

# パターン認識における主成分分析 ——顔画像認識を例として——

(株)NTT データ\* 坂野 鋭

( 受付 2000 年 10 月 16 日 ; 改訂 2001 年 4 月 23 日 )

## 要 旨

本稿では、画像パターン認識において、主成分分析がどのように使われているか、また、顔画像認識を例にして主成分分析の方法が画像パターン認識のためにどのような改良を行われているかについて解説する。まず、統計学とパターン認識技術の関係について述べ主成分分析がその中でどのような位置をしめるのかについて概論する。その上で顔画像認識問題の諸課題、具体的には照明変動による顔画像の変動、姿勢変動に起因する非線形性などをとりあげ、それぞれの課題に対する対応策を議論する。

キーワード：パターン認識，コンピュータビジョン，主成分分析，顔画像認識。

## 1. はじめに

### 1.1 パターン認識と主成分分析

パターン認識とは、入力信号に対して何らかのクラス名を割り当てるための技術、ということが出来る。例えば、スキャナから取り込んだ手書き文字やマイクロフォンから入力された音声を変換する技術やカメラから入力された画像から車両を同定する技術に代表される入力信号から対象を判別もしくは識別する技術がこれにあたる。

このような認識機械を実現するためには、スキャナ、カメラ、マイクなどの機器から入力された信号を処理する技術が必須であることは容易に理解されるが、この他に統計的な処理が本質的な役割を果たす(例えば石井 他(1998))。このことは主として、パターン認識の対象が、基本的に定義の不可能なものであることに起因している。例えば、手書き文字の「あ」がなぜ「あ」であるかということについて定義が存在するわけではない。つまり、パターン認識機械は、事前に収集された「実例」から対象に関する知識を「学習」する必要性に迫られる。

統計的な処理はこの「学習」のための中核的な技術として位置付けられており、様々な手法が用いられているが、その中でも、主成分分析は少々特異な地位を得ている。

統計学、特に多変量解析の分野においては主成分分析は多変数データからの説明変数を抽出する手法として用いられるのに対して、パターン認識の分野においては、

- ・ 多次元信号の圧縮
- ・ 信号からの特徴抽出

\* 開発本部 技術開発部 情報設計技術グループ：〒104-0033 東京都中央区新川 1-21-2 茅場町タワービル 7F。

- ・特定カテゴリに関する特徴の抽出

などの様々な目的で用いられる。つまり、単なる統計的な手法としてのみならず、画像信号を処理する仕組みとしても用いられているわけである。以下、顔の認識、すなわち顔画像による個人の識別問題を例として画像パターン認識の分野で行われている主成分分析に関する様々な適用方法及び改良手法について解説する。

顔画像の認識技術は、

- ・入退室管理などの個人認証の手段として(坂野(1999))
- ・ユーザにカスタマイズされたユーザインタフェースの構築
- ・不審人物監視

など様々な応用があり、パターン認識の分野では活発に研究されている領域である。

顔画像の認識に関しては 70 年代から先駆的な研究が行われていたが、研究が活発となったのは 91 年に MIT の M. Turk らが発表した “Recognition Using Eigenface”(固有顔による認識) というちょっと風変わりなタイトルの論文がきっかけであった (Turk and Pentland (1991))。

この論文は主成分分析による次元圧縮と平均値に基づく最近接識別規則を顔画像の認識に適用しただけのものであったが、その影響は大きかった。その意義は、照明、姿勢などの条件を固定すれば顔画像の認識はあまり大きな技術的工夫を必要とせず実現できることを示した点にある。実際、制限された条件の下では、単純に画像を重ねあわせる程度のことでもかなりの性能で個人同定が可能であることが実験的に示されてきた(典型的な例が中山 他(1998)である)。

しかし、研究が進むうちに、

- ・顔の姿勢変動による画像の変化。つまり、カメラに対する顔の方向が変化した場合に発生する顔画像の変化。
- ・表情の変化。
- ・照明条件の変化による画像の変化。

などの課題を克服することが当初考えられていたより困難であることが認知され、多くの研究者を引き付けるチャレンジングな分野となっている。また、先に述べたアプリケーションを実現するためには

- ・登録のために少数のサンプルしか必要としないこと。
- ・一秒程度の処理時間で実行可能であること。
- ・テンプレートのサイズがせいぜい数キロバイトの大きさに収まること。例えば IC カードなどへの応用を考えれば 1 キロバイトが上限となる。

などの要求も存在する。

後者の問題はサンプル数の問題を除き、CPU、記憶装置の性能向上により緩和されていくことが明らかである。しかし、サンプル数問題 (Jain and Waller (1978), Raudys (1981)) は、数百次元から数千次元の高次元特徴空間において数十程度の少数のサンプルで正確な統計量をいかにして推測するか? という困難な課題であり、近い将来の原理的な解決は望めないと考えられる。

また、姿勢、照明、表情などによる変動の問題は計算論的視覚の理論、統計的パターン認識、信号処理の全ての観点からの研究により克服されると考えられる。

これまでこうした問題を解決するためには多くの手法が提案され、テストされているが、それらの多くについては包括的な解説(赤松(1997), Chellapa et al. (1995), Pentland and T. Choudbury(2000))に譲り、ここでは、主成分分析に関連した技術のうち、著者が興味深いと考

える研究を紹介する．本論文では次節で顔画像認識機械の基本動作を解説した上で，次元圧縮，カテゴリからの特徴抽出などへの応用例を紹介し，姿勢変動に伴う非線形分布や照明変動による画像の変化に対応するための技術への主成分分析の応用及び改良手法について解説する．

## 1.2 顔画像認識機械の基本動作

前節でも述べた通り，パターン認識の分野においては主成分分析は様々な場面で色々な問題を解決するために用いられる．具体的にパターン認識における主成分分析の使われ方を解説する前に，顔認識機械の基本的な動作について説明しておこう．

標準的な顔認識機械は

- ・ CCD カメラなどで実現される画像入力部
- ・ 入力画像から対象部分を抽出したり，雑音除去処理，記述の変換などを行う画像処理部
- ・ 画像処理部から送られた信号を識別する認識部
- ・ 登録用顔データからテンプレート(後述)を生成する学習部
- ・ CRT ディスプレイなどで構成される認識結果の出力部

からなる．登録時には，画像入力部から入力された画像は画像処理部でなんらかの処理により，対象とする顔部分画像の分離，顔位置，大きさの正規化，雑音除去，あるいは方向性抽出フィルタなどによる変換の処理を受け，特徴ベクトル  $\vec{x}$  に変換される．このときの特徴の記述方法は画素そのものである場合もあり，Gabor filter などによる方向性記述への変換などの処理が行われる場合もある．カメラなどから入力された複数の顔画像は画像処理部で，顔画像部分の分離などの処理を受け，学習部で，テンプレートと呼ばれる各個人を代表するデータに変換される．このときにテンプレートとして用いられるデータは，平均値などの統計量であることもあり，人工神経回路網の重みである場合もある．

認識時には，画像入力部から入力された画像は画像処理部で登録時と同様の処理により，対象とする顔部分画像の分離など，学習時と同様な画像処理による前処理を受けた後，特徴ベクトル  $\vec{x}$  に変換される．認識部では事前に登録された認識辞書，もしくはテンプレートと呼ばれているデータと  $\vec{x}$  との比較が行われる．つまり何らかの尺度による類似度が計算され，類似度の最も大きなテンプレートが属するカテゴリ名が認識結果として出力される．

例えば最も簡単な例としては画素そのものを特徴として扱い，学習データの単純平均を認識辞書として持ち，認識辞書と入力の二乗誤差を類似度(というより，この場合は相違度だが)として用いる様な認識系を考えることが出来る(本稿では，この方法を画像マッチングと呼ぶことにする)．

この認識系の場合には，単純にテンプレートを点として表現しているわけだが，データの分布形状によっては点だけでは表現できない場合は数多く見られる．主成分分析及びその改良手法は，この問題を解決するための方法としてしばしば用いられる．また，テンプレートを単純な点として表現した場合でも平均値の様な統計量の計算精度をどのように向上させるかは重要な課題である．主成分分析はこの問題の補助的な手段として用いられる場合も多い．以下，これらの問題に対して，主成分分析がどのように用いられるかを概観する．

なお，本論文では，解説する手法の効果を示すために 2 種類のデータを用いて実験を行う．使用するデータは以下の 2 種類であり，節の主題に応じて適切に使用するものとする．

NTTdata (株)NTT データで作成した顔画像データ．正面，無表情，照明条件一定で，任意の数字を発声している．人数は 79 人で，学習用 10 枚，テスト用 10 枚のデータを手動で切り出し， $16 \times 12$  の画像に正規化している．このデータは目の位置で顔位置と大きさの正規化を施してあるほか，変動の激しい口部分を切り取ってあるので，比較的認識の簡単なデータと



図 1. NTTdata 顔画像データの例 .

