

論 文

部分空間の正準角を利用した三次元パターンマッチングの一基礎検討
——顔と顔写真との区別を目的として——前田 賢一^{†a)} 山口 修[†] 福井 和広^{††}

A Fundamental Discussion on 3-Dimensional Pattern Matching Using Canonical Angles between Subspaces——For the Purpose of Differentiating Face and Its Photograph——

Ken-ichi MAEDA^{†a)}, Osamu YAMAGUCHI[†], and Kazuhiro FUKUI^{††}

あらまし 三次元の物体認識のためには、三次元のパターン間の類似性を定義する必要がある。複数方向からカメラで撮影された三次元のパターンを対象として、三次元パターン同士をマッチングする方法について考察する。マッチングの方式として、三次元パターンをカメラで撮影した画像の集合を部分空間で近似し、部分空間の間の正準角を利用したマッチングを提案する。特に、実物と写真を区別するためには、正準角のうち、最少の正準角以外を利用する必要がある、その使い方として、単独で使う場合と、積や和を利用する場合を比較する。また、その効果を実験によって確認する。

キーワード パターン認識, 三次元, 部分空間, 正準角, 相互部分空間法

1. ま え が き

近年、パターン認識で扱われる問題は、二次元から三次元へと範囲を広げつつあり、この傾向は将来に向かって継続するものと想定される。本論文では、三次元パターン認識の一例として、実物と写真とを見分けるといふ問題を取り上げる。必然的に三次元のパターンを認識する必要が生じ、その具体的な方式が必要とされる。

パターン認識のためには、一般に、パターン間の類似性を定義することが必要であり、その類似性の程度によって認識結果を決定するというプロセスがとられる。したがって、三次元の物体を対象とする場合には、まず、三次元のパターン間の類似性を定義する必要がある。

三次元パターンを表現する方法はいろいろ考えられるが、最も普通に考えられるのは、三次元空間のボクセルに、それぞれの点の濃度値を割り当てる方法である。しかし、一般に可視光では内部は見えないため、表面のみが問題となることになる。例えば、本論文で扱う顔認識を例にすると、文献[1]のサーベイがある。表面は、いわゆる距離画像として表現される。しかし、それを求めるには、レンジファインダやステレオカメラを使うのが一般的であり、従来の二次元用認識装置をそのまま利用するわけにはいかない。

ここでは、あくまで二次元の素直な拡張として三次元認識を扱い、従来の二次元を対象とした場合と同じ、単一のカメラで認識するという立場をとる。

三次元パターンの表現法の一つとして、外側からの「見え」を利用することが考えられており、その一つとして、パラメトリック固有空間法[2]が知られている。それは、三次元の対象物を複数の方向から観察した画像の集合で表現するものである。この集合は、固有空間中の多様体上に分布するため、一方向から観測して得られた画像との距離を、多様体と点との距離として定義し、それが一方向から観測された画像と三次元物体とのマッチングの尺度となる。

[†] (株) 東芝 研究開発センター, 川崎市

Corporate Research & Development Center, TOSHIBA Corporation, 1 Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, 212-8582 Japan

^{††} 筑波大学大学院システム情報工学研究科, つくば市

Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, 305-8573 Japan

a) E-mail: ken.maeda@m.ieice.org

しかし、この方式では、あらかじめ登録しておく対象は三次元を反映しているものの、1回の認識対象はカメラで一瞬を撮影したもの、すなわち二次元であった。このため、前述の応用例で、実物と写真とを見分けることを想定した場合、入力固有空間内の1点で表現されるものであるとすると、たまたまある方向から撮影した写真と、同じ方向から観測した実物とを見分けることはできないことになってしまう。

ここで提案するのは、こうした問題を解決するための三次元パターンのマッチング方法である。すなわち、リアルタイム処理が可能であり、具体的な応用として、三次元の対象物と写真との区別をつけることが可能であるような方式である。

