

最近の技術動向： 相互部分空間法への拡張と その応用事例

福井和広

筑波大学大学院システム情報工学研究科

部分空間法から相互部分空間法への拡張

部分空間法は、各カテゴリのパターンの拡がりを部分空間として捉え、入力ベクトル \mathbf{x} がどのカテゴリの部分空間に最も近いかに基づいて識別を行う。近さを測る指標 S は、部分空間を張る辞書ベクトルを φ_i 、辞書ベクトルの本数を r とすると次式で表される。

$$S = \sum_{i=0}^{r-1} \frac{(\mathbf{x}, \varphi_i)^2}{\|\mathbf{x}\| \|\varphi_i\|} \quad (1)$$

上式は入力ベクトル \mathbf{x} の部分空間への射影長の自乗を表しているが、同時に入力ベクトルと部分空間の成す角度 θ_1 の余弦の自乗 $\cos^2 \theta_1$ でもある。このように部分空間法の類似度の意味には2つの解釈が可能であるが、本稿で部分空間法を角度に基づく方法と見なして、部分空間法に対する一連の拡張を説明することにする。

部分空間法はパターンの拡がりを1つの代表ベクトルに替えて部分空間で表すことで、単純類似度法 ($r=1$ の場合) に比べてパターン変動に対する吸収能力を大きく向上させた。この自然な拡張として、入力側も部分空間に置き換えて、入力部分空間 P と辞書部分空間 Q の成す最小角 θ_1 に基づく識別法が提案された (図-1)。この方法は相互部分空間法 (Mutual Subspace Method (MSM))²⁾ と呼ばれ、次式で定義される類似度 S に基づいて識別を行う。

$$S = \max_{\substack{\mathbf{u} \in P, \mathbf{v} \in Q \\ \|\mathbf{u}\| \neq 0, \|\mathbf{v}\| \neq 0}} \frac{(\mathbf{u} \cdot \mathbf{v})^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2} \quad (2)$$

上式は2つの部分空間の成す最小角 θ_1 の余弦の自乗 $\cos^2 \theta_1$ を表している。この最小角 θ_1 を求める問題は、2つの部分空間 P と Q に含まれる任意ベクトル $\mathbf{u} \in P$ と $\mathbf{v} \in Q$ のペアの中で、最大相関係数を持つペアを探索する問題となる。実はこの問題に対する定式化は多変量解析における正準相関分析¹⁾と同じものである。ただ

し相互部分空間法は正準相関分析の特別な場合に相当すると見なされるので、概念的には正準相関分析の方が広いと言える。

相互部分空間法は2つのパターン集合の分布の類似性を効率良く測る方法となっている。たとえば動画像や多視点画像を用いた顔認識への適用では、カメラに対して相対運動する顔の動画像列、あるいは複数カメラから得られる多視点顔画像データ分布形状の類似度を定量的に測ることで、従来法に比べてきわめて高い識別性能を実現している³⁾。1枚の入力画像パターンのみを用いて部分空間法で顔認識を行うとすると、正面から見るときわめて似ているが少し角度を変えて見ると異なるような顔をうまく識別できない。この不確定性に加えて、頭部姿勢、表情変化、照明条件などの変動による顔パターンの変化は、1枚の画像パターンのみを用いた識別をさらに難しくしている (図-2(a) 参照)。これに対して、動画像列や多視点画像を用いた相互部分空間法は、顔の3次元形状を反映しているパターン分布の形状を比較することで陰に3次元的な識別を行い、高精度な識別を実現する (図-2(b) 参照)。

正準角に基づく相互部分空間法の一般化

前章では2つの部分空間の成す最小角のみを考えたが、両者の間には正準角 (canonical angle)^{☆1} と呼ばれる角度が定義できる。先に述べた最小角はこれらのうちで最小のものに対応している。これらの正準角は2つの部分空間の幾何学的な関係は規定しているので、最小正準角に加えて他の正準角も類似度の計算に考慮することで、相互部分空間法の識別性能をさらに向上できる。ここで複数の正準角に基づいた方法を一般化された相互部分空間

