

簡単な顔表情モデルに対する数理解析の試み

舟久保 登

Key Words : face expression model, mathematical analysis, principal component analysis, discriminant analysis, independent component analysis

1. まえがき

視覚パターン処理・認識の対象分野として、最近人間の顔表情を取り扱う研究が盛んである¹⁾。伝統的な文字認識、またすぐお隣の顔に基づく個人識別と比べ、そこにはこれまでにない内容の異なった関心があるように思われる。その第1はこれが、いわゆる感性情報に属するものであることであろう。そこで顔表情は確かに画像として呈示されるものであるけれども、その背後には様々で入り組んだ心の状態が存在している可能性があり、それをどう数量化できるかは大きな課題である。また顔表情の持つインタラクティブ性にも注目を引かされる。特に人間の場合、相手との情報交換における顔表情の役割は大きく、その際意識しないかも知れないが、絶えずそれを修正しているであろう。この事柄は将来のロボットなど機械と人間との共生を考える際、必須の問題を提起させる。さらに顔表情が本来的には変化を伴う動画像であることも指摘しておきたい。

さて、この論文を書こうとした目的は二つある。その第1は、学部の授業である「メディアシステム演習Ⅱ」の中の一課題としてパターン識別を扱っているが、この対象に履修学生が興味をひきそうな顔表情を選んだことにある。そのためプログラム内で設定できる少ない画素数の簡単な顔表情標準パターンを用意する必要があった。そして結果的に図1に示すよ

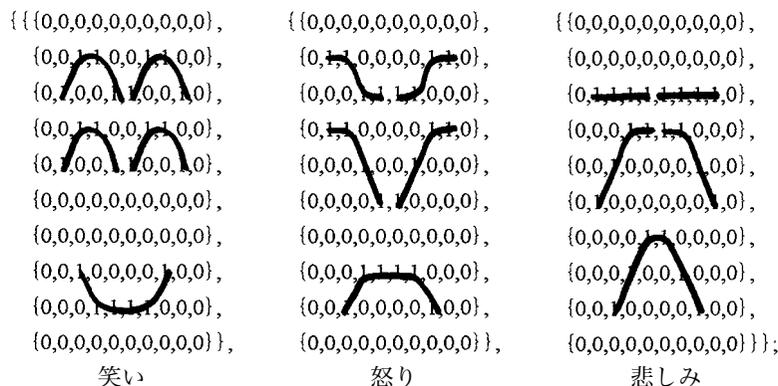


図1 顔表情標準パターン

うな横10×縦10画素から成る笑い，怒り，悲しみの3顔表情を直観的かつほとんど無意識に決めたが，これらがどんな性質を持つようになっているのかに作成者として関心をそそられた次第である。また第2は，筆者のゼミの大学院修士生が修論のテーマに顔表情を取り上げているので，この研究の一つの先行的なひな型を示そうと意図したことである。ただしそちらの取り扱い対象は実際の標準的な6顔表情写真，こちらの方は簡単な三つのモデルというわけで，問題内容やそれが含む複雑さにおいて相当な違いが存在するのは当然である。しかし逆にそうだからこそ，単純なものを使っての試行的な研究も必要といえるであろう。

2. 主成分分析に基づいた数理解析

2.1 対象である顔表情モデル

ここで対象とした顔表情モデルは前頁で図1に示したが，これらについてさらに説明を加えよう。

顔表情が表す感情には，一応典型的なものと考えた笑い，怒り，悲しみの3種を取り上げることとした。そしてこれら各々に対して，横10×縦10画素で画素値は1，0の2値から成るモデル画像を設定したのである。先に述べたようにこれら絵の内容自体は筆者が常識的に作ったもので，さして根拠を持つ絵柄ではないが，パターン識別プログラミングのための標準パターンとして使用するというのもあって，なるべく互いにそれらが重ならないようにしたい気持ちで作成した記憶がある。念のためこの顔パターンを，データ内容としては図1と同じものであるけれども，画像の形で図2に示す。（この結果は残念なことに，図1を作成したときのような自然な顔表情を見せていない。やはり画素数と画素濃度値が少な過ぎた感じであるが，ここではあくまでモデルと割り切って，これを用い議論を進めるとする。）

以下数理解的解析を行うのに，これら顔表情画像全体を対象とする他に，目（左右対称に作ったので右の片目）と口の部分だけでも切出して，取り扱った。この理由は解析の場合を増やして検討の事例を多くしたかった他に，目や口は顔全体に比して絵の画素数が少ないゆえ，後に述べる固有値問題の解法の簡単化法に対する数値的確認にとって，好都合と考えたからである。