基本的なアイデアは、パラメトリック固有空間法と同様に、複数の方向から観測した画像の集合を三次元物体の表現の基本とし、それを部分空間で近似して、部分空間の間の正準角を利用するというものである。これは、二次元のパターン認識で使われた相互部分空間法 [3] を三次元に拡張したものである [4]~[7]。これらの研究で議論されたのは、最小正準角以外も利用するということであった。文献 [8] では、最初にこの可能性が指摘され 2 個の正準角に対応する固有値の和が提案された。引き続き文献 [4], [5] でも、複数の正準角の余弦の 2 乗値の和を用いるという方式に言及された。

文献 [6], [7] では、三次元の物体 (顔) と写真とを区別するという具体例に対して、3 番目に小さい正準角 (第 3 正準角: 以下, n 番目に小さい正準角を「第 n 正準角」と呼ぶ) が効果があるということが示された。

ここでは、更に詳細に検討して、第 2 正準角や第 4 正準角の効果も実験する。また、これらを組み合わせた使い方のバリエーションとして、和 (平均) と積についても比較し、その意味するところを議論する。

2. 問題へのアプローチ

2.1 相互部分空間法の概要

三次元パターンマッチング法は、従来の二次元パターンに対して用いられた方法の自然な拡張である方が望ましいと思われる。本論文で取り上げる方式は、相互部分空間法の三次元への拡張である。相互部分空間法とは、部分空間法 [9] を拡張したもので、既知のパターン集合 (辞書) だけでなく、未知のパターン集合 (入力) も部分空間で表現し、両者の部分空間のなす最小の角 (正準角) の余弦の 2 乗値を類似度として用いる方式である [3]。部分空間の間の正準角とは、

principal angle あるいは *canonical angle* と呼ばれ、互いに直交する方向で測った部分空間が交わる角度である [10], [11]。

辞書の部分空間は、通常の部分空間法で行われるように、様々な収集パターンから主成分分析によって作成することができる。また、入力に関しても同様であるが、カメラと相対的に移動し、見える角度が異なるように撮影された動画を用い、適当な時間窓で区切ることにより、自然に様々なパターンが得られる。

この方式は、もともとは手書き漢字認識のために開発されたものであるが、後に顔認識にも応用された [12]。当初は、三次元ということよりも、入力変動するということを利用して認識に利用したいという思想で応用されたものである。

正準角を求めるには、辞書部分空間への射影行列を P 、入力部分空間への射影行列を Q としたとき、積 PQP 、または QPQ の固有値を計算し、これが正準角の余弦の 2 乗であることを使う。この行列は次元数が高いため、実際には、等価な問題に変換し、

$$X = (x_{ij}) \quad (1)$$

ただし、

$$x_{ij} = \sum_{m=1}^M (\psi_i, \phi_m)(\phi_m, \psi_j) \quad (2)$$

または、

$$x_{ij} = \sum_{n=1}^N (\phi_i, \psi_n)(\psi_n, \phi_j) \quad (3)$$

なる行列 X の固有値を計算する。すなわち、

$$Xz = \mu_k z \quad (4)$$

を解く。ただし、ここで、 M 次元の辞書部分空間の正規直交基底を $\{\phi_m\}_{m=1}^M$ 、 N 次元の入力部分空間の正規直交基底を $\{\psi_n\}_{n=1}^N$ とする (正規直交基底の作成法は付録 1. に示す)。このとき、正準角を $\{\theta_k\}_{k=1}^K$ とすると、

$$\cos^2 \theta_k = \mu_k \quad (5)$$

として求めることができる。この最小の正準角 (最大固有値) を部分空間の間の類似度としたのが相互部分空間法である。

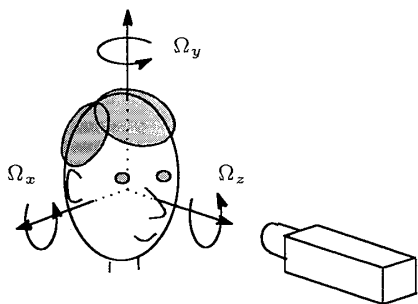


図1 顔の回転
Fig. 1 Rotations of a face.