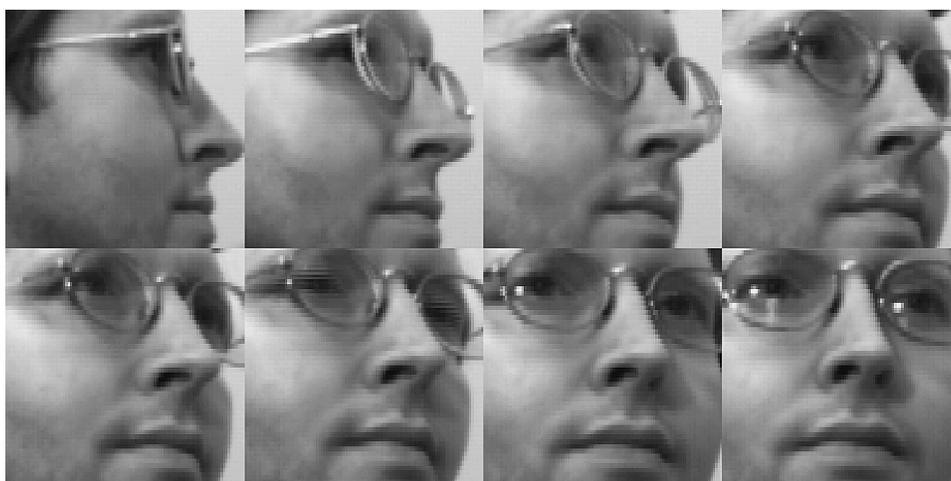


図 2. UMIST 顔画像データベースの例 .

いうことが出来る . 図 1 に画像の例を示す .

UMIST マンチェスタ大学情報理工学研究所で作成されたデータである (Graham and Allinson (1998)). 正面から横顔までの各方向から照明条件一定で撮影されており, 20 人の人物を含み, 一人あたり 40-60 枚の画像で構成される . 実験のためには, このうち 15 人の画像を対象に様々な方向の 10 枚の画像を学習データとし, 残りの 3-50 枚をテスト用とした . なお, このデータからも顔画像を手動で抽出し,  $15 \times 15$  のサイズに正規化してある . ただし, このとき目などの顔部品による正規化は行っていない . 図 2 に画像の例を示す .

2. 顔画像認識への主成分分析の応用

2.1 次元圧縮——もしくは「顔であること」の特徴抽出

前章でも述べた通り，最も単純な顔画像認識技術は，画像そのものの重ね合わせ，つまり，登録時に何らかの方法で作成した個人の標準的な顔画像と入力された顔画像の画素ごとの自乗誤差を計算し，誤差の最も小さな個人を回答として出力する方法である．この原理に基づく認識系においては，処理量の削減や，識別のために有効な特徴の抽出もしくは選択のために特徴ベクトルの次元数を削減する処理が用いられることが多い．特徴量の削減のためには，テーブルを用いて指定した特徴量のみを選択する特徴選択法の他，判別分析や主成分分析の様な次元圧縮法が用いられる．

一般に特徴選択法にはよい設計アルゴリズムが無いために (Jain et al. (2000))，実用的な認識系では判別分析もしくは主成分分析が用いられる事が多い．光学的文字読み取り装置などの場合には，入力されるカテゴリが既知であるために判別分析が有効とされるのに対して，顔画像認識を始めとする個人識別問題では本人と未知の人を識別することを要求されるため，主成分分析が事実上唯一の次元圧縮法となる．

こうした方向の研究では前出の M. Turk による Eigenface (以下，固有顔)の方法が有名である．Turk による固有顔の方法は，最初に顔画像を画素値を成分とするベクトルとみなし，主成分分析を行い，顔画像の特徴をもった比較的低次元の部分空間を得，認識処理をこの部分空間の中で行う．

つまり，最初に顔を記述する部分空間を作るための様々な人物の顔を含む  $M$  個の学習データ

$$X_{all} = \{\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_M\}$$

から平均値

$$\vec{\mu}_{all} = \frac{1}{M} \sum_{i=1}^M \vec{x}_i$$

及び共分散行列

$$C_{all} = \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^M (\vec{\mu} - \vec{x}_i)(\vec{\mu} - \vec{x}_j)^T$$

を計算し，固有値問題

$$(2.1) \quad C_{all} \vec{v} = \lambda \vec{v}$$

を解くことで固有値  $\lambda_j$ ，固有ベクトル  $\vec{v}_j$  を得る．ここで添字  $j$  は固有値の大きな順に並べるものとする．(なお，ここでは用いられる平均値，共分散行列などには様々な人物の顔が含まれる．理想的には全ての人類の顔を用いることが望まれるため， $_{all}$  の添字を用いた．)

次に，個人の登録を行う．登録フェーズでは個々人につき複数枚(ここでは  $m$  枚としよう)の顔画像

$$X^k = \{\vec{x}_1^k, \dots, \vec{x}_i^k, \dots, \vec{x}_m^k\}$$

を用意し，上位  $d$  次元の固有ベクトルを用いて作った射影行列  $\eta = \{\vec{v}_1, \dots, \vec{v}_d\}$  を乗ずることで  $d$  次元のベクトル  $\{\vec{\xi}_1^k, \dots, \vec{\xi}_m^k\}$  を作り，その平均値  $\vec{\mu}^k$  をテンプレートとして登録する．ただし，ここで上付き添字  $k$  は個々人の ID 番号を示すものとする．

認識時には，入力された顔であることのみがわかっている画像  $\vec{x}$  と射影行列を乗ずることで圧縮されたベクトル  $\vec{\xi}$  を作り，登録されているテンプレート  $\vec{\mu}^k$  との自乗誤差

$$(2.2) \quad D^2 = \|\vec{\mu}^k - \vec{\xi}\|^2$$

表 1. 固有顔法, 画像マッチング法における認識率 (NTTdata 1) .

| 認識法   | 画像マッチング | 固有顔   | 固有顔   | 固有顔   | 固有顔   | 固有顔   |
|-------|---------|-------|-------|-------|-------|-------|
| 次元数   | 192     | 1     | 5     | 10    | 20    | 30    |
| 累積寄与率 | -       | 48.2% | 70.9% | 81.7% | 89.7% | 93.1% |
| 認識率   | 92.23%  | 18.1% | 84.3% | 94.8% | 96.7% | 97.1% |

を計算し, 自乗誤差が最小となった  $k$  を回答として出力する .

この方法は主成分分析を次元数の削減に用いた例である . この様な適用法に関しては次元数の削減は計算量とテンプレートのメモリ容量の削減に効果があり, 認識率は顔画像の圧縮率とのトレードオフになるという誤解をよく受けるが, 事実として, 主成分分析による次元圧縮を用いた場合には, 一般に認識率の向上が見られる . このことを示す実験結果を表 1 に示す .

