、また、 $5/(n - F)$ は、少数サンプルのための補正項である。

この標準偏差の推定値から、 $2.5\hat{\sigma}$ よりも大きな誤差 $|\varepsilon_i|$ をもつデータを例外値として判定することができる。

このような方法で例外値を検出し、例外値を取り除いたデータに対して最小 2 乗法を適用すると、LMedS 推定の漸近有効性を向上させることができる。

7.1.3 オプティカルフローの推定への応用

勾配法を用いてオプティカルフローを推定する場合、考えている領域内のすべての画素が同時に同じ運動をしている場合には最小 2 乗法を用いても問題は無いが、領域内に複数の運動が含まれている場合、隠れがある場合や影が写っているような場合には、最小 2 乗法では正確な推定値が得られない。一方、ロバスト推定を用いると主要な運動のみにモデルをあてはめることができるため、より正確な推定値を得ることができる。

例えば、Black 等 [8] は、M-estimator を用いる方法を提案している。また、梅山 [81] は、LMedS 推定を用いる方法を示している。Sawhney 等 [65] は、ビデオ映像からの主要な動きの推定に M-estimator を利用してしている。また、Black 等 [9] は、あらかじめ輝度情報で領域分割し、ロバスト推定を用いて各領域に複数のモデルの内から最もあてはまりのよいモデルを選択することにより、オプティカルフローを推定する方法を示している。その他、栄藤等 [18] は、動画像符号化やパノラマ画像生成のために、M 推定を拡張したロバスト推定とパラメータ化ブロック相関を用いた動き検出方法を提案している。

7.1.4 不連続を含むデータの復元への応用

次に、ロバスト統計の考え方のちょっと変わった応用例として、不連続性を例外とみなす滑らかさの評価基準について述べる。

正則化 [63, 84] は、与えられたデータだけでは解が一意に定まらないような問題 (不良設定問題) を「滑らかさ」などの適切な条件を付加することにより、良設定問題に変換して、近似解を安定に得るための手法である。コンピュータビジョンでも、奥行情報の復元やオプティカルフローの推定、スネークス [29] などで盛んに利用されている。

正則化では、未知の量 z から観測過程 A を通して観測されたデータ y から、観測データと解の差を表

すペナルティー項と滑らかさ等の制約を与える安定化項との重み付き和

$$E = \sum_i \|Az_i - y_i\|^2 + \lambda \sum_i \|(Pz)_i\|^2 \quad (81)$$

を最小化することにより z を推定する。ここで、重み λ は、正則化パラメータと呼ばれている。

正則化の条件として「滑らかさ」を用いた場合、不連続な部分までもが滑らかになり、不連続性の復元が難しいことが知られている。この問題に対して、Geman 等 [21] は、線過程 (line process) を導入して、不連続性を解決しようとしている。ロバスト推定の考え方を適用すると、比較的簡単に、不連続な部分を例外値として自動的に除外し、それ以外の部分を滑らかに復元することが可能となる [80]。

一般的には、式 (81) のように、観測データと解の差を表すペナルティー項と滑らかさ等の制約を与える安定化項は、それぞれの 2 乗誤差で定義されるが、これらの項を M-estimator の考え方を適用して、

$$E = \sum_i \rho(Az_i - y_i) + \lambda \sum_i \rho((Pz)_i) \quad (82)$$

のように変形すると、第 1 項でデータに含まれる例外値を自動的に除外し、第 2 項で不連続点を例外値として自動的に除外することができる。

Black 等 [8, 9] は、オプティカルフローの推定にこの方法を利用し、良好な結果を得ている。また、梅山 [80] は、不連続がある輪郭の抽出のために、ロバストな滑らかさ基準を導入したスネークスを提案している。

8 EM アルゴリズム

画像処理やコンピュータビジョンの問題を困難にしている原因の一つとして、扱う対象が複数同時に存在することが多いことが挙げられる。例えば、画像中には一般に複数の物体が存在している。ひとつの物体の認識や識別は比較的易しい場合もあるが、複数存在する場合には一つ一つの物体を切り出さなければひとつの物体のみを認識する手法は使えない。しかし、まだ認識もできていない物体を切り出すのは容易なことではない。