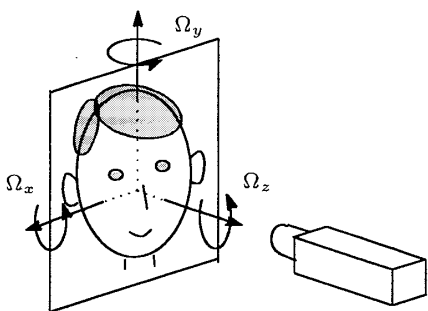


図2 顔写真の回転
Fig. 2 Rotations of a face photograph.

2.2 正準角の利用法

動画を利用した顔認識においては、最小の正準角だけが用いられていた[12]. 更に環境条件の変動を吸収する工夫が加えられたバージョン[8]では、複数正準角の利用が提案されたが、三次元との関係が明確に示されてはなかった.

しかし、少なくとも同一人物の実際の顔と写真を区別するという問題(図1, 図2のように、単一のカメラによって撮影された画像から両者を区別する問題)を扱うためには、最小の正準角では不十分であることが指摘されている. その理由は、写真が撮影されたのと同じ方向から見た顔は、写真と一致するためである. すなわち、この場合、最小の正準角は0度となる. 2個の部分空間に一致する部分がある状況は、模式的に、図3に示される.

これを解決するためには、最小の正準角以外を明示的に利用する必要がある. 文献[6], [7]では、これを目的として、第3正準角を利用するという方式が提案されていた. その理論的根拠は、部分空間の相違を論じた文献[13]で示された最大の正準角を使うということからスタートしている. すなわち、二つの部分空間への射影を、それぞれ P , Q としたとき、

$$PQP\mathbf{y} = \mu_{\min}P\mathbf{y} \quad (6)$$

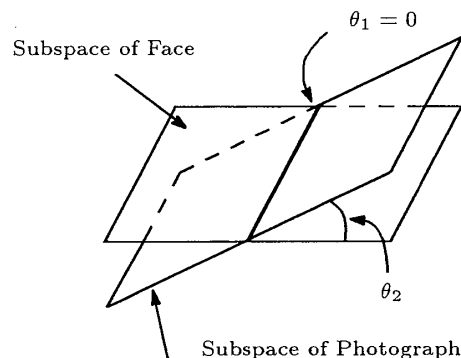


図3 一致する部分を有する2個の部分空間
Fig. 3 Two subspaces which have a common part.

の最小固有値(すなわち最大の正準角)を求めるという方法である(\mathbf{y} を固有ベクトルとする一般化固有値問題). ただし、

$$\|P\mathbf{y}\| = 1 \quad (7)$$

とする.

ここで、三次元物体である顔を縦横2方向に微少回転させることを考えると、パターンは近似的に3次元^(注1)部分空間上に分布すると考えることができる(付録2. 参照). 回転軸は3種類あるが、写真との区別に役立たない方向は考えないこととした. すなわち、図1, 図2において、カメラが顔のほぼ正面にあるとすると、考慮する回転は、 Ω_x , Ω_y だけであり、 Ω_z は考慮しない.

顔の三次元の実物では、回転によって異なる部分が見えたり、陰影の状況が変化する. ところが顔の写真では、形状は投影変換(あるいは近似的にアフィン変換)で変形するだけであり、陰影は全体的に明るさが増えるだけである. この違いは、最小正準角には反映せず、第2正準角と第3正準角の差として現れると考えられる.