表に示された通り, 部分空間次元数 10 に達した段階で既に画像マッチングをしのぐ認識率を得ている . この様な現象が発生する原因は, 1) 次元の圧縮に伴う平均値の推定精度の向上, 2) 主成分分析により, 顔の識別に関係の無い情報が落ちたこと, などいくつか考えられるが, 対象による依存性があることもあり, 明確な理由はわかっていない . いずれにせよ, 直感的には顔画像認識のために有効な特徴が抽出されたものとみなすことが出来, その意味で主成分分析による次元圧縮は「顔であることの特徴抽出」と理解されている .

また, 顕わに識別能力の高い特徴を抽出する次元圧縮法として知られている判別分析を顔画像認識に適用する方法も検討されているが (Belhumer and Kriegman (1996)), 先に述べた通り, 多くのアプリケーションでは未知の人物と本人を識別することが要求されるためあまり広くは用いられていない .

ここで述べた通り, 一定の照明条件の下で顔画像の入力を正面の画像に限定し, 顔の位置や大きさが正確に正規化されるのであれば, 固有顔の方法で個人の同定がほぼ可能であることが示された .

しかしながら, 顔は 3 次元物体であるため, ある程度の姿勢変動があった場合には明らかにこれらの条件は満たされない . 例えば, 10 度から 20 度以上の傾きが生じた場合には鼻による遮蔽などが発生するため目の位置などで正規化を行ったとしても単純な画像マッチングで認識が困難になるのは明らかである .

この様な問題に対する最も簡単な対処法は様々な角度から撮影した顔画像を学習データとして用いることである . しかしながら, 固有顔法では単に平均値一点のみをテンプレートとして用いているために, 姿勢変動によって発生した顔画像の分布を近似することが困難である . この様な場合にはさらに複雑なデータの広がり表現できるテンプレートを用いることが必要である . 主成分分析を用いたこの様な例としては部分空間でカテゴリごとの分布を近似する部分空間法がある . 次節では部分空間法を用いた顔の認識について解説する .

## 2.2 個人の顔の特徴抽出 — 部分空間法

複雑な形状に分布したデータを表現する方法は, パターン認識の世界では長らく中心的な研究課題の一つであった . 最も初期には全ての学習データをそのままテンプレートとする  $k$ -最近接法, ポテンシャル関数法などのノンパラメトリック法に始まり, 90 年前後の多層パーセプトロンの研究, 最近の SVM (Support Vector Machine) の研究など枚挙に暇が無い (例えば, Duda et al. (2001)).

部分空間法は Watanabe によって提案されたパターン認識独特の主成分分析の利用法である (Watanabe and Pakvasa (1973)). 部分空間法においては, 登録時に与えられた複数の学習データに対して主成分分析を施すことで学習データの分布を近似する部分空間を計算し, 未知の入

カデータと部分空間の類似度を計算することで認識処理を実行する。

このときの部分空間の構成法としては、いわゆる主成分分析、すなわち共分散行列の対角化によって得られた主成分を基底とする部分空間を用いるもの、共分散行列の代わりに自己相関行列を用いるもの、あるいは、初期の部分空間を前記のどちらかの方法で作リ、部分空間を誤り訂正過程で修正して行くものなど多岐にわたる（オヤ(1986)）また、部分空間と入力データとの類似度（もしくは相違度）としては、角度、部分空間への射影の大きさ、自乗誤差、あるいは固有値を考慮した誤差など様々なものが用いられる（黒沢(1996)）。また、角度を類似度として用いた部分空間法は特に複合類似度とも呼ばれる（飯島(1989)）。

こうした様々な部分空間法は、色々なシチュエーションで使用されているが、ここでは最も古典的な Watanabe の CLAFIC を取り上げる。

CLAFIC は部分空間の構成に自己相関行列を用い、認識時には、データを部分空間に射影したときの大きさを類似度として用いる手法である。即ち、登録時には個々人につき複数枚（ここでは  $r$  枚としよう）の顔画像

$$X^k = \{\vec{x}_1^k, \dots, \vec{x}_i^k, \dots, \vec{x}_r^k\}$$

を用意し、平均値

$$\vec{\mu}^k = \frac{1}{r} \sum_{i=1}^r \vec{x}_i^k$$

及び自己相関行列

$$C^k = \frac{1}{r} \sum_{i=1}^r \sum_{j=1}^r \vec{x}_i^k \vec{x}_j^{kT}$$

を計算し、固有値問題を

$$(2.3) \quad C^k \vec{v}^k = \lambda^k \vec{v}^k$$

を解くことで個々人毎に固有値  $\lambda_j^k$ 、固有ベクトル  $\vec{v}_j^k$  を求め、例えば上位  $q$  個の固有ベクトルをテンプレートとして登録する。ただし、ここでも上付き添字  $k$  は個々人の ID 番号を示し、添字  $j$  は固有値の大きな順に並べるものとする。

CLAFIC における類似度  $s$  は

$$(2.4) \quad s = \frac{1}{|\vec{x}|} \sum_{i=1}^q (\vec{v}_i \cdot \vec{x})$$

で定義される。ここで  $q$  は認識に用いる部分空間の次元数であり、累積寄与率などから決定される。

この様に部分空間法は主成分分析で求められた固有ベクトルで形成された部分空間でデータの分布を近似し、入力されたデータと、部分空間の類似度を計算することで認識処理を実施する。つまり、データの分布の変動の少ない成分を比較の尺度とする興味深い手法である。しかしながら一方で、この様な方法で激しい姿勢変動を含むような顔画像をうまく近似出来るのは明らかではない。

大きな姿勢変動がある場合の部分空間法の有効性を示すために UMIST データでの実験結果を表 2 に示す。なお、使用する部分空間の次元数は個々人により一定とし、認識率が最大となる次元数を表示した。固有顔についても認識率が最大となる結果を示した。画像マッチング法や固有顔法ではかなり低い認識率になっているのに対して、部分空間法では高い認識率が得られている。この結果から部分空間によるデータの近似が、様々な方向の顔データの表現に有効であることがわかる。

表 2. 部分空間法, 固有顔法, 画像マッチング法における認識率 (UMIST) .

| 認識法 | 画像マッチング | 固有顔   | 部分空間法 |
|-----|---------|-------|-------|
| 次元数 | 192     | 40    | 5     |
| 認識率 | 54.2%   | 53.9% | 95.7% |

しかしながら, 1 章で述べたアプリケーションを実用化するためにはさらに高い認識率が求められる. 特に, 本論文で用いている様な小規模データではほぼ 100% の認識率が要求される. 次章では高精度化のためにパターン認識の分野で行われている主成分分析の改良手法を中心に解説する.

### 3. さらに高精度な認識系の実現のために

前章までの解説の通り, パターン認識の分野では, 主成分分析もしくはその主たる応用である部分空間法は重要な技術であり, 文字認識, 顔画像認識の分野を中心に積極的な改良が進められてきた. 本節では顔画像認識の分野で行われてきた主成分分析に関連する技術の改良手法を解説する.

#### 3.1 相互部分空間法

主成分分析を用いた手法の中で筆者が最も興味深いと考える研究は山口らによって発表された相互部分空間法を用いた顔認識技術である (山口 他 (1997), 前田・渡辺 (1985)). 前節で述べた通り, 通常の部分空間法は, 登録時に与えられた複数の学習データに対して主成分分析を施すことで部分空間を計算し, 未知の入力データとの角度, 距離などの類似度を計算することで認識処理を実行する. 即ち,  $\vec{v}_i$  を認識辞書となる主成分だとすると前節で述べたとおり, CLAFIC における類似度  $s$  は

$$s = \frac{1}{|\vec{x}|} \sum_{i=1}^q \vec{v}_i \cdot \vec{x}$$

で定義される. ここで  $q$  は認識に用いる部分空間の次元数である.

一方で, 相互部分空間法では学習データから主成分分析で主成分を計算するところまでは同じであるが, 認識時には, 複数の顔画像を入力画像とし, これらの入力画像に対して主成分分析を施し, 登録された主成分との類似度を計算することで認識処理を実行する.