☆1 principal angle とも呼ばれる。

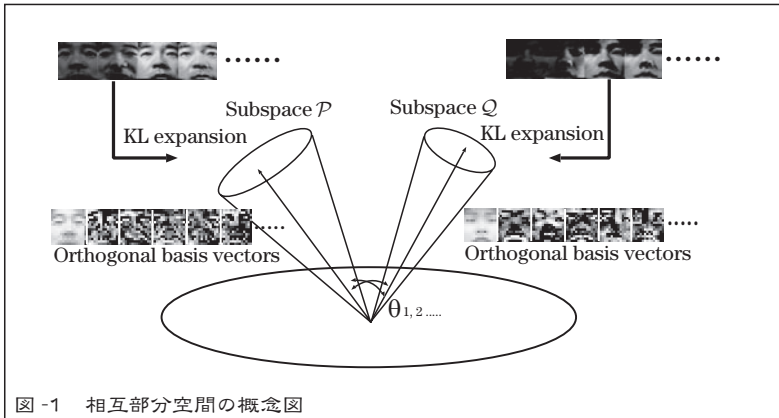


図-1 相互部分空間の概念図

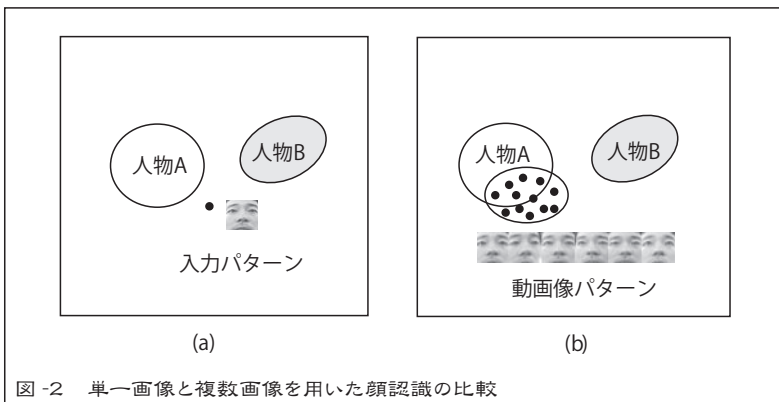


図-2 単一画像と複数画像を用いた顔認識の比較

法と呼ぶことにする。

●正準角とは

図-3は2つの1次元部分空間(ベクトル)の成す角度を示しており、これは最も簡単な正準角の例である。また2つの2次元部分空間(平面)に対しては、図-4に示すように2つの正準角 θ_1 と θ_2 が定義できる。この図では3次元空間に2つの2次元部分空間を押し込んでいるために、両者の間には1次元の共通部分空間が存在し、最小正準角 θ_1 は零度となっている。一般的には、 f 次元ベクトル空間における m 次元部分空間 P と n 次元部分空間 Q (便宜上、 $m \geq n$ 、かつ両者に共通部分空間が存在しない場合)の間には n 個の正準角が定義でき、第1正準角 θ_1 は2つの部分空間の成す最小角となる。第2正準角 θ_2 は最小正準角 θ_1 に直交する方向において測った最小角、同様に第3正準角 θ_3 は第1正準角と第2正準角に直交する方向で測った最小角である。以下同様に n 個の正準角 $\theta_i (i = 1, \dots, n)$ が順次求まる。これを式で書くと次のようになる。

$$\cos^2 \theta_i = \max_{\substack{\mathbf{u}_i \perp \mathbf{u}_j (j=1, \dots, i-1) \\ \mathbf{v}_i \perp \mathbf{v}_j (j=1, \dots, i-1)}} \frac{|\langle \mathbf{u}_i, \mathbf{v}_i \rangle|^2}{\|\mathbf{u}_i\|^2 \|\mathbf{v}_i\|^2} \quad (3)$$