文献[6], [7]では、実験的に第3正準角が有効であることが示されていたが、第2正準角や第4正準角が有効か否かという実験はなされていなかった. また、組合せによる効果も議論されていなかった. 一方、文献[4], [5], [8]では、組合せの例として、固有値の和(平均)が提案されたが、実物と写真とを区別するという効果に関しては議論されていなかった.

ここでは、組合せの例として、固有値の平均と積に関しても取り上げる.

(注1): 部分空間の次元数は算用数字で表す.

3. 実験

3.1 単一正準角の効果

ここでは、最小正準角から第4正準角までの効果を、実験によって検証した結果を述べる。データは、10名(男:7名, 女:3名)の被験者から収集し、表1のようなデータセットを作成した。

このデータセットのうち、3と4を既知パターンとし、辞書の作成に使い、残りのデータで照合実験を行った。なお、顔の検出・切出し、正規化などは、文献[8]の方法を用いている。

写真のデータとして、3の最初の正面向きの1枚を表示して、いろいろな方向に動かしたものを用意した(表2)。

なお、実験を通じて、正規化パターンを縦横 14×15 の210次元で表現し、辞書と入力の部分空間の次元数は、それぞれ、7次元とした。入力の動画シーケンス(図4)から主成分分析で作成した基底の例を図5に示す。左側がより重要な主成分である。

この次元数は、付録2.に述べるとおり、パターンの変動が微少な回転だけであれば、3次元でよいはずであるが、実際には、それ以外の変動(位置の変化、表情の変化、等)が加わり、その変動が回転より主要な変動である場合が考えられる。主要なものから3次元だけとった場合には、目的とする回転による変動が無視されてしまう危険性がある。これを防ぐことを目的として、多めに7次元とした。しかし、たまたまの変化であれば、辞書と入力で同じであることはまれであると考えられ、共通する変動が回転であれば、両方

表1 実験データセット

Table 1 Data set for the experiment.

データセット	首振りの種類	数/人
1	縦の首振り (1)	100
2	横の首振り (1)	100
3	縦の首振り (2)	100
4	横の首振り (2)	100
5	縦の首振り (3)	100
6	横の首振り (3)	100
7	いろいろな方向の首振り	100

表2 写真データセット

Table 2 Picture data set.

データセット	首振りの種類	数/人
8	いろいろな方向の回転 (1)	100
9	いろいろな方向の回転 (2)	100
10	いろいろな方向の回転 (3)	100
11	いろいろな方向の回転 (4)	100

の部分空間の重なりは、回転によってもたらされる3次元の部分が多くなると考えられる。

以下の表は、実験結果として、本人の実物に対する最小類似度と本人の写真に対する最大類似度を比較したものである。利用した正準角ごとに、それぞれ、表3, 表4, 表5, 表6に示す。

まず、表3から分かるように、ある「本人の実物

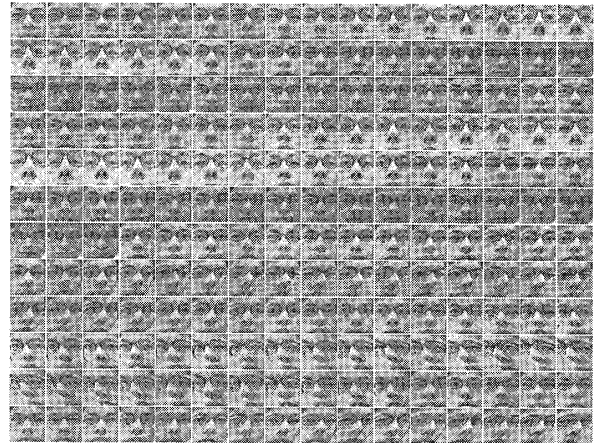


図4 入力動画シーケンスの例

Fig. 4 Example of input image sequence.



図5 部分空間の基底例

Fig. 5 Example of the basis of subspace.

表3 最小正準角を使った場合の類似度

Table 3 The similarity values using the smallest canonical angle.