さらに通常主成分分析において良くなされるように，対象集合の平均を0とする前処理を

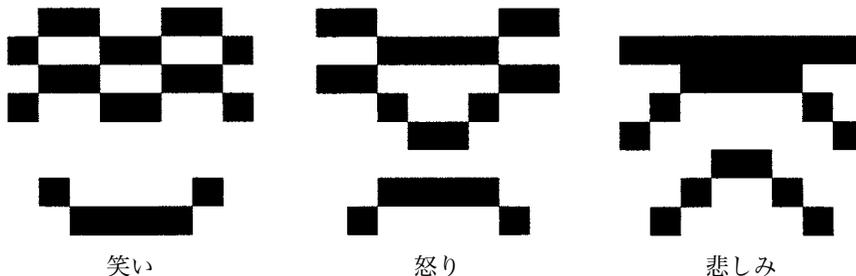


図2 顔表情モデル画像

行っている。したがってこの操作後では、各画像は画素が正負の実数値をとるようになるので、正について赤、負に青の色を用いた画像表現を使う工夫をしている。これは図3の結果となる（図3などカラーの図は最後にまとめてある）。

2.2 目画像モデルに対する数理解析

目（右目）画像モデルは横4×縦3＝12画素から成るもので、図2の顔表情モデル画像の向って左端から2列、上端から4行目から始まる部分領域である。これらを各々下のよう表すことにする。

$$\begin{cases} f_e^{(1)}(i, j) : \text{笑 い (1) の目 (e) 画像} \\ f_e^{(2)}(i, j) : \text{怒 り (2) } \\ f_e^{(3)}(i, j) : \text{悲しみ (3) } \end{cases}, \begin{matrix} i = 1 \sim 4 \\ j = 1 \sim 3 \end{matrix}$$

すると上述した平均画像は

$$\bar{f}_e(i, j) = \frac{1}{3} \times \sum_{k=1}^3 f_e^{(k)}(i, j)$$

であるから、この平均化処理操作後の画像は

$$\begin{cases} g_e^{(1)}(i, j) = f_e^{(1)}(i, j) - \bar{f}_e(i, j) \\ g_e^{(2)}(i, j) = f_e^{(2)}(i, j) - \bar{f}_e(i, j) \\ g_e^{(3)}(i, j) = f_e^{(3)}(i, j) - \bar{f}_e(i, j) \end{cases}$$

となる。

次にこの三つの目画像についての主成分分析を行うために、画素に対する共分散行列を計算する。

$$c_p(m, n) = \frac{1}{3} \times \sum_{k=1}^3 g_e^{(k)} \left(\left\lfloor \frac{m+3}{4} \right\rfloor, m - \left\lfloor \frac{m-1}{4} \right\rfloor \times 4 \right) \cdot g_e^{(k)} \left(\left\lfloor \frac{n+3}{4} \right\rfloor, n - \left\lfloor \frac{n-1}{4} \right\rfloor \times 4 \right)$$

$m = 1 \sim 12, n = 1 \sim 12, \lfloor \cdot \rfloor$ は \cdot の整数部分

しかしこうして得られる共分散行列は、上に記したように12行×12列の行列となり、主成分分析を求める際に必要な固有値問題の解法においては、行列の次元が大きき難しい*。ところが一方この共分散行列は元来笑い、怒り、悲しみの三つの目画像から作り出したものであることを考えると、そのランク（rank, 階数）は3、また対象集合についての平均操作も考慮に入れると、2しかない筈である。そしてこの場合は付録に添付した行列の特異値分解に基づく方法により、画素でなく感情の種類に対する共分散行列を経るやり方を用い、3次元の行列の固有値問題解法に問題を簡単化できるのである²⁾。

* この事柄は顔全体表情画像を後で扱うときに、真に深刻な障害となる。なぜならこの場合の共分散行列は、100（＝10×10）行×100列の本当に大行列となるから。

その場合の目画像に対する共分散行列は,

$$c_c(m, n) = \frac{1}{11} \times \begin{pmatrix} \sum_{i,j} g_c^{(1)}(i, j)^2 & \sum_{i,j} g_c^{(1)}(i, j) \cdot g_c^{(2)}(i, j) & \sum_{i,j} g_c^{(1)}(i, j) \cdot g_c^{(3)}(i, j) \\ \sum_{i,j} g_c^{(2)}(i, j) g_c^{(1)}(i, j) & \sum_{i,j} g_c^{(2)}(i, j)^2 & \sum_{i,j} g_c^{(2)}(i, j) \cdot g_c^{(3)}(i, j) \\ \sum_{i,j} g_c^{(3)}(i, j) g_c^{(1)}(i, j) & \sum_{i,j} g_c^{(3)}(i, j) \cdot g_c^{(2)}(i, j) & \sum_{i,j} g_c^{(3)}(i, j)^2 \end{pmatrix}$$

$$= \begin{pmatrix} 0.162 & -0.081 & -0.081 \\ -0.081 & 0.222 & -0.141 \\ -0.081 & -0.141 & 0.222 \end{pmatrix}$$

となり, この行列についてならば固有値問題は容易に解け*, 以下のように答が得られる.

$$\begin{array}{l} \text{固有値:} \quad \lambda_1 \quad \lambda_2 \quad \lambda_3 \\ \quad \quad \quad 0.363 \quad 0.243 \quad 9.24 \times 10^{-18} \\ \text{固有ベクトル:} \quad \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \\ \quad \quad \quad \begin{pmatrix} 0 \\ 0.707 \\ -0.707 \end{pmatrix} \quad \begin{pmatrix} 0.816 \\ -0.408 \\ -0.408 \end{pmatrix} \quad \begin{pmatrix} 0.577 \\ 0.577 \\ 0.577 \end{pmatrix} \end{array}$$