前田の定式化では部分空間同士の類似度は角度で定義され, 入力画像列から得られた主成分を  $\vec{v}_j'$  とすると部分空間同士の角度は, 行列

$$(3.1) \quad X_{ij} = \sum_{l=1}^q (\vec{v}_i' \cdot \vec{v}_l) (\vec{v}_l \cdot \vec{v}_j')$$

の固有値問題

$$(3.2) \quad \lambda \vec{x} = X_{ij} \vec{x}$$

を解くことにより最大固有値として与えられる (前田・渡辺 (1985)).

つまり, 登録時の画像と認識時の画像双方の変動の少ない部分同士を比較する事で認識率の向上を図れる優れた手法であり, 実験的にも絶大な威力を示す. ただし, この方法は, 与えられた顔画像データの分布が非線形であった場合には, 十分な精度を達成できないという問題を持つ.

この問題は相互部分空間法に限らず、線形主成分分析を用いた認識アルゴリズムに共通の課題である。特に顔画像認識の場合、顔の姿勢が大きく変動する際には、図 3(詳細は後述)に示す様に強度の非線形性が存在することが知られている。

この様な非線形性の問題を解決するために、パターン認識の分野では様々な技術が研究開発されてきた。次節では、これらの技術と顔の認識への応用に関して概観するとともに、著者らの最近の仕事である、相互部分空間法と非線形主成分分析の融合した新しい顔画像認識技術を解説する。

### 3.2 核非線形相互部分空間法

前節で述べた物体の姿勢変動に起因する強度の非線形性の問題を解決するために様々な方法が提案されてきた。例えば、村瀬らは画像登録時に、様々な角度で撮影するとともに撮影したときの角度を記録し、これをパラメータとした回帰分析の問題として扱うことで姿勢変動に伴う非線形性を記述した(村瀬・Nayar(1994), 村瀬(1995))。しかし、姿勢変動に不変な顔の認識を考えた場合、何らかのパラメータが与えられることは考えづらく、非線形主成分分析の手法を考える事が必要となる。

従来、非線形主成分分析に関しては統計学、人工神経回路網の分野で活発な研究が行われてきた。統計学の分野では高次モーメントを共分散行列に何らかの形で組み入れるという戦略が取られているが(例えば Gnanadesikan(1977) Saito and Kariya(1988))人工神経回路網の分野では入力と出力の多層パーセプトロンを結合したいわゆる砂時計型ネットワークのアプローチ(Diamantaras and Kung(1996), Oja(1982)), が取られている。

これらの方法を表情や姿勢の変動のあるデータに適用した例としては、EM アルゴリズムにより表情変動に対応した複数の部分空間を作ることを目標としたアプローチがある(Frey et al.(1998))。実験結果は Turk 流の Eigenface より高性能を示しているが、表情ごとの部分空間を作ることに成功しているかどうかかわからないなど問題も多い。また、Bar et al.(1998)は非線形写像を用いて表情変動に不変な部分空間を構成する手法を提案しているが、アルゴリズム自体がまともに構成されておらず、有為な結果とは言い難い。また、Moghaddam(1999)は砂時計型パーセプトロンを用いた非線形主成分分析を部分空間の構成に利用し、線形部分空間法より認識精度が落ちるという結果を示しているが、この実験では認識対象を正面顔、無表情に限定しており、非線形構造がないデータに対して非線形主成分分析を適用すると性能低下が起こるという経験に対する典型的な例となっている。

こうした非線形主成分分析に関する研究で興味深いものの一つが Schölkopf et al.(1998)により提案された核非線形主成分分析である。(Schölkopf は Kernel PCA と命名している。前田・村瀬(1999)はカーネル非線形主成分分析という訳を与えているが著者はカーネルに積分核の意味での「核」の日本語訳をあてた。)

核非線形主成分分析は一言で言うと特徴ベクトルを無限次元もしくは有限であっても極めて高次元の関数空間に非線形写像した後で特異値分解の方法で(データ数  $-1$ )次元の主成分を抽出する手法である。関数空間での内積を計算する核関数が存在する場合にはこの特異値分解は有限時間で計算することが出来、陰的にはあるが固有ベクトルが明確に定義されるという意味で興味深い非線形主成分分析の方法である。

この方法には、他の人工神経回路網的アプローチと異なり、固有ベクトルが明確に定義されるため部分空間法や相互部分空間法への応用が簡便であるという優位性がある。以下では Schölkopf et al.(1998)に従い核非線形主成分分析を導出し、これを部分空間法に適用する方法、及び、相互部分空間法を導出する。

近年、SVM に関する研究が盛んに行われているが、SVM に代表されるカーネル法は歴史的

にはパターン認識の古典的な手法であるポテンシャル関数法にその起源を持つ(アイゼルマン他(1978)). ポテンシャル関数法の基本的なアイデアは, ガウス関数の様な「ポテンシャル」関数によってデータをばかしてしまうことで, 有限なデータから確率密度関数を近似することである. 例えば, カテゴリ  $k$  についての  $m$  個の学習データ  $\{\vec{x}_1^k, \dots, \vec{x}_j^k, \dots, \vec{x}_m^k\}$  が与えられた場合, 入力データ  $\vec{x}$  を識別するプロセスは, 入力データと全学習データについてのポテンシャル関数

$$(3.3) \quad s^k = \sum_{i=1}^m k(\vec{x}_i^k, \vec{x})$$

を計算し,  $s$  を最大にする  $k$  を回答として出力することである. つまり, この関数が密度関数の近似として機能する.

ここで用いられるポテンシャル関数は, 二つの引数を持つ意味での核関数であるが, こうした核関数が正定値で対称である場合, Mercer の定理により,

$$(3.4) \quad k(x, y) = \sum_{i=1}^{\infty} \frac{\psi_i(x)\psi_i(y)}{\lambda_i}$$

なる級数が存在し, 各点で必ず収束する(証明は近藤(1953)). これをさらに

$$(3.5) \quad k(x, y) = \sum_{i=1}^{\infty} \frac{\psi_i(x)\psi_i(y)}{\lambda_i} = \sum_{i=1}^{\infty} \frac{\psi_i(x)}{\sqrt{\lambda_i}} \frac{\psi_i(y)}{\sqrt{\lambda_i}} = (\Psi(x) \cdot \Psi(y))$$

の様に書くと, 核関数を用いた計算が無次元もしくは極めて高次元の空間での内積の計算を行っていることに対応することがわかる. Schölkopf の核非線形主成分分析は, この空間での主成分分析を定式化することで, 非線形主成分分析を実現した手法である.

核非線形主成分分析を導入するために形式的に関数空間への非線形写像

$$(3.6) \quad \Psi: \mathcal{R}^N \rightarrow \mathcal{F}, \quad \vec{x} \mapsto X$$

を考える. すると, 空間  $\mathcal{F}$  での共分散行列に対応する線形演算子は

$$(3.7) \quad \bar{C} = \frac{1}{m} \sum_{j=1}^m (\Psi(\vec{x}_j)\Psi(\vec{x}_j)^T)$$

と書ける. 特に  $\mathcal{F}$  が無限次元の場合には  $\Psi(\vec{x}_j)\Psi(\vec{x}_j)^T$  は  $\mathcal{F}$  上の線形演算子

$$(3.8) \quad X \mapsto \Psi(\vec{x}_j)(\Psi(\vec{x}_j) \cdot \vec{X})$$

となる. 但し, ここで  $\mathcal{F}$  は  $L_2$  の関数空間

$$(3.9) \quad \infty > (\Psi(\vec{x}) \cdot \Psi(\vec{x}))$$

であることを仮定している.