ここで、 $\mathbf{u}_i \in P$ 、 $\mathbf{v}_i \in Q$ 、 $\|\mathbf{u}_i\| \neq 0$ 、 $\|\mathbf{v}_i\| \neq 0$ 、また (\cdot) と $\|\cdot\|$ はそれぞれ内積とノルムを表す。

●正準角の計算

ここでは f 次元ベクトル空間における m 次元部分空間 P と n 次元部分空間 Q (便宜上、 $m \geq n$ 、かつ両者に共通部分空間が存在しない場合)の間で定義される n 個の正準角の計算法について説明する。以下では計算法1から計算法3の順に解くべき行列が簡単化されているが、すべて n 個の同一の正準角セットが求まる。

計算法1:

正準相関分析の定式化によると、正準角を求める問題は部分空間 P と Q への射影を表す射影行列 P と Q の積 PQ (あるいは QP)の固有値問題に帰着する¹⁾。ここで A と B を以下に示すようにそれぞれ部分空間 P と Q の正規直交基底ベクトルを列に並べた行列とすると、 $P = AA^T$ と $Q = BB^T$ となる。

$$A = [\Psi_1 \Psi_2 \dots \Psi_m] \quad (4)$$

$$B = [\Phi_1 \Phi_2 \dots \Phi_n] \quad (5)$$

$f \times f$ 行列 PQ (あるいは QP)の第 i 番目に大きい固有値 λ_i が第 i 正準角 θ_i の余弦の自乗 $\cos^2 \theta_i$ となる。この計算は次のように幾何学的に解釈できる。行列 PQu はあるベクトル $u \in P$ を部分空間 Q に射影し、続けて部分空間 P へ射影する操作を表している。したがって射影ベクトル PQu とベクトル u の方向が同じ、つまり $PQu = \lambda u$ となる場合に、ベクトル u とベクトル Qu の成す角度が部分空間の成す最小角に対応する。

計算法2:

$f \times f$ 行列 PQ (あるいは QP)の固有値問題は、小さいサイズ $m \times m$ 行列 $C (= A^T B B^T A)$ (あるいは $n \times n$ 行列 $C' (= B^T A A^T B)$)の固有値問題に変換できる。この変換により計算法1に比べて少ない計算量で $\cos^2 \theta_i$ が行列 C (あるいは C')の第 i 固有値 λ_i として求まる^{☆2}。

$$C = (c_{ij}), (i, j = 1 \sim m), c_{ij} = \sum_{k=1}^n (\Psi_i \cdot \Phi_k)(\Phi_k \cdot \Psi_j) \quad (6)$$

$$C' = (c_{ij}), (i, j = 1 \sim n), c_{ij} = \sum_{k=1}^m (\Phi_i \cdot \Psi_k)(\Psi_k \cdot \Phi_j) \quad (7)$$

計算法3:

上記固有値の平方根 $\sqrt{\lambda_i}$ 、つまり第 i 番目に小さい正準角 θ_i の余弦 $\cos \theta_i$ は、行列 $D = A^T B$ あるいは $D' = B^T A$ の第 i 番目に大きい特異値 μ_i として求まる。