そこで付録(1)にある行列の特異値分解に基づいた,

$$\mathbf{u}_k = \frac{1}{\sqrt{\lambda_k}} X \mathbf{v}_k$$

なる \mathbf{u}_k と \mathbf{v}_k の間に成り立っている関係式を用いれば,

$$X = \begin{pmatrix} g_e^{(1)}(1,1) & g_e^{(2)}(1,1) & g_e^{(3)}(1,1) \\ \vdots & \vdots & \vdots \\ g_e^{(1)}(4,3) & g_e^{(2)}(4,3) & g_e^{(3)}(4,3) \end{pmatrix} = (\mathbf{g}_e^{(1)}, \mathbf{g}_e^{(2)}, \mathbf{g}_e^{(3)})$$

であるから, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ から $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ が得られることになる。(ただし上記の固有値 λ_3 は λ_1 と λ_2 に比して極端に値が小さく, よって数値計算における誤差と考えられるので無視する。この事柄はいま3感情の目画像を対象としているが, しかし前処理に平均操作を行っているので1次元減って2次元感情空間となっていることを考慮すれば, この点からも納得される。) よって結局

$$\mathbf{u}_1 = \frac{1}{\sqrt{\lambda_1}} X \mathbf{v}_1 = \frac{1}{\sqrt{0.363}} (\mathbf{g}_e^{(1)}, \mathbf{g}_e^{(2)}, \mathbf{g}_e^{(3)}) \begin{pmatrix} 0 \\ 0.707 \\ -0.707 \end{pmatrix} = 1.174 (\mathbf{g}_e^{(2)} - \mathbf{g}_e^{(3)})$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{\lambda_2}} X \mathbf{v}_2 = \frac{1}{\sqrt{0.243}} (\mathbf{g}_e^{(1)}, \mathbf{g}_e^{(2)}, \mathbf{g}_e^{(3)}) \begin{pmatrix} 0.816 \\ -0.408 \\ -0.408 \end{pmatrix} = 0.828 \{2\mathbf{g}_e^{(1)} - (\mathbf{g}_e^{(2)} + \mathbf{g}_e^{(3)})\}$$

* Excelを利用する神田公生氏のフリーソフトを使用した。記して感謝する。

この二つの主成分の画像の形を図4に掲げた。

さらに三つの目について、これら2成分の含有量を算出すると、

$$d_e^{(k,l)} = (\mathbf{u}_k, \mathbf{g}_e^{(l)}) \quad \text{ここで } (\cdot, \cdot) \text{ は内積}$$

$$k = 1, 2 \quad l = 1, 2, 3$$

により、表1のようなになる。そしてこの内容からわれわれには、 \mathbf{u}_2 成分によりまず笑い目が残りの怒り目と悲しみ目から区別され、次に \mathbf{u}_1 成分により \mathbf{u}_2 成分では区別され得なかった怒り目と悲しみ目の分けられることが判明する。図5はこの間の状況を直観的に把握しやすいよう、描いたものである。またそこには、顔表情（感情）は円環状に配置されるという文献にあった記述に則して³⁾、念のためこれら3点を通る円も描き加えてみた。

表1 目画像の成分値

	\mathbf{u}_1	\mathbf{u}_2
$\mathbf{g}_e^{(1)}$	-0.000	4.414
$\mathbf{g}_e^{(2)}$	4.698	-2.207
$\mathbf{g}_e^{(3)}$	-4.698	-2.207

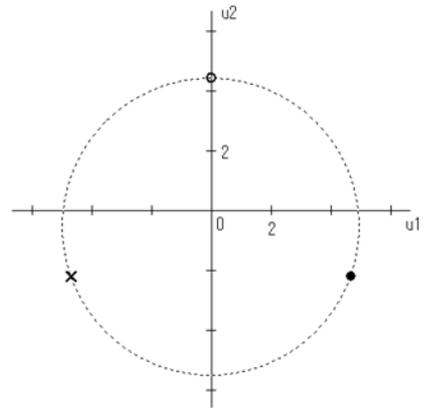


図5 目画像の成分分布
(○ : 笑い, ● : 怒り, × : 悲しみ)

2.3 口画像モデルに対する数理解析

上と同じような解析を、今度は3感情における口部分に対して行った。ここで扱った口部分とは図2の顔画像中より、各々左から3列目、上から7行目の画素から始まる横6×縦3画素部分である。(この範囲は図1および後に示す固有ベクトルの図などからも、容易に読み取れよう。)

さてこれについて平均化操作の前処理をし、共分散行列について簡単化した方法で求めた固有値と固有ベクトルは次のようになった。

$$\begin{array}{l} \text{固有値:} \quad \lambda_1 \quad \lambda_2 \quad \lambda_3 \\ \quad \quad \quad 0.432 \quad 0.118 \quad 1.04 \times 10^{-17} \\ \\ \text{固有ベクトル:} \quad \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \\ \quad \quad \quad \begin{pmatrix} 0.816 \\ -0.408 \\ -0.408 \end{pmatrix} \quad \begin{pmatrix} -7.44 \times 10^{-21} \\ -0.707 \\ 0.707 \end{pmatrix} \quad \begin{pmatrix} 0.577 \\ 0.577 \\ 0.577 \end{pmatrix} \end{array}$$