$\mathcal{F}$  上での主成分  $V \in \mathcal{F} \setminus \{0\}$  を考えると, これは行列  $\bar{C}$  の固有値問題

$$(3.10) \quad \lambda V = \bar{C}V$$

の解となる.  $\mathcal{R}^N$  の時と同様に, これは

$$(3.11) \quad \lambda(\Psi(\vec{x}_k) \cdot V) = (\Psi(\vec{x}_k) \cdot \bar{C}V)$$

と書ける. また, 主成分はデータの線形結合であるから, ある係数  $\alpha$  により

$$(3.12) \quad V = \sum_{i=1}^m \alpha_i \Psi(\vec{x}_i)$$

となることは明らかである．

式 (3.11)(3.12) を連立すると，

$$(3.13) \quad \lambda \sum_{i=1}^m \alpha_i (\Psi(\vec{x}_k) \cdot \Psi(\vec{x}_i)) = \frac{1}{m} \sum_{i=1}^m \alpha_i \left( \Psi(\vec{x}_k) \cdot \sum_{j=1}^m \Psi(\vec{x}_j) (\Psi(\vec{x}_j) \cdot \Psi(\vec{x}_i)) \right)$$

を得る．

ここで新しい  $m \times m$  行列

$$(3.14) \quad K_{ij} = (\Psi(\vec{x}_i) \cdot \Psi(\vec{x}_j))$$

を定義すると，式 (3.13) は

$$(3.15) \quad \begin{aligned} \lambda \sum_{i=1}^m \alpha_i K_{ki} &= \frac{1}{m} \sum_{i=1}^m \alpha_i \left( \Psi(\vec{x}_k) \cdot \sum_{j=1}^m \Psi(\vec{x}_j) K_{ji} \right) \\ &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^m \alpha_i K_{kj} K_{ji} \end{aligned}$$

ここで  $\alpha_i$  は結合係数を並べたものなのでベクトルとみなすことが出来て，

$$(3.16) \quad m\lambda\alpha K = \alpha K^2$$

と書ける．両辺に  $K^{-1}$  を乗ずると

$$(3.17) \quad m\lambda\alpha = \alpha K$$

なる固有値問題を解く事によって  $\alpha$  の具体的な表式が得られることがわかる．これは特異値分解などの既知の方法で計算することが出来る．

固有ベクトル  $\alpha$  が求まると，任意の特徴ベクトル  $\vec{x}$  から  $V$  によって作られる部分空間への写像は

$$(3.18) \quad V \cdot \Psi(\vec{x}) = \frac{1}{\lambda} \sum_{i=1}^m \alpha_i (\Psi(\vec{x}_i) \cdot \Psi(\vec{x}))$$

によって計算できることがわかる．

但し， $\Psi(\cdot)$  は一般に極めて高次元，時には無限次元のベクトルであるため， $\Psi(\vec{x}) \cdot \Psi(\vec{y})$  を計算するためには極めて長時間もしくは無限の時間を必要とする．しかしながら，ここで考えている  $\Psi(\cdot)$  の内積は式 (3.5) を満足することがわかっている．従って  $\Psi(\vec{x}) \cdot \Psi(\vec{y})$  は単に関数  $k(\vec{x}, \vec{y})$  を計算することに帰着される．従って，

$$(3.19) \quad V \cdot \Psi(\vec{x}) = \frac{1}{\lambda} \sum_{i=1}^m \alpha_i k(\vec{x}_i, \vec{x})$$

の様になり，高次元の固有ベクトル  $V$  を顕わに求めなくても，その写像を計算することが出来るようになる．これにより，非線形写像を施したあとの関数空間での主成分分析のアルゴリズムが与えられたことになる．

なお，核非線形主成分分析は，核関数  $k(x_i, x_j)$  が等方性を持ち，データ間の距離  $\|\vec{x}_i - \vec{x}_j\|$  を用いて  $k(\vec{x}_i, \vec{x}_j) = r(\|\vec{x}_i - \vec{x}_j\|)$  と書ける場合には，原空間での距離  $\|\vec{x}_i - \vec{x}_j\|$  を

$$\Delta_{ij} = \sqrt{2(1 - r(\|\vec{x}_i - \vec{x}_j\|))}$$

と非線形変換したものを距離として布置を求める計量的多次元尺度構成法と対応付けられる (Williams (2001))．これについて，Williams は kernel MDS (kernel MultiDimensional Scaling; 核非線形多次元尺度構成法) という用語を提唱している．詳細については原論文を参照されたい．

また核非線形主成分分析で求められた部分空間を認識辞書に用いたパターン認識法が津田 (1999), 前田・村瀬 (1999) により提案されたカーネル非線形部分空間法であり, 実験的に高い識別性能が示されている.

次に, 相互部分空間法への適用を考える. 相互部分空間法への適用のためには辞書として登録されている部分空間と入力画像系列から計算される部分空間の角度を有限時間で計算出来ることを示せば良い.

これまで用いてきた関数空間での固有ベクトル  $V$ , 固有値  $\lambda$ , 学習データ  $\vec{x}_i$  を辞書側と呼び, 認識対象として入力された画像系列を  $\vec{x}'$ , 関数空間での固有ベクトルを  $W$ , 固有値  $\nu$ , と書き, 入力側と呼ぶことにする. つまり,  $V$  を「辞書側固有ベクトル」,  $W$  を「入力側固有ベクトル」, などと呼ぶことにする.

$V, W$  はそれぞれ辞書側, 入力側データから計算された基底ベクトルであり, ノルムが正規化されていると仮定すると, 辞書登録された画像群と入力された画像系列の類似度を評価するためには  $(V \cdot W)$  の値の大きさを評価すればいいことになる.  $(V \cdot W)$  は  $\mathcal{F}$  上の積分であるから, この値を顕わに有限の時間で計算することは出来ない. しかし, 式 (3.12) を用いて  $V \cdot W$  の表式を書くと

$$(3.20) \quad W = \sum_{j=1}^{m'} \alpha'_j \Psi(\vec{x}'_j)$$

より

$$(3.21) \quad \begin{aligned} V \cdot W &= \left( \sum_{i=1}^m \alpha_i \Psi(\vec{x}_i) \cdot \sum_{j=1}^{m'} \alpha'_j \Psi(\vec{x}'_j) \right) \\ &= \sum_{i=1}^m \sum_{j=1}^{m'} \alpha_i \alpha'_j (\Psi(\vec{x}_i) \cdot \Psi(\vec{x}'_j)) \end{aligned}$$

である. ここで再び式 (3.5) を代入すると

$$(3.22) \quad V \cdot W = \sum_{i=1}^m \sum_{j=1}^{m'} \alpha_i \alpha'_j k(\vec{x}_i, \vec{x}'_j)$$

となり, 有限の計算時間で類似性が評価できることがわかる.

上に示した計算手順により核非線形主成分分析を相互部分空間法に適用することが可能であることが示された. 具体的に核非線形主成分分析で求められた部分空間同士の角度を計算するためには, 式 (3.1)(3.2) に (3.22) を代入すればよい. こうして導出された核非線形主成分分析に基づく相互部分空間法は核非線形相互部分空間法と呼ばれる(坂野・武川 (1999), 坂野 他 (2001)).

なお, 実データに適用する場合には具体的な核関数として SVM の場合と同様に動径基底関数 (Radial Basis Function; RBF)

$$(3.23) \quad k(\vec{x}, \vec{y}) = \exp\left(\frac{-\|\vec{x} - \vec{y}\|^2}{\sigma^2}\right),$$

多項式

$$(3.24) \quad k(\vec{x}, \vec{y}) = (1 + \vec{x} \cdot \vec{y})^m,$$

もしくは

$$(3.25) \quad k(\vec{x}, \vec{y}) = (\vec{x} \cdot \vec{y})^m$$

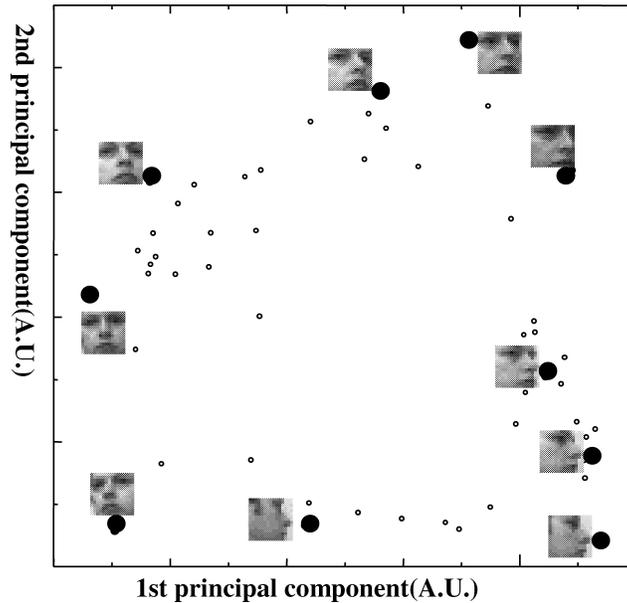


図 3. 線形主成分分析による顔画像の散布図.