複数物体の認識のほかにも、画像が複数の領域からなっているときに領域を分割する問題、あるいは、画像中の複数の物体が移動している場合にそれぞれの運動パラメータやオプティカルフローを求める問題などはいずれも複数性によって困難が生じている。

このような問題に対処するためのアプローチとして、複数対象のモデル化の問題を確率分布の混合分布の最尤推定問題としてとらえる方法がある。その解は、EM アルゴリズムを利用することにより求めることができる。また、局所最適解ではあるがアルゴリズムの収束性も示すことができる。

8.1 隠れ変数と EM アルゴリズム

複数の対象があって、それぞれが確率分布の形であらわされているとする。複数の対象全体はその (重み付きの) 和

$$p(x; \theta) = \sum_{i=1}^k \xi_i p_i(x; \psi_i) \quad (83)$$

で表現することができるとする。ここで k は対象の個数、 $p_i(x; \psi_i)$ は i 番目の対象の確率分布、 ξ_i はその重みをあらわす。この形の分布は (有限) 混合分布と呼ばれている。最尤推定では、 x のサンプル点に基づいて、尤度を最大にするようなパラメータ θ を求めることが目的であるが、一般にこのままでは複雑すぎて解けない。

そこで、もし仮に、各サンプル点がどの対象に属しているかがわかったとすれば、あとはそれぞれの対象ごとの確率分布の推定を行えばよくなる。どの対象に属しているかというのを観測できない変数（または隠れた変数）とみなすと以下で述べるように EM アルゴリズムが適用可能となる。

今、観測される変数を x 、観測できない変数を z とすると、EM アルゴリズムは、次の 2 つのステップの計算を $t = 0, 1, 2, \dots$ に対して繰り返し行なう。まず、パラメータを適当な値 $\theta = \theta^{(0)}$ に初期化しておく。

1. (Expectation) step (x, z) の対数尤度の (与えられたデータ x とパラメータ $\theta^{(t)}$ に対する) 条件付き期待値を計算する。つまり、

$$Q(\theta | \theta^{(t)}) = E[\log p(x, z; \theta) | x, \theta^{(t)}] \quad (84)$$

を計算する。

2. (Maximization) step $Q(\theta | \theta^{(t)})$ を最大にする θ を $\theta^{(t+1)}$ とおく。

この各繰り返しステップで得られるパラメータは尤度を単調に増加させることが知られている。

混合分布の場合に EM アルゴリズムを適用すると、 ξ に関しては Q の最大化が陽に解けて、 ψ_i は各対象に関する重み付きの最適化問題となり、対象ごとに独立に解くことができる。

8.1.1 正規混合分布に対する EM アルゴリズム

各対象の確率分布が正規分布としてあらわされる正規混合分布

$$p(x; \theta) = \sum_{i=1}^k \xi_i \phi(x; \mu_i, \sigma_i^2), \quad (85)$$

の場合には、さらに ψ_i も陽に解けて、EM アルゴリズムの各繰り返しステップは以下のように書ける。ここで N 個のサンプル x_1, \dots, x_N があったとしよう。

1. 各サンプル x_j に対する、観測できない変数 z の条件付き確率

$$\begin{aligned} w_{zj}^{(t)} &\equiv p(z | x_j; \theta^{(t)}) \\ &= \frac{\xi_z^{(t)} \phi(x_j; \mu_z^{(t)}, \sigma_z^{2(t)})}{\sum_{i=1}^k \xi_i^{(t)} \phi(x_j; \mu_i^{(t)}, \sigma_i^{2(t)})}, \end{aligned} \quad (86)$$

を計算する。

2. 重み確率を次の式で更新する

$$\xi_i^{(t+1)} = \frac{1}{N} \sum_{j=1}^N w_{ij}^{(t)} \quad (87)$$

3. 各正規分布のパラメータを以下の式で更新する

$$\mu_i^{(t+1)} = \frac{1}{N \xi_i^{(t+1)}} \sum_{j=1}^N w_{ij}^{(t)} x_j, \quad (88)$$

$$\sigma_i^{2(t+1)} = \frac{1}{N \xi_i^{(t+1)}} \sum_{j=1}^N w_{ij}^{(t)} x_j^2 - \mu_i^{(t+1)2}. \quad (89)$$