こうして得られた値は、この場合の共分散行列は違っているにもかかわらず不思議なことに前節の目として次に述べる顔全体のそれと、固有値および固有ベクトルの番号と符号の順は異なっているが、完全に一致している。したがってこれから計算される画素値についての固有ベクトルも式の形は前と同じで、

$$\mathbf{u}_1 = \frac{1}{\sqrt{0.432}} \begin{pmatrix} \mathbf{g}_m^{(1)} & \mathbf{g}_m^{(2)} & \mathbf{g}_m^{(3)} \end{pmatrix} \begin{pmatrix} 0.816 \\ -0.408 \\ -0.408 \end{pmatrix} = 1.242 \{ 2\mathbf{g}_m^{(1)} - (\mathbf{g}_m^{(2)} + \mathbf{g}_m^{(3)}) \}$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{0.118}} \begin{pmatrix} \mathbf{g}_m^{(1)} & \mathbf{g}_m^{(2)} & \mathbf{g}_m^{(3)} \end{pmatrix} \begin{pmatrix} \approx 0 \\ -0.707 \\ 0.707 \end{pmatrix} = 2.058 (\mathbf{g}_m^{(3)} - \mathbf{g}_m^{(2)})$$

である (m は口の意)。しかし勿論これを求める際の固有値自身の値および $(\mathbf{g}_m^{(1)}, \mathbf{g}_m^{(2)}, \mathbf{g}_m^{(3)})$ は別であるゆえ、最終的な $\mathbf{u}_1, \mathbf{u}_2$ の値はそれなりに相違している。これを示したのが図6であり、前節の図4と比較すれば、確かに画像として違っていることは、見ただけで納得できよう。

そこで最後にこの軸に対する三つの感情の口に対する2成分の含有量を求めた。その計算法は前と同様で

$$d_m^{(k,l)} = (\mathbf{u}_k, \mathbf{g}_m^{(l)})$$

であり、この結果を表2に掲げた。それを見ると今度は \mathbf{u}_1 成分により笑いの口が怒りおよび悲しみの口と先ず分けられ、その後で \mathbf{u}_2 成分を用いて怒りの口と悲しみの口が分離されるわけである。このようにこの際使われる成分番号はずれているが、最初3感情を二つに分け、それから分離できなかった2感情を分けるという、同じ階層型分類ができることが面白い (図7参照)。この事実は顔全体の場合も含め、次章の判別分析に活用する。

表2 口画像の成分値

	\mathbf{u}_1	\mathbf{u}_2
$\mathbf{g}_m^{(1)}$	9.104	0.000
$\mathbf{g}_m^{(2)}$	-4.552	-4.116
$\mathbf{g}_m^{(3)}$	-4.552	4.116

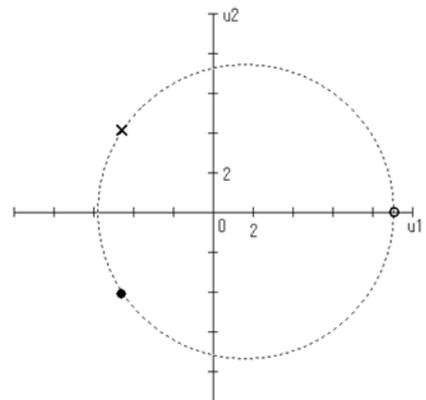


図7 口画像の成分分布
(○ : 笑い, ● : 怒り, × : 悲しみ)

2.4 顔全体画像モデルについての数理解析

これまで顔の中から目と口という、感情の表現を対象とする場合に重要と考えられる部分（部品）を取り上げて解析してきたが、最後のこの節では顔全体をそのまま扱う。これにより加わる情報は、モデルであるため実写のものほど顕著ではないけれども、目と口の関係を含む顔全体における部品の配置具合などである。（笑い泣きの表情他は、これに該当する情報かもしれない。）

さてまずこの3感情の顔画像について、平均化操作の前処理を施す。その結果は先に示した図3に挙げたようなものになる。これからも分るようにこの場合の画像は100画素から構成されており、したがって画素値に対する共分散行列を作成すると100×100の行列となつて、相当大規模ゆえ直接固有値問題を解こうとするのは困難である。したがってこれまでと同様に付録に記した方法で求めることにし、3感情についての共分散行列を用意すると、それは

$$c_f(m, n) = \begin{pmatrix} 0.112 & -0.056 & -0.056 \\ -0.056 & 0.099 & -0.043 \\ -0.056 & -0.043 & 0.099 \end{pmatrix}, \text{ 添字 } f \text{ は顔全体の意}$$

である。そして実際にこの行列の固有値、固有ベクトルは下のようを得られた。

$$\begin{array}{l} \text{固有値:} \quad \lambda_1 \quad \lambda_2 \quad \lambda_3 \\ \quad \quad \quad 0.168 \quad 0.142 \quad 4.62 \times 10^{-18} \\ \\ \text{固有ベクトル:} \quad \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \\ \quad \quad \quad \begin{pmatrix} 0.816 \\ -0.408 \\ -0.408 \end{pmatrix} \quad \begin{pmatrix} 4.03 \times 10^{-20} \\ -0.707 \\ 0.707 \end{pmatrix} \quad \begin{pmatrix} 0.577 \\ 0.577 \\ 0.577 \end{pmatrix} \end{array}$$