などが用いられる．ここで  $\sigma, m$  などのパラメータは写像の非線形性を制御するパラメータでカーネルパラメータと呼ばれる．

以下に UMIST データに対して核非線形相互部分空間法を適用した結果を示す．なお詳細な実験条件については坂野・武川 (1999), 坂野 他 (2001) を参照されたい．

核非線形相互部分空間法には従来の非線形手法と比較して興味深い性質がある．図 3 は UMIST データベースの一人目の画像列を線形主成分分析した 1 番目と 2 番目の主成分で張られた 2 次元の部分空間にデータを写像したものである．主成分分析は図中に黒丸で示した点で表されるデータを用いて行い, それ以外のデータを点でプロットした．なお, 黒丸で表したデータには対応する顔画像を示した．

この散布図から分かる, もしくは推測されることは

- ・姿勢変動に起因する非線形性の存在
- ・データが円弧の様な一次元の関数で近似出来そうに見える

の 2 点である．こうした非線形性は, 例えば, 正面顔と横顔を重ね合わせた画像が顔画像ではなくなること, つまり, 顔画像に属する要素の中に, 線形結合が成立しないものがあるために発生すると理解されているが, その性質や非線形性を定量的に説明する理論は成立していない．

図 4 に相互部分空間法, 核非線形相互部分空間法の認識率の部分空間次元数の依存性を, 図 5 に核非線形主成分分析で得られた上位 2 位の主成分で張られる部分空間への散布図を示す．図 4 では相互部分空間法の認識率は次元数 3-4 程度になるまで低いままなのに対して, 核非線形相互部分空間法では次元数が 1 でもほぼ 100% に近い認識率が得られている．また, 図 5 の散布図を見ると, 一部 outlier が見られるものの, ほぼ 1 次元の部分空間にフィッティングされていることが分かる．つまり, データに含まれる顔の運動の自由度 1 を核非線形主成分分析がうまく近似し, 相互部分空間法の雑音除去能力によりそのことが認識率に反映されたと考え

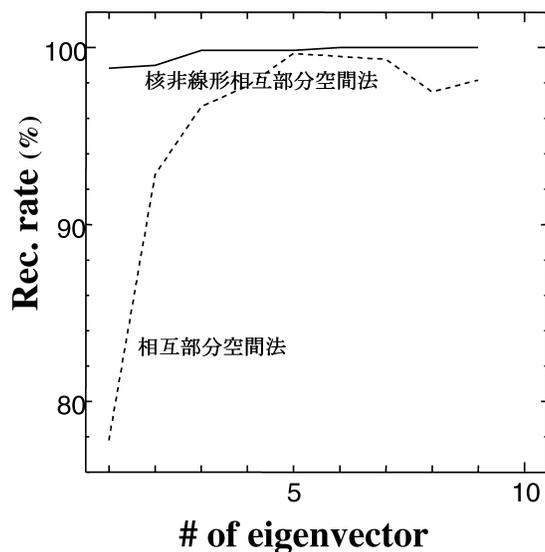


図 4. 認識率の次元数依存性.

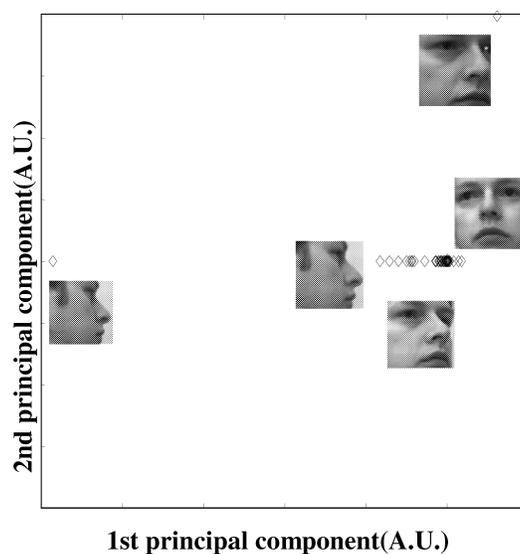


図 5. 核非線形主成分分析による顔画像の散布図.

られる。つまり、これらのデータは顔の物理的な運動の自由度で画像空間に現れる非線形構造を非線形部分空間が高精度に近似すれば姿勢変動に伴う画像変動は克服されることを示唆している。

さらに、相互部分空間法では部分空間の次元数が 5 次元の時にピーク性能を出し、以降次元数が上がるごとに性能が低下する。これは、特徴空間における部分空間の重複部分が増加しているためと考えられる。これに対して、核非線形相互部分空間法では部分空間の次元数が高く

なっても認識率の低下が起こっていない。これは高次元空間への写像により、部分空間の重なり部分が生じにくくなっていることを示唆している。つまり、登録人数が増加した場合にも認識率が低下しない、もしくは減少が少ないという可能性を示している。

### 3.3 幾何学的、光学的拘束に基づく方法

前節までに解説した方法論は基本的には学習時に色々な条件で撮影された顔画像を学習することにより、多様な顔画像の変形をいかにして吸収するかというアプローチであった。姿勢変動や表情による顔画像の変動を考えた場合、このアプローチは正統的であり、実用的な価値も高い。しかしながら、顔画像の変動には、姿勢、表情などの他に照明条件による変動が非常に大きい(懐中電灯を下から照らした顔などを想起されたい)。しかも、姿勢、表情などと違い、登録時にユーザに「照明条件を変えてください」というわけには行かない。無論、実装では顔画像登録の際に様々な照明条件での顔画像を撮像することが考えられるが登録設備のコストの増大につながるなどの問題を生ずる。

この様な照明変動に対する対策としては、

1. エッジなど照明変動に対しての変動が小さな特徴を用いること。
2. レンジファインダなどの手段で3次元構造を取得し、認識時に変動をシミュレーションすること。
3. 様々な照明変動を学習データに含め、照明変動に不変な認識辞書を設計すること。

などの方法が考えられるが、1.の方法は照明変動に不変な特徴がそもそも個人の特徴を表しているかどうか疑わしく実験的にも高い認識率は得られていない(例えば, Adini et al.(1997))。第2の3次元モデルを用いる方法では、入力された顔画像について、姿勢、照明方向などを推定する問題を抱え込むことになるがこれを効率的に解決する方法は発見されていない。第3の様々な照明条件で撮影した顔画像を学習する方法は先に述べた通り実装上の理由から現実的ではない。

現在、画像パターン認識の分野で研究されているのは主として、視覚の計算理論に基づくアプローチである。視覚の計算理論とは、主に D. Marr によって創始された人間の視覚の機能を可能にする計算過程を明らかにすることにより脳の機能を明らかにしようとする学問体系である。工学的には脳が行うのと同様な視覚機能を計算機で実現するための学問と位置付けることが出来る。この分野では陰影、テクスチャ(模様)、運動などの情報から対象の3次元構造を復元するための各種の方法が提案されている(徐・辻井(1999))。これらの手法は原理的には全て顔の認識のために応用することが可能であり、実際に幾つかの応用例も存在するが、主成分分析との関連で特に興味深いのは、Georghiades らによる Illumination cone と呼ばれる光学的な拘束条件を用いた方法である (Georghiades et al. (1998))。

説明の都合上、初めに、物体とカメラが固定され、光源方向のみが変動するときの3次元物体認識問題を考える。ここで物体の表面反射が、Lambert モデル、つまり

$$(3.26) \quad I = \rho(\vec{n} \cdot \vec{l})$$

の様に画像の輝度  $I$  が光源の単位入射ベクトル  $\vec{l}$  と物体表面の単位法線ベクトル  $\vec{n}$  のみで決定されているとする。