この更新式を見ると、正規分布の各パラメータが $w_{zj}^{(t)}$ で重みづけされた平均と分散になっていることがわかる。

8.1.2 EM アルゴリズムの応用

混合分布の EM アルゴリズムの応用として、Sekita[68] は複数直線のあてはめに対して EM アルゴリズムを適用した。Akaho[2] は、位置・尺度が変化する任意形状の複数物体を認識するための手法を EM アルゴリズムから導いている。また、Ayer et al[6] は画像のシーケンス中における複数物体の動きパラメータの分布を混合分布でモデル化し、EM アルゴリズムにより推定された結果に基づいて領域分割を行なっている。

隠れマルコフモデルの応用例として、Bose et al[10] はテキスト認識のモデルに隠れマルコフモデルを用いている。また、Starner et al[71] は手話 (American Sign Language) の認識に隠れマルコフモデルを用いている。これらはいずれもあらかじめ分類するクラスの数だけのモデルを用意し、それぞれのモデルのパラメータを学習データを用いて推定しておく。新たにテストデータが与えられると、それぞれのモデルに対してそのデータの尤度を計算し最も値が高くなるモデルのクラスを出力する。Zhang[85] では画像を隠れ状態を含む MRF でモデル化し、領域分割への応用を行なっている。

Fan et al[20] は、画像フレーム間の動きパラメータを部分ブロックの直交展開として定式化し、そのパラメータ推定問題に対して EM アルゴリズムを適用している。なお、混合分布や隠れマルコフモデルの場合は観測できない変数が離散値をとる確率変数であったが、Fan et al の扱っている問題では連続値の確率変数の場合になっている。

9 モデル選択

コンピュータビジョンでは、例えば、画像の領域分割における分割数の決定、オプティカルフローの推定のためのモデルの決定、正則化手法における正則化パラメータ λ の決定などのように、複数のモデルの中からデータをうまく説明するモデルを選択しなければならないことがある。また、モデルのパラメータを精度良く推定するためにもデータをうまく説明するモデルをデータに即して決定することが必要となる。

重回帰分析において、平均 2 乗誤差は学習データに対するモデルのあてはまりの良さを評価する基準であると考えられる。しかし、モデル $f(x, \theta)$ の自由度を大きくすると学習データに対する平均 2 乗誤差をどんどん小さくすることができる。そのモデルが学習データに対していくら良くあてはまったとしても、必ずしも未知のデータに対しても良くあてはまるとは限らないので、平均 2 乗誤差をモデル選択の基準として用いることはできない。未知データに対しても有効に働くようなモデル、つまり、データの背後の確率統計的関係をうまく捉えたモデルを決定するためには、モデルの自由度に関するバイアスをもたないようなモデルのあてはまりの良さに関する評価基準が必要となる。

9.1 モデル選択手法

統計の分野では、このようなモデル選択の問題は古くから議論されており、特に有名なのは赤池の AIC (An Information Theoretical Criterion)[3, 62] や Rissanen の MDL (Minimum Description Length)[56, 57] などの情報量基準、交差確認法 (Cross-Validation 法) や bootstrap 法 [15, 16] などのデータからのサンプリングに基づく手法である。

9.1.1 情報量基準

AIC は赤池により最大対数尤度と期待平均対数尤度の間の偏りに関する解析的評価から導出されたモデル選択のための基準であり、モデルの自由度を F とすると、

$$AIC = -2(\text{最大対数尤度}) + 2F \quad (90)$$

のように定義される [3]。一方、MDL は情報圧縮のための最適な符号化法を決定するための基準として Rissanen により提案された記述長最小化原理 (Principle of Minimum Description Length) が一般化されたもので、重回帰分析でのモデル選択の場合には、

$$MDL = -(\text{最大対数尤度}) + \frac{F}{2} \log n \quad (91)$$

のように定義される [56, 57]。これらの評価基準を用いると、学習データに対するあてはまりの良さに大きな差がある場合には、第 1 項に大きな差があらわれ、あてはまりの良いモデルが選ばれ、第 1 項に大きな差が無い場合には、第 2 項が働いて自由度の小さいモデルが選ばれる。従って、これらの基準を用いたモデル選択では、まず、考えているいくつかのモデルのパラメータを学習データから最尤法により求め、そのパラメータを用いた時の対数尤度 (最大対数尤度) とそのモデルの自由度から AIC あるいは MDL を計算し、その値が最も小さくなるモデルを選択すればよい。モデル選択の基準は、AIC や MDL 以外にもいくつか提案されているが、基本的には、データに対するあてはまりの良さとモデルの自由度とのバランスを取ることににより、最良のモデルを選択するような基準となっている。