前述したようにこの結果は、固有値こそ値は違っているが、固有ベクトルは口と完全に一致し（ただし計算誤差に由来すると解釈される 4.03×10^{-20} は除く）、目とはベクトルの番号のみが順序で入れ換っている。

そこで最後にこの値に基づき、画素値についての固有ベクトルを計算する。（前と同じ理由で、 \mathbf{u}_1 と \mathbf{u}_2 のみである。）

$$\mathbf{u}_1 = \frac{1}{\sqrt{0.168}} (\mathbf{g}_f^{(1)}, \mathbf{g}_f^{(2)}, \mathbf{g}_f^{(3)}) \begin{pmatrix} 0.816 \\ -0.408 \\ -0.408 \end{pmatrix} = 1.991 \{ 2\mathbf{g}_f^{(1)} - (\mathbf{g}_f^{(2)} + \mathbf{g}_f^{(3)}) \}$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{0.142}} (\mathbf{g}_f^{(1)}, \mathbf{g}_f^{(2)}, \mathbf{g}_f^{(3)}) \begin{pmatrix} \approx 0 \\ -0.707 \\ 0.707 \end{pmatrix} = 1.877 (\mathbf{g}_f^{(3)} - \mathbf{g}_f^{(2)})$$

図8はこの二つの成分の画像の形での表示である。

さらに3感情の顔全体についての、これら3成分の含有量を計算して求め、これまでと同様に表3と図9にした。図から分るようにこの分析は口と同じく、まず \mathbf{u}_1 成分により笑いの

モデル顔が怒りと悲しみの2モデル顔と分かれ、その後で u_2 成分により怒り顔が悲しみ顔から分離される。

表3 顔全体像の成分値

	u_1	u_2
$g_f^{(1)}$	33.180	0.000
$g_f^{(2)}$	-16.590	-26.266
$g_f^{(3)}$	-16.590	26.266

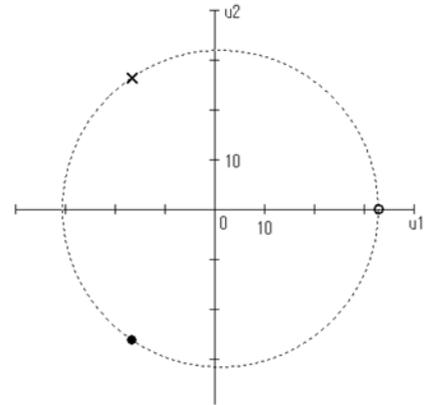


図9 顔全体像の成分分布

(○ : 笑, ● : 怒り, × : 悲しみ)

以上まとめると目、口、顔全体のすべての場合において、第一番目の成分（それは u_1 、ないし u_2 成分のどちらであるかは決められないが）が笑いを分離し、その後で第2番目の成分が残った怒りと悲しみを区別している。つまりモデル画像においては画像パターンとして、第一的には笑いが独立しており、怒りと悲しみは同類に属するというわけである。顔表情は直接は顔面にある各種筋肉の変化により決まり、この量が心の感情から制御されているという図式であろうが、これと上の解析結果との関係はどうだろうか。これについては現実の顔表情に関する分析結果が、真に待たれるところである。

3. 判別分析による数理解析

画像パターンを識別するという目的の場合には、以上議論してきた主成分分析の方法よりも、同じ多変量解析の分野に属する判別分析の手段の適用の方が、直接的と思える。そこでこの章では、これについて若干考察する。

ところでこの判別分析に現われる評価関数中にクラス間分散やクラス内分散なる量が使われているように、元来この分析法は各クラスが多数のメンバーを持っていることを前提としている。しかしながらわれわれの場合、笑い、怒り、悲しみの3クラスにおいて各々ただ一つのモデルパターンしか用意していないので、例えばクラス内分散は計算できない。そこで下のように考えた。

前章のどの場合についても、まず第1段階では笑いと、怒り・悲しみを一体とした2クラスを区別する成分が求められた。よって笑いのモデルの1メンバーから成る第1のクラス、怒りと悲しみの各モデルの2メンバーで構成される第2クラスを想定し、二つのクラスを分ける判別分析問題としたのである。そして2クラス判別分析はその解法を非常に簡単とし、

その判別関数（ベクトル）はFisherにより次式で与えられている。

$$\psi = \frac{1}{S_w^{-1}}(\mu_1 - \mu_2), \quad \begin{array}{l} \mu_i \text{ は } i \text{ クラスの平均ベクトル} \\ S_w \text{ は, クラス内分散行列} \end{array}$$

けれども上のような工夫をしても S_w の行列式の値が0となり逆行列が計算できず、さらに一般化逆行列（Moore-Penrose）に拡張しても、これを求める操作は不可能であった。そんな次第で結局、上式における S_w^{-1} の意味まで遡って検討する必要が生じたのである。