☆2 行列 C と C' は n 個の同じ非零の固有値を持つ。

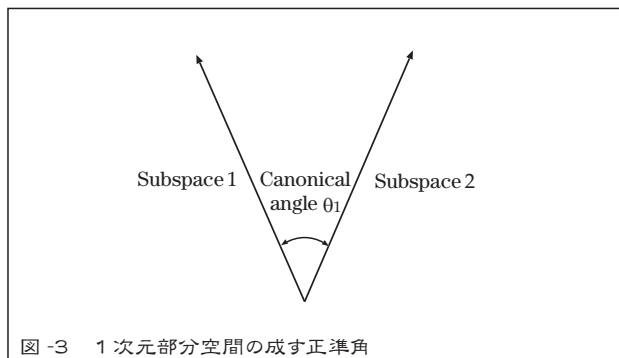


図-3 1次元部分空間の成す正準角

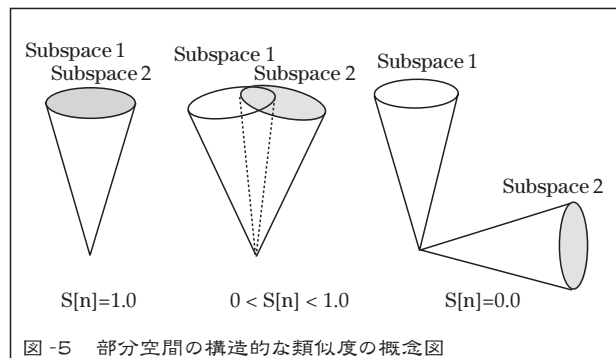


図-5 部分空間の構造的な類似度の概念図

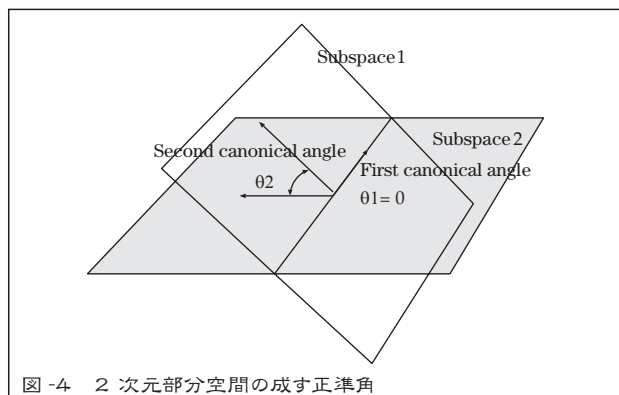


図-4 2次元部分空間の成す正準角

●部分空間の構造的な類似度

2つの部分空間の成す正準角により部分空間の構造的な類似度を定量的に測ることができる。たとえば $\cos^2\theta$ の n 番目までの平均値 $S[n] = \frac{1}{n} \sum_{i=1}^n \cos^2\theta_i$ は次のような特性を持つ。図-5に示すように、2つの部分空間が完全に一致するときに最大値 1.0 となり、両者が離れるにつれて次第に低下してゆき、両者が完全に直交するときに最小値 0.0 となる。このほかにも平均に替えて累積を用いても同じような特性を持つ類似度が得られる。

ただし類似度 $S[n]$ が有効となるのは、比較する2つの分布が多様なデータを含んでいる場合のみである。たとえば画像認識においては対象物が運動している場合や、多視点カメラにより異なる画像を安定に獲得できる場合などは、この条件を満足する。

カテゴリ部分空間の直交化による 識別性能の向上

一般化された相互部分空間法（以下では、単に相互部分空間法）は、部分空間法に比べてパターン変動に対する高い吸収能力を実現する。しかしながら、その識別能力は不十分であり、性能改善の余地が多分に残されている。これは部分空間法ベースの方法全般に言えることだが、あるカテゴリ部分空間はそのカテゴリ分布を近似するという点では最良となっているが、他のカテゴリ部分空間との関係を考慮されていないために、識別という観点では

必ずしも最良の部分空間にはなっていないためである。

本来、高い認識性能を実現するためには、全カテゴリに共通的な特徴はできるだけ抑制し、各カテゴリの差異が顕著に現れる特徴を選択的に用いることが有効である。これは文字認識において、“犬”と“大”を識別するためには、点の有無を判定基準とすることが有効であることから理解できる。