この様な場合に Shashua は幾つかの実験を通して光源変動の問題が画像の線形結合で解決できることを示唆した (Shashua (1992))。この示唆に触発された Belhumeur and Kriegman (1996) は物体が凸であり、その表面が Lambert 反射モデルに従い、光源が無遠にある点光源とみなせるとき、任意の光源方向から照射された物体の画像の輝度  $I^{(4)}$  が、独立な3方向の光源のも

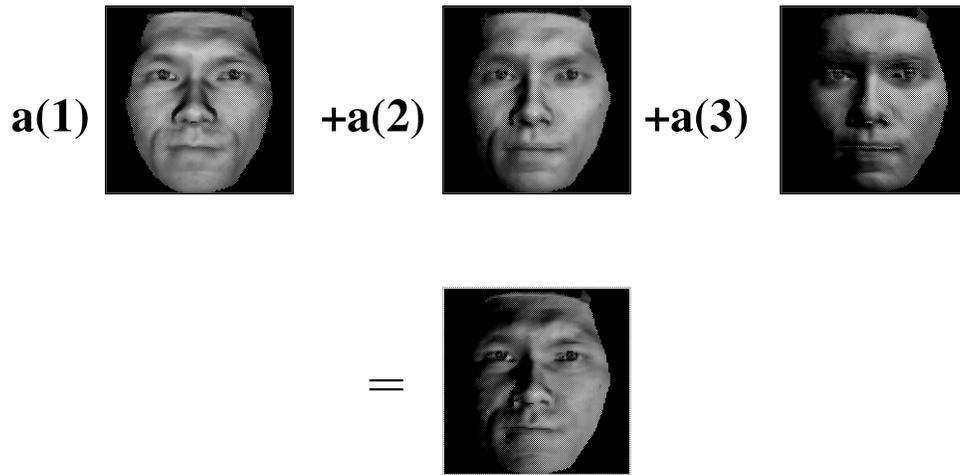


図 6. 線形結合により異なる照明条件の画像が生成される例 .

とで得られた 3 枚の画像  $I(1), I(2), I(3)$  の線形結合 ,

$$(3.27) \quad I(4) = \max \left( \sum_{j=1}^3 a(j)I(j), 0 \right)$$

により表されることを示した(図 6 参照). ここで,  $a(j)$  は結合係数である. このことは, 光源変動による画像の変化が画像空間中の 3 次元の部分空間で記述されることを示している. つまり, 実例の画像からこの部分空間の基底ベクトル,  $\hat{I}(1), \hat{I}(2), \hat{I}(3)$  を生成出来れば, 光源変動に起因する全ての画像が記述できたことになる. Belhumeur and Kriegman (1998) はこの部分空間から, 成分が負となる部分を除いた半部分空間を Illumination cone と名づけた.

従って, これを顔認識の問題に適用すれば, 表情変動, 姿勢変動がない場合の光源不変認識アルゴリズムが構成出来ることになる. Georghiadis らは, このアイデアに基づく認識アルゴリズムを提案し, Illumination cone と入力画像の二乗誤差を計算する認識系を構成し, 光源変動に対するロバスト性を実験的に示した.

無論, この時に用いられる 3 枚の画像は十分にかげ離れた条件で撮影されていなくてはならない. 例えば, 右側からだけ上, 中, 下, の様に撮影した画像からでは顔全体を良好に再構成する事は出来ない. また, 最適な画像再構成のためには問題の 3 枚の画像  $\hat{I}(1), \hat{I}(2), \hat{I}(3)$  は直交していることが要求される. 従って, Illumination cone のアイデアを顔認識系に应用する場合には, この 3 次元部分空間の基底ベクトルは, 複数の別方向からの照明を当てた顔画像データから得られた自己相関行列を特異値分解することで計算される.

つまり, 登録時には個人  $k$  について  $m$  枚の登録画像,  $I^k(1), \dots, I^k(m)$  を用意し, これから計算された自己相関行列の第 3 固有値までに対応する固有ベクトル  $\hat{I}^k(1), \hat{I}^k(2), \hat{I}^k(3)$  をテンプレートとして登録する.

認識時には, 入力された顔画像  $I$  に対して個々人に対応する Illumination cone との自乗誤差

$$(3.28) \quad (s^k)^2 = I^2 - \sum_j^3 (\hat{I}^k(j) \cdot I)^2$$

を計算し自乗誤差の最も小さな  $k$  を認識結果として出力する. Georghiadis らの実験では, 照

明変動した正面顔画像についてはほぼ 100% の認識率が得られている。

Illumination cone で物体の 3 次元構造が記述されることは因子分解法で複数枚の画像を特異値分解することで 2 次元画像から 3 次元が再構成されるプロセスと類似している (Tomasi and Kanade (1992), 解説として 徐・辻井(1999), 松山 他(1998)を参照されたい)。

因子分解法は動画から 3 次元を再構成する手法のひとつである。撮影対象となる物体と光源が固定され、カメラが動いた場合に撮像された、 $F$  フレームの動画画面上で、物体上の  $P$  個の特徴点が追跡できたとする。このときに画像上の  $P$  個の特徴点の座標で構成される  $F \times 2P$  の行列  $W$  は特異値分解によって  $W = R \times S$  というカメラの回転パラメータ  $R$  と物体の 3 次元形状  $R$  に分解される。手続き的にはこのときの特異値分解は画像上の  $2P$  個の特徴点座標で構成されたベクトルを主成分分析することに相当する。

このときカメラの自由度 3 が特異値分解により残るわけであるが、Illumination cone の場合には撮影対象となる物体とカメラが固定され、光源の自由度 3 が特異値分解により現れるわけである。この様な観点から Illumination cone の方法は光学的因子分解法と呼ばれることもある。

Illumination cone の方法は、原理的に照明条件による変動を全て吸収することが出来るが、一方で、複数の画像の対応点を厳密に計算する必要があるために顔の姿勢が変動した場合には適用できないという問題があった。この問題に対し、春山 他 (1998) は Maki et al. (1998) の提案した Geotensity 拘束を用いることにより光源、姿勢変動が存在した場合にもロバストな認識手法を提案している。

上記の手法は、コンピュータビジョンの理論の基礎的な側面の検証を行う意味で意義深いですが、一方で、問題を顔の認識に限れば鼻の影で生ずる非線形性を処理できない、姿勢の変動に対処するためにはどうしても画像間の対応付けの問題を解かなくてはならない、など問題も多く、今後も多くの研究が期待される。

#### 4. おわりに

以上、顔画像の認識問題を例にとり、パターン認識における主成分分析の使われ方と独自に行われて来た改良について述べてきた。主成分分析は将来においてもパターン認識の分野では重要な技術でありつづけ、様々な改良が行われるであろう。

しかしながら、ここで紹介した技術は顔画像認識の課題のうち姿勢変動、照明変動などの個別の問題に対処したものに過ぎず、顔画像認識の問題をトータルに解決するものではない。また顔の経時変化、表情変動の問題に関してはアプローチすら見えていない状態である。

これらの問題の解決のためにはここで紹介したパターン認識、コンピュータビジョンの技術の他にも信号処理やコンピュータグラフィックス、心理学、生理学など様々な分野の協力が重要だと考えられる。このような学際的な領域に成長しつつある顔画像認識技術の研究において主成分分析がどのような役割を果たしうるか、今後の興味も尽きることはない。

#### 謝 辞

有益なコメントをして頂いた査読者の皆様、有益な討論をしていただき、貴重な助言をしていただいた統計数理研究所 伊庭幸人助教授、UMIST 画像データベースの使用を許可して頂いたマンチェスター大学情報理工学研究所の Allinson 教授、実験データを提供していただいた(株)NTT データ 春山智氏、顔画像を提供していただいた(株)NTT データ 中山直樹氏、データの整備などにご協力いただいた沼野ゆき氏、そして、文献の収集などに協力して頂いた電子