9.1.2 データからのサンプリングに基づく手法

情報量基準は、漸近展開により、モデルの自由度に関するバイアスを式の形で求めて補正する評価基準であるが、データからのサンプリングに基づく手法では、コンピュータの計算パワーを利用して学習データからのサンプリングを繰り返し、その平均的な振舞いを知る事ことにより、バイアスの少ない評価基準を得ようとする。その代表的な方法には、交差確認法 (Cross-Validation 法) や bootstrap 法 [15, 16] などがある。

交差確認法では、大きさ n の学習データを推測用のデータと評価用のデータに分け、推測用データから求めた推定値を用いたモデルを評価用データにあてはめることにより、モデルの良さを評価する。データの分け方としては、推測用のデータとして $n - 1$ 個のデータを用い、残りの 1 個のデータを評価に用いることが多い。そして、 n 個のデータの各々がちょうど 1 回ずつ評価用データとなるように n 通りの分け方を考え、その平均値によりモデルを評価する。モデル選択の場合には、各モデル毎に交差確認法でモデルの良さ (例えば、平均 2 乗誤差) を評価し、それらを比較することにより、最良のモデルが決定できる。

bootstrap 法では、 n 個の学習データから重複を許して m 個の標本をランダムに選び、それに基づいてパラメータを計算し、そのパラメータに基づいてモデルの良さ (例えば、平均 2 乗誤差) を評価するという操作を多数回繰り返し、それらの平均を比較することにより、最良のモデルを決定する。

こうしたデータからのサンプリングに基づく手法は、多数回のサンプリングとそれに基づいてパラメータを推定するための膨大な計算が必要となるが、情報量基準のような漸近展開を用いていないので、手法を適用する際に要求される条件がかなり緩やかである。しかし、bootstrap 法におけるサンプリングの回数を選び方や交差確認法における評価用データと推測用データの分割の仕方については、必ずしも最善の方法が知られているわけではない。

画像処理やコンピュータビジョンの応用では、MDL が頻繁に用いられているが、場合によってはデータからのサンプリングに基づくモデル選択手法も適用して見る価値があると思う。

9.1.3 領域分割への応用

画像を局所的な特徴が一樣な領域に分割する領域分割は、コンピュータビジョンにおける最も基本的な手法のひとつであり、小さな領域から一樣な領域を統合する統合法、大きな領域を分割する方法、確率的弛緩法を用いる方法等の多くの手法が提案されている [74]。領域の統合をどこで止めるか、あるいは、領域の分割とどこで止めるかといった分割数の決定の自動化は、これらの手法に共通する課題である。モデ

ル選択手法は、その解決策を与える有力な手段であり、MDLがよく利用されている。

例えば、Leclerc[42]は、画像の領域分割の問題をMDLの観点から定式化し、連続変形法によるアルゴリズムを提案している。Kanungo等[28]は、Leclerc[42]が定式化した画像の領域分割の問題の準最適解を高速に求めるために、領域統合に基づく手法を提案している。また、Zhu等[87, 88]は、それらをさらに発展させ、マルチバンドの画像の領域分割に適用している。さらに、Zheng等[86]は、動画画像圧縮への応用のための動きによる領域分割にMDLを用いている。

9.1.4 正則化パラメータの決定への応用

次に、ちょっと変わったモデル選択の応用例として、正則化の際の正則化パラメータの決定方法について述べる。

正則化[63, 84]は、不良設定問題を「滑らかさ」などの適切な条件を付加することにより、良設定問題に変換して、近似解を安定に得るための手法であり、正則化パラメータを変化させることにより、取り得る解の自由度が変化するので、適切なモデルを得るためにはデータに応じて適切に正則化パラメータを設定する必要があり、これもある種のモデル選択と解釈できる。