ユークリッド距離に基づく識別法では、特徴抽出の有効性を測る指標としてフィッシャー基準がよく知られている。これに対して、部分空間法などの角度(相関)ベースの識別法では、各カテゴリ部分空間がお互いにどの程度直交関係にあるかが、有効な基準となる。2つのカテゴリ部分空間が直交関係にあるということは、一方のカテゴリ部分空間に含まれる特徴は、他のカテゴリ空間には含まれないということを意味する。したがって直交化されたカテゴリ部分空間を用いて識別を行うということは、各カテゴリのみが持つ特徴を用いた識別を行うことになり、高い識別性能が実現できる。

直交化を実現する方法として、各カテゴリ分布の自己相関行列の総和を用いた白色化が挙げられる。この方法は有効な特徴抽出法として提案され、これを前処理として組み込んだ部分空間は直交部分空間法と呼ばれ、その有効性が示されている。この拡張として自己相関行列の代わりに各カテゴリ部分空間への射影を表す射影行列の総和を白色化することで直交化したカテゴリ部分空間に相互部分空間法を適用した直交相互部分空間法(OMSM)^{☆3}が提案された。OMSMは相互部分空間法に比べて顔画像認識の識別性能を大きく向上させることが示されている⁴⁾。

このほかにも、各カテゴリ部分空間の差異を表す一般化差分部分空間へ各カテゴリ部分空間を射影することで、カテゴリ部分空間の関係を直交化に近付ける方法も提案されている。この射影を組み込んだ相互部分空間法は制約相互部分空間法(Constrained Mutual Subspace Method(CMSM))と呼ばれ、OMSMと同様に、顔認識、物体認識などの問題において有効であることが確認されている⁶⁾。

☆3 Whitened Mutual Subspace Method (WMSM)とも呼ばれている。

カーネル関数を用いた非線形識別への
拡張

一般に動画像や多視点カメラから得られる画像パターンの分布は非線形構造を持つために、各カテゴリのパターン分布を他のカテゴリ部分空間と重ならないように、線形部分空間で表すことは難しい。この場合にはカーネル主成分分析 (Kernel PCA) を適用してパターン分布を非線形部分空間で表し、これを用いて先に述べた一連の方法を適用するのが有効である。

●カーネル主成分分析 (Kernel PCA)

KPCA は、入力ベクトル \mathbf{x} を非線形写像関数 $\phi(\mathbf{x})$ によりきわめて高次元の特徴空間 \mathcal{F} へ写像し、写像されたベクトル $\{\phi(\mathbf{x})\}$ に対して通常の線形主成分分析 (PCA) を適用する方法である。特徴空間 \mathcal{F} 上に生成された線形部分空間は、入力空間から見ると非線形写像先の空間に生成されているので、“非線形部分空間”と呼ばれる。しかし実体はあくまでも高次元空間特徴空間上の線形部分空間である。

カーネル主成分分析について簡単に述べる^{☆4}。まず d 次元ベクトル空間 (入力空間と呼ばれる) のベクトル \mathbf{x} を非線形写像関数 $\phi(\mathbf{x})$ により、入力空間に比べて遥かに高い次元 d_ϕ の特徴空間 \mathcal{F} に非線形写像する。

$$\phi : \mathbf{x} \rightarrow \phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \dots, \phi_{d_\phi}(\mathbf{x}))^T \quad (8)$$