いま上の式において分子は二つのクラスの平均ベクトルの差であり、これは両者の違いを表すベクトルで、これが2クラスを分ける成分となっていることは、直観的にも充分納得される場所である。それでは分母は何を示しているのだろうか。実は普通の（ユークリッド）距離に対してMaharanobisの距離があり、それが持つ意味と同様に、この分母はクラス内分散における分散の大きさにより、ベクトルの要素の各距離を修飾しているのである。つまり同じ距離の値だとしても、その分散が大きかったとしたら、その距離値は小さな分散の場合に比して、効果が少ないというわけである。

このように意味が分かると、ここでの場合1メンバーの第1クラス、2メンバーの第2クラスで統計的分布などないゆえに、分母の項は無視（1とする）して良いであろう。したがって結論的に結果は非常に簡単となり、

$$\psi = \mu_1 - \mu_2$$

である。具体的に前章のデータを当てはめると、

$$\begin{cases} \mu_1 = g_k^{(1)} \\ \mu_2 = \frac{1}{2}(g_k^{(2)} + g_k^{(3)}) \end{cases}$$

で、値は目画像の場合 ($k=l$) に対して表4のようなになる。

表4 目画像の ψ

-0.5	0.5	0.5	-0.5
1	-0.5	-0.5	1
-0.5	0	0	-0.5

そうするとこれは先の u_1 ないし u_2 と定数倍の差を無視して同じである（目の場合では図4の u_2 ）。何のことはない単純なここでのモデルでは、主成分分析と判別分析は同内容の解析結果を与えるものである。

4. 独立成分分析による解析について

主成分分析、判別分析など一連の数理解析の流れの中で、最新の話題は独立成分分析 (Independent Component Analysis) である。そこでここでの対象についてもこの方法による解析を行ってみようとするのは、当然な要求であろう。しかし残念ながら現時点においてこの課題は手を付けた段階に過ぎず、したがってその結果の報告は次の機会に譲らざるを得ない。そこでここではこの方法について筆者が気に懸っている二つの事項について述べておく。

その一つは勉強したところによると、独立成分分析が解析の際に前提とするデータの統計的分布は、通常広く仮定される正規分布でない他の分布であることである⁴⁾。(実はこの条件はもっと強い拘束性を持ち、正規分布の場合には独立成分分析の意味が失せてしまうらしい。) ところでわれわれの際は笑い、怒り、悲しみの3クラスに対して各クラスのパターンが一つずつという、極端に片寄った分布を扱う状況になっている。先の第3章における判別分析の場合は、この分布条件は問題を全く単純化してしまい、結局内容的にはその前の主成分分析と等価にしてしまったが、今度の独立成分分析については如何であろうか。

第2は成分を表す軸に対する含有量を算出する際に、双直交 (陪直交) の概念・性質を用いているらしい事柄である⁵⁾。はるか以前、筆者が初めて文字を対象としたパターン認識の研究を行っていたとき、その特徴抽出・識別のための基盤は有名な飯島のパターン認識理論であったが、その理論では直交性でなく、これを拡張した陪直交性が重要な役割を果たしていた。そんなわけで対象とするパターンを直交の性質を有する主成分などで無理に展開するのではなく、より弱い独立性の要請のみにより解析する独立成分分析には両者にある共通点があると思われ、非常に関心を呼びさまされる。

以上のような次第で、独立成分分析に対する筆者の思い入れは現在強い。

5. むすび

笑い、怒り、悲しみの3感情の顔表情モデルをたまたま作成したのを機会に、その直観かつ偶然的な内容がどうなっているかの数理解析を行った。このためまず主成分分析を使用した。この際必要となる大規模な行列の固有値問題解法に、別の性質から見てその行列の階数が低くなっている場合、行列の特異値分解に基づく簡単化した求め方を、解説・整理した。そして実際に主成分分析を適用した結果は、顔全体、部分である目、口のいずれを対象にしたときにも、まず第1成分による笑いと怒り・悲しみの分離、第2成分により怒りと悲しみの区分という、階層的分離を示すことが得られた。これは勿論直接は顔画像の絵柄に対する解析結果であるが、それを通して快適なないし好ましい感情と不快なないし嫌な感情の対立が反映しているようで、面白い。この事柄は別の驚きなどさらに3感情を加えた修士課程の学生の研究において、さらなる検討が加えられるよう期待している。

続く判別分析においては、結論的に主成分分析と同じ結果が得られた。この事実は対象と

するパターンが各感情当たり一つの計三つと極端に少ないので、当然予期されたものであったが、実際それが出るのを確認できたことは、それなりの意義があったと考える。すなわち、主成分分析はパターン自体の表現性、一方判別分析はその間の分離性と、評価の着眼点に違いがあるけれども、ここでのように単純な場合は、表裏一体となるわけである。

最後の独立成分分析に対しては、興味と期待は大きいのであるが、具体的な定量的議論を述べるに到らなかったことは残念であった。しかし本文中にも記したようにこれについての検討は引き続き行う予定であるので、適当な機会に報告するつもりである。