正則化手法における正則化パラメータ λ の決定に関しては、いくつかの手法が提案されている。Shahraray等[69]は、交差確認法を変形したGeneralized Cross-Validationを用いた方法を2次元点列からの輪郭の復元に利用している。また、Thompson等[76]は、画像の復元問題を例に、Total Predicted Mean Square ErrorやGeneralized Cross-Validationなどのいくつかの基準の比較を行っている。最近では、MacKay[43]が、経験Bayes推定の観点から、正則化パラメータ λ の決定法を導出している。

10 おわりに

本稿では、基本的な統計手法、特に、多変量データ解析と最近画像処理等でも盛んに利用されるようになってきた統計手法について、具体的な応用例を示しながら、その考え方を示した。本稿が、さまざまな画像処理課題を解決する際に、統計手法やその考え方をうきかけとして、少しでも役に立てば幸いである。

参考文献

- [1] S.Akaho: Translation, scale and rotation invariant features based on higher-order autocorrelations, Bulletin of the Electrotechnical Laboratory, Vol.57, No.10, pp.973-981, 1993.
- [2] Akaho, S.: "The EM algorithm for multiple object recognition", Proc. of ICNN'95, pp.2426-2431, Perth, 1995.
- [3] Akaike, H.: "A new look at the statistical model identification", IEEE Trans. on Automatic Control, Vol.AC-19, NO.6, pp.716-723, 1974.
- [4] 赤松: コンピュータによる顔の認識の研究動向, 電子情報通信学会誌, Vol.80, No.3, pp.257-266, 1997.
- [5] 赤松, 佐々木, 深町, 末永: 濃淡画像のマッチングによるロバストな正面顔の識別法—フーリエスペクトルのKL展開の応用—, 電子情報通信学会論文誌, Vol.J76-D-II, No.7, pp.1363-1373, 1993.
- [6] Ayer, S. and Sawhney, J. S.: "Layered representation of motion video using robust maximum-likelihood estimation of mixture models and MDL encoding", Proc. of ICCV'95, pp.777-784, 1995.
- [7] D.H.Ballrad and C.M.Brown, *Computer Vision*, Prentice-Hall (1982).
- [8] Black, M. J. and Anandan, P.: "The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, Computer Vision and Image Understanding", Vol.63, No.1, pp.75-104, 1996.
- [9] Black, M. J. and Jepson, A. D.: "Estimating optical flow in segmented images using variable-order parametric models with local deformations", IEEE Trans. on PAMI, Vol.18, No.10, pp.972-986, 1996.