特徴空間 \mathcal{F} 上の写像に対して、主成分分析を行うためには、異なるベクトル \mathbf{x} と \mathbf{y} のそれぞれの写像 $\phi(\mathbf{x})$ と $\phi(\mathbf{y})$ の内積を計算する必要がある。しかし、特徴空間 \mathcal{F} 上において、この内積 $(\phi(\mathbf{x}) \cdot \phi(\mathbf{y}))$ を直接計算することは、ベクトルの次元がきわめて高いために計算困難 (無限次元空間では不可能) となる。ところが非線形写像 ϕ をカーネル関数 $k(\mathbf{x}, \mathbf{y})$ を介して定義すると、内積 $(\phi(\mathbf{x}) \cdot \phi(\mathbf{y}))$ は入力ベクトル \mathbf{x} と \mathbf{y} のみから計算できる。つまり $(\phi(\mathbf{x}) \cdot \phi(\mathbf{y})) = k(\mathbf{x}, \mathbf{y})$ の関係が成立する。これが“カーネルトリック”と呼ばれる計算技法である。具体的な非線形写像 ϕ が存在するためには、関数 $k(\mathbf{x}, \mathbf{y})$ が対称であり、かつ任意の有限個の点集合に対してカーネル行列 (後で定義) が半正定値になるという条件 (Mercer の条件) を満たす必要がある^{☆5}。これらの条件を満たす関数はカーネル関数と呼ばれ、多項式関

数やガウシアン関数が代表的である。

$$k(\mathbf{x}, \mathbf{y}) = (1 + (\mathbf{x}, \mathbf{y}))^n \quad (9)$$

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2}\right) \quad (10)$$

このガウシアン関数を適用した場合には、入力ベクトルは無限次元空間へ写像されることになる。

m 個のベクトル $\mathbf{x}_i (i = 1, \dots, m)$ に対するカーネル主成分分析は、何段階かの導出過程を経て、最終的にはカーネル関数を介して得られる以下の $m \times m$ のカーネル行列 \mathbf{K} の固有値問題に帰着する。

$$\begin{aligned} \mathbf{K}\mathbf{a} &= \alpha\mathbf{a} \\ K_{ij} &= (\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)) \\ &= k(\mathbf{x}_i, \mathbf{x}_j) \end{aligned} \quad (11)$$

求めるべき非線形部分空間 \mathcal{V} の第 i 正規直交基底ベクトル \mathbf{e}_i は、 m 個のベクトル $\{\mathbf{x}\}$ の写像 $\phi(\mathbf{x})$ の線形和で表される。

$$\mathbf{e}_i = \sum_{j=1}^m a_{ij} \phi(\mathbf{x}_j) \quad (12)$$

ここで a_{ij} は上で定義したカーネル行列 \mathbf{K} の第 i 番目に大きい固有値 α_i に対応する固有ベクトル \mathbf{a}_i の第 j 成分である。ただし \mathbf{a}_i は \mathbf{e}_i が単位ベクトルであるという要請 $(\mathbf{e}_i \cdot \mathbf{e}_i) = 1.0$ により、 $\alpha_i(\mathbf{a}_i \cdot \mathbf{a}_i) = 1.0$ を満足するように基準化されている。この基底ベクトル \mathbf{e}_i 自体は陽には求めることはできないが、写像 $\phi(\mathbf{x})$ の \mathbf{e}_i への射影や、異なる非線形部分空間 \mathcal{V}_1 と \mathcal{V}_2 の各基底ベクトルの内積 $(\mathbf{e}_i^1 \cdot \mathbf{e}_j^2)$ は次に述べるように計算できる。

●カーネル相互部分空間法 (KMSM)

非線形部分空間、つまり特徴空間 \mathcal{F} 上に生成された線形部分空間の成す複数の正準角に基づいて識別する方法を、非線形相互部分空間法 (Kernel Mutual Subspace Method) (以降では KMSM) と呼ぶ^{☆6}。2つの非線形部分空間の成す正準角は両者の基底ベクトル同士の内積が計算できれば、計算法2あるいは計算法3により容易に求まる。たとえばカテゴリ1の第 i 基底ベクトル \mathbf{e}_i^1 とカテゴリ2の第 j 基底ベクトル \mathbf{e}_j^2 との内積は、式(12)に基づいて以下のように求まる。他の基底ベクトルの組合せも同様に求まる。

$$(\mathbf{e}_i^1 \cdot \mathbf{e}_j^2) = \left(\sum_{s=1}^m a_{is}^1 \phi(\mathbf{x}_s^1) \cdot \sum_{t=1}^m a_{jt}^2 \phi(\mathbf{x}_t^2) \right) \quad (13)$$

$$= \sum_{s=1}^m \sum_{t=1}^m a_{is}^1 a_{jt}^2 (\phi(\mathbf{x}_s^1) \cdot \phi(\mathbf{x}_t^2)) \quad (14)$$

$$= \sum_{s=1}^m \sum_{t=1}^m a_{is}^1 a_{jt}^2 k(\mathbf{x}_s^1, \mathbf{x}_t^2) \quad (15)$$