技術総合研究所 坂野貴子博士に感謝します。

### 参 考 文 献

- Adini, Y., Moses, Y. and Ullman, S. (1997). Face recognition: The problem of compensating for changes in illumination direction, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **19**, 721–732.
- アイゼルマン, M. A., ブラヴェルマン, E. M., ロゾノエル, L. I. (1978). 『パターン認識と学習制御』 (北川敏夫, 林 順雄 訳), 共立出版, 東京.
- 赤松 茂 (1997). コンピュータによる顔の認識—サーベイ—, 信学論, **J80-D-II(8)**, 2031–2046.
- Bar, S. D., Edelman, S., Howell, A. J. and Buxton, H. (1998). A similarity based method for generalization of face recognition over pose and expression, *Proceedings of IEEE Automatic Recognition of Face and Gesture '98*, 118–123.
- Belhumeur, P. B. and Kriegman, D. (1996). What is the set of images of an object under all possible illumination conditions?, *Proceedings of International Conference of Computer Vision*, 1053–1059.
- Belhumeur, P. B. and Kriegman, D. (1998). What is the set of images of an object under all possible illumination conditions?, *International Journal of Computer Vision*, **28(3)**, 245–260.
- Belhumeur, P. B., Hespanha, J. P. and Kriegman, D. J. (1997). Eigenface vs. Fisharfaces: Recognition using class specific linear projection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **19**, 711–720.
- Chellapa, R., Wilson, C. J. and Sirohey, S. (1995). Human and machine recognition of face: A survey, *Proceedings of IEEE*, **83(5)**, 704–740.
- Diamantaras, K. I. and Kung, S. Y. (1996). *Principal Component Neural Networks*, Wiley, New York.
- Duda, R. O., Hart, P. E. and Stork, D. G. (2001). *Pattern Classification*, 2nd ed., Wiley, New York.
- Frey, B. J., Colmenarez, A. and Huang, T. S. (1998). Mixtures of local linear subspaces for face recognition, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 32–37.
- Georghiades, A. S., Kriegman, D. J. and Belhumeur, P. N. (1998). Illumination cones for recognition under variable lighting: Faces, *Proceedings of Computer Vision and Pattern Recognition*, 52–58.
- Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*, Wiley, New York. (丘本 正, 磯貝恭史 訳(1979) 『統計的多変量解析』, 日科技連出版社, 東京)
- Graham, D. B. and Allinson, N. S. (1998). Characterizing virtual eigensignatures for general purpose face recognition, *Face Recognition: From Theory to Applications* (eds. H. Wechsler et al.), p. 446, Springer, Berlin.
- 春山 智, 坂野 鋭, 武川直樹 (1998). 陰影・幾何拘束を用いた光源・姿勢変動にロバストな顔認識アルゴリズム, 信学技報, PRMU98–134.
- 飯島泰蔵 (1989). 『パターン認識理論』, 森北出版, 東京.
- 石井健一郎, 上田修功, 前田英作, 村瀬 洋 (1998). 『よくわかるパターン認識』, オーム社, 東京. (部分空間法を始めとするパターン認識については多くの良書があるが, 本書は特に優れている.)
- Jain, A. K. and Waller, W. G. (1978). On the optimal number of features in the classification multivariate Gaussian data, *Pattern Recognition*, **10**, 365–374.
- Jain, A. K., Duin, R. P. W. and Mao, J. (2000). Statistical pattern recognition: A review, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **22**, 4–37.
- 徐 剛, 辻井三郎 (1999). 『三次元ビジョン』, 産業図書, 東京.
- 近藤次郎 (1953). 『積分方程式』, 培風館, 東京.

- 黒沢由明 (1996). 部分空間法とベイズ識別の関係から導かれる新しいパターン認識方式, 信学技報, PRU96-104.
- 前田英作, 村瀬 洋 (1999). カーネル非線型部分空間法によるパターン認識, 信学論 (D-II), J82-D-II(4), 600-612.
- 前田賢一, 渡辺貞一 (1985). 局所構造を導入したパターン・マッチング法, 信学論 (D), J68-D(3), 345-352.
- Maki, A., Watanabe, M. and Wiles, C. (1998). Geotensity: Combining motion and lighting for 3D surface reconstruction, *Proceedings of International Conference of Computer Vision '98*, 1053-1060.
- 松山隆司, 久野義徳, 井宮 淳 編 (1998). 『コンピュータビジョン: 技術評論と将来展望』, 新技術コミュニケーションズ, 東京.
- Moghaddam, B. (1999). Principal manifolds and Bayesian subspaces for visual recognition, *Proceedings of International Conference of Computer Vision*, 1131-1136.
- 村瀬 洋 (1995). CVCV-WG 特別報告: コンピュータビジョンにおける技術評論と将来展望 (VI) — 固有空間法による画像認識 —, 情報処理学会研究報告, CV97-9, 59-66.
- 村瀬 洋, Nayar, S. K. (1994). 2次元画像照合による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), 2179-2187.
- 中山直樹, 春山 智, 坂野 鋭 (1998). 撮動法による顔画像の認識, 信学技報, PRMU98-90.
- Oja, E. (1982). A simplified neuron model as a principal component analyser, *J. Math. Biol.*, **15**, 267-273.
- オヤ, エルツキ (1986). 『パターン認識と部分空間法』, 産業図書, 東京.
- Pentland, A. and Choudbury, T. (2000). Face recognition for smart environments, *IEEE Computer*, **33**, p. 50.
- Raudys, S. J. (1981). Determination of optimal dimensionality in statistical pattern recognition, *Pattern Recognition*, **11**, 263-270.
- Saito, T. and Kariya, T. (1988). A generalization of principal component analysis, *J. Japan Statist. Soc.*, **18**(2), 187-193.
- 坂野 鋭 (1999). バイオメトリック個人認証技術の現状と課題, 信学技報, PRMU99-29.
- 坂野 鋭, 武川直樹 (1999). 核非線型相互部分空間法と顔認識への応用, 信学技報, PRMU99-116.
- 坂野 鋭, 武川直樹, 中村太一 (2001). 核非線形相互部分空間法による物体認識, 信学論, J84-D-II(8) (掲載予定).
- Schölkopf, B., Smola, A. and Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, **10**, 1299-1319.
- Shashua, Ammon (1992). Illumination and view position in 3D visual recognition, *Advances in Neural Information Processing*, **4**, 404-411.
- Tomasi, C. and Kanade, T. (1992). Shape and motion from image stream under orthography: A factorization method, *International Journal of Computer Vision*, **9**(2), 137-154.
- 津田宏治 (1999). ヒルベルト空間における部分空間法, 信学論 (D-II), J82-D-II(4), 592-599.
- Turk, M. and Pentland, A. (1991). Face recognition using eigenfaces, *Proceedings of Computer Vision and Pattern Recognition*, 568-591.
- Watanabe, S. and Pakvasa, N. (1973). Subspace method of pattern recognition, *Proceedings of 1st International Joint Conference of Pattern Recognition*, 25-32.
- Williams, C. K. I. (2001). On a connection between kernel PCA and metric multidimensional scaling, *Advances in Neural Information Processing Systems 13* (eds. T. K. Leen, T. G. Diettrich and V. Tresp), MIT Press, Cambridge, Massachusetts (to appear).
- 山口 修, 福井和宏, 前田賢一 (1997). 動画像を用いた顔認識システム, 信学技報, PRMU97-50.

## Principal Component Analysis in Pattern Recognition —From the Viewpoint of Facial Image Recognition—

Hitoshi Sakano

(NTT Data Corporation)

In this article, we introduce “How to use” the principal component analysis in facial image recognition. We also introduce some improvement in the fields.

First, we describe the role of principal component analysis in image recognition technology. And we point out some difficulties in facial image recognition technology, for example image change caused by illumination change, nonlinear distribution caused by head pose change.

Finally, we introduce some improvement of principal component analysis and how to solve the problems.

---

Key words: Pattern recognition, computer vision, principal component analysis, facial image recognition.