- [10] Bose, C. B. and Kuo, S.-S : “Connected and degraded text recognition using hidden markov model”, Proc. of ICPR’92, pp.116-119, 1992.
- [11] R.Chellappa, C.L.Wilson, and S.Sirohey: Human and machine recognition of faces: a survey, Proc. of the IEEE, Vol.83, NO.5, pp.705-740, 1995.
- [12] Dempster, A., Laird, N. and Rubin, D. : “Maximum likelihood from incomplete data via the EM algorithm”, J. Roy. Statist. Soc. B, Vol.39, pp.1-38, 1977.
- [13] K.I.Diamantaras and S.Y.Kung, “Principal Component Neural Networks: Theory and Applications,” John Wiley & Sons, 1996.
- [14] J.Durbin, “The fitting of time-series models”, Rev. Inst. Int. de Stat., Vol.28, No.3, pp.233-244, 1960.
- [15] Efron, B. : “Bootstrap methods: another look at the jackknife”, The Annals of Statistics, Vol.7, No.1, pp.1-26, 1979.
- [16] Efron, B. : “Estimating the error rate of a prediction rule: improvements in cross-validation”, J. American Statist. Assoc., Vol.78, pp.316-331, 1983.
- [17] Efron, B. : “The bootstrap method for assessing statistical accuracy”, Behaviormetrika, Vol.17, pp.1-35, 1985.
- [18] 栄藤, 安慶: “パラメータ化ブロック相関 – グローバル動き保証とビデオモザイクのための2次元パラメトリック動き予測 –”, 信学技報, PRMU-97-93, 1997.
- [19] T.Ezzat and T.Poggio: Facial analysis and synthesis using image-based models, Proc. of Inter. Conf. on Automatic Face and Gesture Recognition, pp.116-121, 1996.
- [20] Fan, C. M., Namazi, N. M. and Penafiel, P. B. : “A new image motion estimation algorithm based on the EM technique”, IEEE Trans. on PAMI, Vol.18, No.3, pp.348-352,1996.
- [21] Geman, S. and Geman, D. : “Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images”, IEEE Trans. on PAMI, Vol.6, No.6, pp.721-741, 1985.
- [22] F.Goudail, E.Lange, T.Iwamoto, K.Kyuma, and N.Otsu: Face recognition system using local autocorrelations and multiscale integration, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.18, No.10, pp.1024-1028, 1996.
- [23] Gu, H., Shirai, Y. and Asada, M. : “MDL-based segmentation and motion modeling in a long image saequence of scene with multiple independently moving objects”, IEEE Trans. on PAMI, Vol.18, No.1, pp.58-64, 1996.
- [24] 日本カラーデザイン研究室, 配色イメージブック, 講談社, 1984.
- [25] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W.A. : *Robust Statistics: An Approach Based on Influence Functions*, John Wiley & Sons, 1986.
- [26] O.Hasegawa, K.Itou, T.Kurita, S.Hayamizu, K.Tanaka, K.Yamamoto, N.Otsu, “Active agent oriented multimodal interface system,” Proc. IJCAI-95, pp.82-87, 1995.
- [27] Huber, P. J. : *Robust Statistics*, John Wiley & Sons, 1981.
- [28] Kanungo, T., Dom, B., Niblack, W. and Steele, D. : “A fast algorithm for MDL-based multi-band image segmentation”, Proc. of CVPR’94, pp.609-616, 1994.
- [29] Kass, M., Witkin, A. and Terzopoulos, D. : “Snakes: Active contour models”, Inter. J. of Computer Vision, pp.321-331, 1988.
- [30] M.Kirby and L.Sirovich: Application of the Karhunen-Loève procedure for the characterization of human faces, Trans. on Pattern Analysis and Machine Intelligence, Vol.12, No.1, pp.103-108, 1990.
- [31] T.Kohonen, *Associative Memory*, Springer-Verlag, 1977.
- [32] M.Kreuts, B.Volpel and H.Jansen: Scale-invariant image recognition based on higher-order autocorrelation features, Pattern Recognition, Vol.29, No.1, pp.19-26, 1996.
- [33] 栗田, 関田, 大津, “複素自己回帰モデルに基づく輪郭形状間の距離,” 電子情報通信学会論文誌, Vol.J73-D-II, No.9, pp.1493-1503, 1990.
- [34] T.Kurita, N.Otsu, and T.Sato, “A face recognition method using higher order local autocorrelation and multivariate analysis,” Proc. of Int. Conf. on Pattern Recognition, Aug. 30 - Sep. 3, The Hague, Vol.II, pp.213-216, 1992.
- [35] 栗田, 加藤, 福田, 坂倉, “印象語による絵画データベースの検索,” 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383, 1992.