☆4 通常は各ベクトルから分布の平均ベクトルを差し引くが、ここでは平均ベクトルを差し引かない場合を考える。

☆5 さらにコーシー・シュワルツ不等式も満たす必要がある。

☆6 最小正準角に基づく方法は非線形核相互部分空間法と名付けられているが⁵⁾、本稿では他の方法との関係を分かりやすくするために、非線形相互部分空間法という名称を使うことにする。

とに相当し、識別に3次元形状情報を考慮することで、高性能な顔認識が実現できることが理解される。

逆に、視点と光源の相対位置が固定され、顔が運動する場合を考える。この場合でも、各画像間で各点の正確な対応付けができれば、これにより顔の3次元向きや大きさを推定し、これに基づいて正規化すれば同様の照明部分空間を生成することができる。しかし、一般的には正

規化のための十分な特徴点抽出やその対応付けが難しいために、3次元的な正規化に替えて、目鼻などの特徴点を基準にした2次元アフィン変換による正規化が適用される場合が多い。顔向き変動が小さい場合には、このような簡易的な正規化でも、パターン分布は線形部分空間である程度近似できるが、顔向きの変動が大きい場合にはそれも困難となる。また完全拡散反射、凸形状等の仮定が成立しない領域が多くなると、線形部分空間による表現はより困難になる。このような場合には先に述べたように線形部分空間の代わりに非線形部分空間を適用することが有効である。

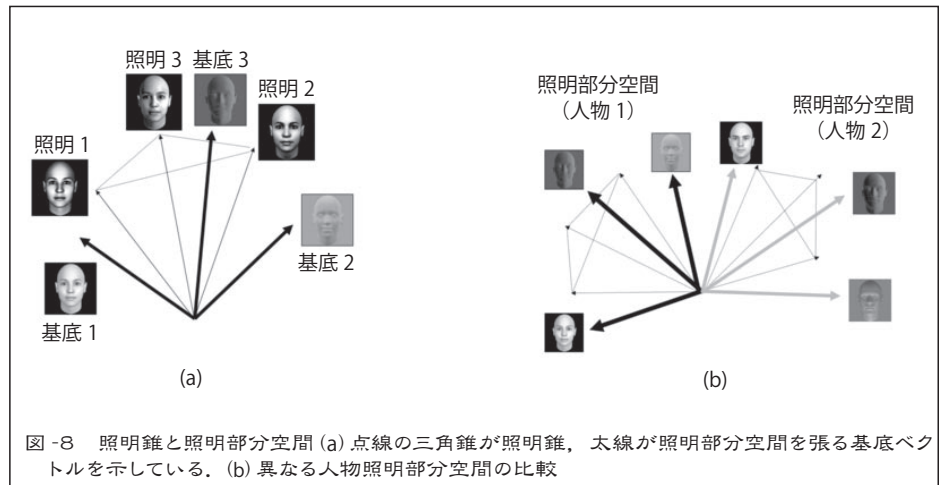
なお、これまでの議論では光源あるいは顔が動く場合であったが、静止した顔に対してマルチカメラシステムを用いて多視点画像を能動的に獲得できれば、上記の議論は成立する。