終りにこの研究の過程で一緒に考え、検討してくれた筆者のゼミの大学院修士生、畠山径久君に謝意を表します。

参考文献

1. 人間とコンピュータによる顔表情の認識 [II], 赤松茂, 電子情報通信学会誌, Vol.85, No.10, pp.766-771 (2002年10月) など
2. Face Recognition: Eigenface, Elastic Matching, and Neural Nets, Jun Zhang et al., Proceedings of the IEEE, Vol.85, No.9, pp. 1423-1435, 1997
3. 顔という知能, 原文雄・小林宏, 共立出版, 2004年8月15日, p.7
4. 独立成分分析, 根本幾・川勝真喜訳, 東京電機大学出版局, 2005年2月10日
5. 独立成分分析, 村田昇, 東京電機大学出版局, 2004年7月10日
6. 逆問題とその解き方, 岡本良夫, オーム社, 1992年3月を参考

付 録

(1) 行列の特異値分解に基づく固有値問題解法の簡単化⁶⁾

主成分分析を行う際には、共分散行列の固有値問題を解かねばならない。しかしこの共分散行列は大規模になることも多く、このときにはその解法に困難が生ずる(顔表情画像全体の画素値に注目した行列は、ここでの場合横100行、縦100列の大きさである)。けれども同じ対象集合の共分散行列の規模が、別の性質についてのものにすると、非常に小さくて済む(例えば顔表情の笑い、怒り、悲しみの3種に対する場合とすれば、固有値問題は3行×3列の行列となる)。以下の内容はこのような状況下で適用できる、固有値問題解法の簡単化である。

X は N 行× M 列の行列とする。ここで N は顔表情画像の画素数、 M は表情数である。すると、

XX^T は画素値に注目した共分散行列で、 $N \times N$ 行列

$X^T X$ は表情に対する共分散行列で、 $M \times M$ 行列

となり、この場合に対する固有値、固有ベクトルを各々、

XX^T の固有値は λ_n 、固有ベクトルは \mathbf{u}_n

$X^T X$ の " λ_m 、固有ベクトルは \mathbf{v}_m

と書くことにしよう。($\{\mathbf{u}_n\}$ 、 $\{\mathbf{v}_m\}$ は正規直交ベクトル)

すると次の式が導けていく。(I は単位行列、添字は次数)

$$\begin{aligned}
 X &= I_N X I_M = \sum_{n=1}^N \mathbf{u}_n \mathbf{u}_n^T \cdot X \sum_{m=1}^M \mathbf{v}_m \mathbf{v}_m^T \\
 &= \sum_{n=1}^N \sum_{m=1}^M \mathbf{u}_n (\mathbf{u}_n^T X \mathbf{v}_m) \mathbf{v}_m^T \cdots \cdots \textcircled{1}
 \end{aligned}$$

さてここで

$$\mathbf{w}_m = \frac{1}{\sqrt{\lambda_m}} \cdot X \mathbf{v}_m$$

とすると, $\{\mathbf{w}_m\}$ は正規直交で, しかも $\{\mathbf{w}_m\}$ は $\{\mathbf{u}_n\}$ (XX^T の固有ベクトル) に一致する. なぜならば,

$$XX^T \mathbf{w}_m = XX^T \left(\frac{1}{\sqrt{\lambda_m}} \cdot X \mathbf{v}_m \right) = \frac{1}{\sqrt{\lambda_m}} X (X^T X \mathbf{v}_m) = \frac{1}{\sqrt{\lambda_m}} X (\lambda_m \mathbf{v}_m) = \lambda_m \left(\frac{1}{\sqrt{\lambda_m}} X \mathbf{v}_m \right) = \lambda_m \mathbf{w}_m$$

ゆえに

$$\mathbf{u}_n^T X \mathbf{v}_m = \mathbf{u}_n^T \sqrt{\lambda_m} \mathbf{u}_m = \sqrt{\lambda_m} \delta_{nm} \quad \delta_{nm} \text{ はクロネッカーの } \delta$$

したがって上の式①に代入すれば,

$$X = \sum_{n=1}^N \sum_{m=1}^M \mathbf{u}_n (\sqrt{\lambda_m} \delta_{nm}) \mathbf{v}_m^T = \sum_{m=1}^M \mathbf{u}_m \sqrt{\lambda_m} \mathbf{v}_m^T$$

この最終的に得られた式を, 行列の特異値分解定理という.

そこでさらに, この式の両辺に右から \mathbf{v}_k を掛ける.

$$X \mathbf{v}_k = \left(\sum_{m=1}^M \sqrt{\lambda_m} \mathbf{u}_m \mathbf{v}_m^T \right) \cdot \mathbf{v}_k = \sqrt{\lambda_k} \mathbf{u}_k$$

この結果

$$\mathbf{u}_k = \frac{1}{\sqrt{\lambda_k}} X \mathbf{v}_k$$

を導けるが, この内容は規模の小さな $X^T X$ の \mathbf{v}_k から, 大規模な XX^T の \mathbf{u}_k を計算できる方法を示している.