- [36] 栗田: “ロバスト統計に基づく画像解析”, 信学誌, Vol.76, No.12, pp.1293-1297, 1993.
- [37] T.Kurita, I.Sekita, and N.Otsu, “Invariant distance measures for planar shapes based on complex autoregressive model,” Pattern Recognition, Vol.27, No.7, pp.903-911, 1994.
- [38] 栗田, 堀田, 三島: Log-Polar 画像の高次局所自己相関特徴を用いた大きさ・回転に不変な 2 次元図形及び顔画像の認識, 信学技報, Vol.PRMU96-212, 1997.
- [39] 栗田, 小林, 三島, “PARCOR 画像の高次局所自己相関特徴を用いた背景変化および並行移動に強いジェスチャー認識,” 信学技報, PRMU96-213, p.159-164, 1997.
- [40] T.Kurita, Y.Kobayashi, and T.Mishima, “Higher order local autocorrelation features of PARCOR images for gesture recognition,” Proc. of Inter. Conf. on Image Processing, Vol.3, pp.722-725, 1997.
- [41] T.Kurita and S.Hayamizu, “Gesture recognition using HLAC features of PARCOR images and HMM based Recognizer,” Proc. of Inter. Conf. on Automatic Face and Gesture Recognition, pp.422-427, 1998.
- [42] Leclerc, Y. G. : “Constructing simple stable descriptions for image partitioning”, Inter. J. of Computer Vision, Vol.3, pp.73-102, 1989.
- [43] MacKay, D. J. C. : “Bayesian interpolation”, Neural Computation, Vol.4, pp.415-447, 1992.
- [44] Masuda, T. and Yokoya, N. : “A robust method for registration and segmentation of multiple range images”, Computer Vision and Image Understanding, Vol.61, No.3, pp.295-307, 1995.
- [45] J.A.Mclaughlin and J.Raviv, “Nth-order autocorrelations in pattern recognition,” Information and Control, Vol.12, pp.121-142, 1968.
- [46] Meer, R., Mintz, D., Rosenfeld, A. and Kim, D. Y. : “Robust regression methods for computer vision: a review”, Inter. J. of Computer Vision, Vol.6, No.1, pp.59-70, 1991.
- [47] Miller, J. V. and Stewart, C.V. : “MUSE: Robust surface fitting using unbiased scale estimates”, Proc. of CVPR'96, pp.300-306, 1996.
- [48] H.Murase and Shree K.Nayar, “Visual learning and recognition of 3-D objects from appearance,” Inter. Journal of Computer Vision, Vol.14, pp.5-24, 1995.
- [49] N.Otsu, “A threshold selection method from gray-level histograms,” IEEE Trans. on System, Man, and Cybernetics, Vol.SMC-9, No.1, pp.62-66, 1979.
- [50] 大津, “判別および最小 2 乗基準に基づく自動しきい値選定法,” 電子通信学会論文誌, Vol.J63-D, No.4, pp.349-356, 1980.
- [51] 大津: パターン認識における特徴抽出に関する数理的研究, 電子技術総合研究所研究報告, Vol.818, 1981.
- [52] N.Otsu and T.Kasvand, “Image restoration by multiplier regression analysis approach,” Proc. of Inter. Conf. on Pattern Recognition, pp.155-158, 1984.
- [53] N.Otsu, and T.Kurita, “A new scheme for practical, flexible and intelligent vision systems,” Proc. IAPR Workshop on Computer Vision, pp.431-435, 1988.
- [54] 大津, 栗田, 関田, パターン認識—理論と応用—, 朝倉書店, 1986.
- [55] A.Pentland, B.Moghaddam, and T.Starner: View-based and modular eigenspaces for face recognition, Proc. of CVPR'94, pp.84-91, 1994.
- [56] Rissanen, J. : “A universal prior for integers and estimation by minimum description length”, The Annals of Stat., Vol.11, No.2, pp.416-431, 1983.
- [57] Rissanen, J. : “Stochastic complexity and modeling”, The Annals of Stat., Vol.14, No.3, pp.1080-1100, 1986.
- [58] Rousseeuw, R.J. : “Least median of squares regression”, J. American Stat. Assoc., Vol.79, pp.871-880, 1984.
- [59] Rousseeuw, R. J. and Leroy, A. M. : *Robust Regression and Outlier Detection*, John Wiley & Sons, 1986.
- [60] Roth, G. and Levine, M. D. : “Geometric primitive extraction using a genetic algorithm”, Proc. of CVPR'92, pp.640-643, 1992.
- [61] Roth, G. and Levine, M. D. : “Extracting geometric primitives”, CVGIP:Image Understanding, Vol.58, No.1, pp.1-22, 1993.
- [62] 坂本, 石黒, 北川: 「情報量統計学」, 共立出版, 1983.
- [63] 坂上, 横矢: “弛緩法と正則化”, 情報処理, Vol.30, No.9, pp.1047-1057, 1989.
- [64] A.Samal and P.A.Iyengar: Automatic recognition and analysis of human faces and facial expressions: a survey, Pattern Recognition, Vol.25, No.1, pp.65-77, 1992.
- [65] Sawhney, H. S. and Ayer, S. : “Compact representations of videos through dominant and multiple motion estimation”, IEEE Trans. on PAMI, Vol.18, No.8, pp.814-830, 1996.