今後の展開

本稿では部分空間法から相互部分空間法、そして非線形識別への一連の理論拡張について解説した。

この理論拡張の中で中心的な役割を果たしているのが、2つの部分空間の幾何学的な関係を規定する正準角であった。この正準角は汎用的な概念であり、これをベースにしている相互部分空間法はきわめて広範な応用を持つ。本稿では顔画像認識を例として取り上げたが、一般的な3次元物体認識、さらには画像パターン以外のさまざまなパターン識別問題へも適用可能である。

近年、サポートベクタマシン (SVM) とさまざまな特徴抽出を組み合わせた識別アルゴリズムが多数提案されているが、SVM は基本的に2クラス識別器のために、多クラス識別への適用には工夫が必要である。これに対して部分空間法、相互部分空間法は複数クラスを前提にしているために、大規模クラスの識別問題に対しても容易に対応できる。実際、文字認識や顔画像認識などの大



規模な識別問題において実用レベルの成果が得られている。しかしながら、これらの誇るべき成果の割には海外での知名度は低い。そこで世界に向けて部分空間法の有効性や潜在能力を改めて問うために、2007年11月に国際学会 Subspace2007 が開催された。この予稿集は Subspace2007 の Web ページ¹¹⁾ からダウンロード可能である。これは前年に国内で開催された Subspace2006 予稿集¹⁰⁾ と併せて部分空間法に関する貴重な資料となっているので、部分空間法に興味を持った方は参考にしていただきたい。最後に、本稿が部分空間法の有効性、潜在能力を再認識する手助けとなってくれば幸いである。

参考文献

- 1) 柳井晴春：多変量データ解析法 理論と応用，行動計量学シリーズ8，朝倉書店。
- 2) 前田賢一，渡辺貞一：局所的構造を導入したパターン・マッチング法，信学会論 (D)，Vol.J68-D，No.3，pp.345-352 (1985)。
- 3) 山口 修，福井和広：顔向きや表情の変化にロバストな顔認識システム “Smartface”，電子情報通信学会論文誌 (D-II)，Vol.J84-D-II，No.6，pp.1045-1052 (2001)。
- 4) 河原智一，西山正志，山口 修：直交相互部分空間法を用いた顔認識，情報処理 CVIM-151，pp.17-24 (2005)。
- 5) 坂野 鏡，武川直樹，中村太一：核非線形相互部分空間法による物体認識，信学論 (D-II)，Vol.J84-D-II，No.8，pp.1549-1556 (2001)。
- 6) 福井和広，山口 修：部分空間法の理論拡張と物体認識への応用，情報処理学会論文誌 コンピュータビジョンとイメージメディア，Vol.46，No.SIG 15 (CVIM 12)，pp.21-34 (2005)。
- 7) 前田英作，村瀬 洋：カーネル非線形部分空間法によるパターン認識，信学論 (D-II)，Vol.J82-D-II，No.4，pp.600-612 (1999)。
- 8) 津田宏治：ヒルベルト空間における部分空間法，電子信学論 (D-II)，Vol.J82-D-II，No.4，pp.592-599 (1999)。
- 9) Fukui, K. and Yamaguchi, O. : The Kernel Orthogonal Mutual Subspace Method and Its Application to 3D Object Recognition, ACCV2007, pp.467-476 (2007)。
- 10) 部分空間法研究会 Subspace2006 予稿集，<http://www.viplab.is.tsukuba.ac.jp/ss2006/>
- 11) ACCV Workshop Subspace2007 Proceedings，<http://www.viplab.is.tsukuba.ac.jp/ss2007/>

(平成 20 年 4 月 14 日受付)

福井和広 (正会員) | kfukui@cs.tsukuba.ac.jp

筑波大学大学院システム情報工学研究科准教授。1988年九州大学大学院工学研究科機械工学専攻修了。(株)東芝研究開発センターマルチメディアラボラトリ主任研究員を経て、2004年筑波大学着任。主な研究テーマはパターン認識、コンピュータビジョン、電子情報通信学会、IEEE 各会員。