(2) 固有値問題解法の簡単化についての数値計算による確認

上述した行列の特異値分解に基づく簡単化が実際どのように働いているかを知るために, 画素に対する共分散行列がそれほど大規模とならない目の場合について (この共分散行列は横12行×縦12列), 直接その固有値問題を計算してみた. この際の共分散行列は本文中の c_p であり, 具体的には下に与えられるものである.

$$c_p = XX^T$$

この結果は少々長いが,

$$\begin{aligned} \mathbf{u}'_1 &= (0.354 \quad 0.354 \quad -0.354 \quad -0.354 \quad 0 \quad -0.354 \\ &\quad 0.354 \quad 0 \quad -0.354 \quad \approx 0 \quad \approx 0 \quad 0.354)^T \\ \mathbf{u}'_2 &= (-0.25 \quad 0.25 \quad 0.25 \quad -0.25 \quad 0.5 \quad -0.25 \\ &\quad -0.25 \quad 0.5 \quad -0.25 \quad \approx -0 \quad \approx -0 \quad -0.25)^T \end{aligned}$$

となって, 簡単化を用いて求めた $\mathbf{u}_1, \mathbf{u}_2$ と

$$\mathbf{u}_1/\mathbf{u}'_1 \approx 3.316, \quad \mathbf{u}_2/\mathbf{u}'_2 \approx 3.312$$

同じ定数倍だけ違っている. この定数倍の差はパターンとしてみた画像にとっては本質的なものではないゆえ, ここで使った簡単化で確かに欲しかった固有ベクトルの得られることが確認できた.

なお厳密に述べるとここでの c_p には全ての要素が0になる行と列があるため, そのままでは固有値問題が解けない. そこでこの0の値の要素は, 他の要素値に比較して充分小さな0.001と置き換えて計算したことを, お断りしておく.

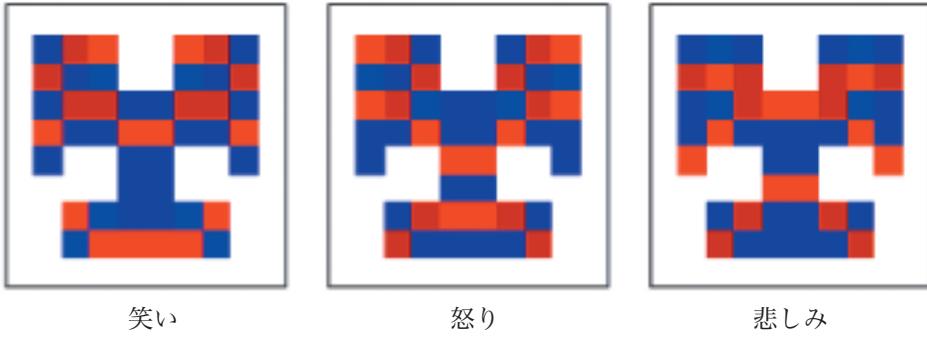


図3 平均0操作後の顔表情画像

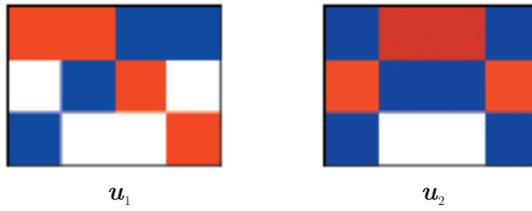


図4 目画像についての u_1 と u_2 成分像

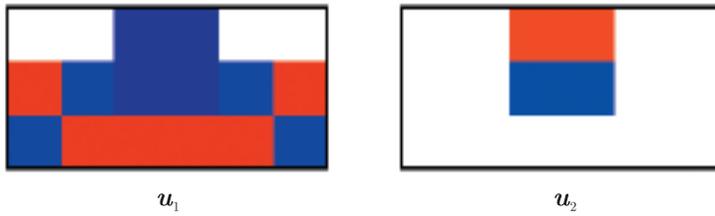


図6 口画像についての u_1 と u_2 成分像

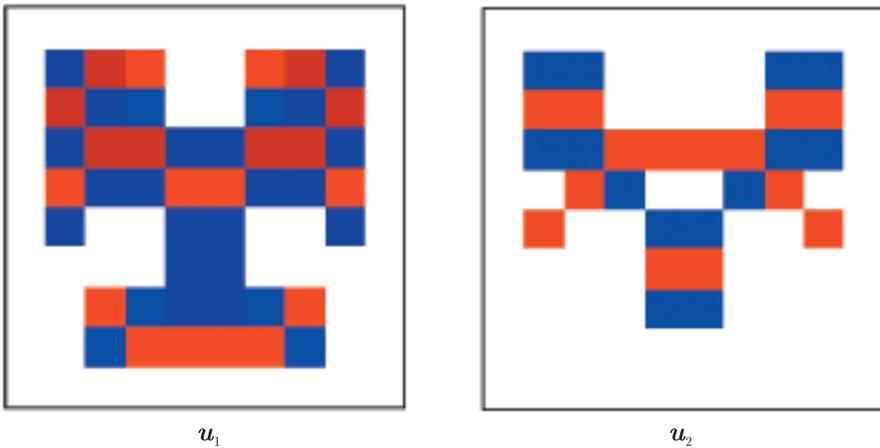


図8 顔全体画像に対する u_1 と u_2 成分像