- [66] 関田, 栗田, 大津, “複素自己回帰モデルによる形の識別,” 電子情報通信学会論文誌, Vol.J73-D-II, No.6, pp.804–811, 1990.
- [67] I.Sekita, T.Kurita and N.Otsu, “Complex autoregressive model for shape recognition,” IEEE Trans. on Pattern Anal. & Machine Intelli., Vol.PAMI-14, No.4, pp.489–496, 1992.
- [68] Sekita, I. : “On fitting several lines using the EM algorithm”, Proc. of CVVC’94, pp.107-109, 1994.
- [69] Shahrraray, B. and Anderson, D. J. : “Optimal estimation of contour properties by cross-validated regularization”, IEEE Trans. on PAMI, Vol.11, No.6, pp.600-610, 1989.
- [70] 渋川, 高橋, 配色辞典 2, 河出書房新社, 1987.
- [71] Starner, T. and Pentland, A. : “Real-time american sign language recognition from video using hidden Markov models”, M.I.T. Media Lab. Perceptual Computing Section Technical Report No.375, 1995.
- [72] Stone, M. : “Cross-validators choice and assessment of statistical predictions”, J. Roy. Stat. Soc., Vol.B36, pp.111-147, 1974.
- [73] D.L.Swets and J.Weng, “Using discriminant eigenfeatures for image retrieval,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.18, No.8, pp.831-836, 1996.
- [74] 高木, 下田 (監修): 「画像解析ハンドブック」, 東京大学出版会, 1991.
- [75] Tanner, M. A. : *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, Springer-Verlag, 1993.
- [76] Thompson, A. M., Brown, J. C., Kay, J. W. and Titterton, D. M. : “A study of methods of choosing the smoothing parameter in image restoration by regularization”, IEEE Trans. on PAMI, Vol.13, No.4, pp.326-339, 1991.
- [77] M.Turk and A.Pentland: Eigenfaces for recognition, Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71-86, 1991.
- [78] M.A.Turk and A.P.Pentland, “Face recognition using eigenfaces,” Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp.586-591, 1991.
- [79] 上田, 中野: “確率的アニメーリング EM アルゴリズム”, 信学論 (D), Vol.J80-D-II, No.1, pp.267-276, 1997.
- [80] 梅山: “ロバスト統計の考え方をを用いた正則化における不連続の検出について”, 信学技報, PRMU95-217, 1996.
- [81] 梅山: “ロバスト回帰に基づく多重オブティカルフローの推定について”, 信学技報, PRMU96-108, 1996.
- [82] N.Yokoya and M.D.Levin, “Range image segmentation basedn on differential geometry: a hybrid approach,” IEEE Trans. on PAMI, Vol.11, No.6, pp.643-649, 1989.
- [83] 横矢, M.D.Levin, “微分幾何学特徴に基づく距離画像分割のためのハイブリッド手法,” 情報処理学会論文誌, Vol.30, No.8, pp.944-953, 1989.
- [84] 横矢, 坂上: “画像理解と最適化原理”, 信学誌, Vol.74, No.4, pp.326-334, 1991.
- [85] Zhang, J. : “The mean field theory in EM procedures for Markov random fields”, IEEE Trans. on Signal Processing, Vol.40, No.10, pp.2570-2583, 1992.
- [86] Zheng, H. and Blostein, S. D. : “Motion-based object segmentation and estimation using the MDL principle”, IEEE Trans. on Image Processing, Vol.4, No.9, pp.1223-1235, 1995.
- [87] Zhu, S. C., Lee, T. S. and Yuille, A. L.: “Region competition: Unifying snakes, region growing, energy/Bayes/MDL for multi-band image segmentation”, Proc. of ICCV’95, pp.416-423, 1995.
- [88] Zhu, S. C. and Yuille, A. L.: “Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation”, IEEE Trans. on PAMI, Vol.18, No.9, pp.884-900, 1996.