本人/他人	実物/写真	最大/最小	類似度
本人	実物	最小値	0.953
本人	写真	最大値	0.989
他人	実物	最大値	0.851

表4 第2正準角を使った場合の類似度

Table 4 The similarity values using the 2nd smallest canonical angle.

本人/他人	実物/写真	最大/最小	類似度
本人	実物	最小値	0.907
本人	写真	最大値	0.830
他人	実物	最大値	0.780

表5 第3正準角を使った場合の類似度

Table 5 The similarity values using the 3rd smallest canonical angle.

本人/他人	実物/写真	最大/最小	類似度
本人	実物	最小値	0.809
本人	写真	最大値	0.740
他人	実物	最大値	0.716

表6 第4正準角を使った場合の類似度
Table 6 The similarity values using the 4th smallest canonical angle.

本人/他人	実物/写真	最大/最小	類似度
本人	実物	最小値	0.592
本人	写真	最大値	0.657
他人	実物	最大値	0.582

(Self Face)」に対する類似度の最小値が、別のある「本人の写真 (Self Photograph)」に対する類似度を下回っているケースがある。すなわち、最小の正準角による類似度を使って、適当なしきい値を設定することによって実物と写真とを区別することはできないことが分かる。

次に、表6によると、第4正準角においても、同じ現象が確認される。更に悪いことに、最小の正準角では、実物同士で他人との類似度の差が十分にあり、最悪のケースでも0.102の差があったが、第4正準角では、最悪0.01の差しかないことが分かる。

これに対して、第2正準角、第3正準角ともに、実物と写真とを区別する能力があり、かつ、本人 (Self) と他人 (Others) とを区別する能力もあることが分かる (表4, 表5 参照)。類似度の差を見ると、第2正準角では、実物と写真との差が、最悪で0.077で、本人と他人との差が、最悪で0.127である。一方、第3正準角では、それぞれ、0.069と0.093である。本人の実物に対する類似度の値が、0.907から0.809に減少していることを考え、比率で見ると、実物と写真とを見分ける能力はほぼ等しいが、本人と他人とを見分ける能力は第2正準角の方が優れていることになる。

別の観点で評価するために、FAR (False Acceptance Rate) と FRR (False Rejection Rate) とをプロットした。それぞれ、図6, 図7, 図8, 図9に示す。これによって、最悪のケース以外に、全体としての類似度の分布を把握することができる。

図6によっても、最小の正準角では、本人の写真の FAR (Self Photograph FAR) が本人の実物の FRR (Self Face FAR) と交差しており、区別が不可能であることが確認される。図9も同様に、第4正準角も適していないことが分かる。

第2正準角では、他人の実物に対する FAR が、本人の写真に対する FAR を下回っているのに対して、第3正準角では、逆に本人の写真に対する FAR が、他人の実物に対する FAR を下回る部分が多い。カーブの落ち方は、第2正準角が第3正準角より急しゅんで

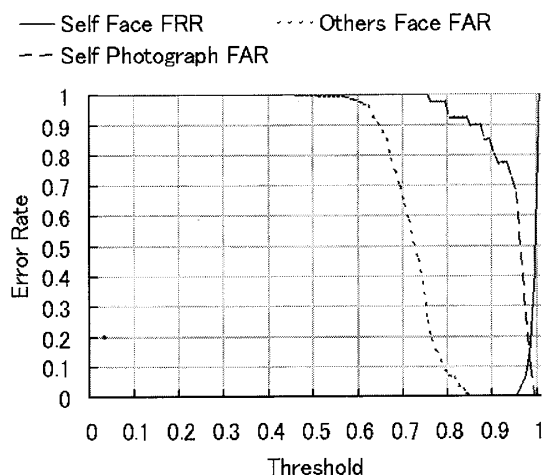


図6 最小正準角を使った場合の FAR-FRR
Fig.6 FAR-FRR curve using the smallest canonical angle.

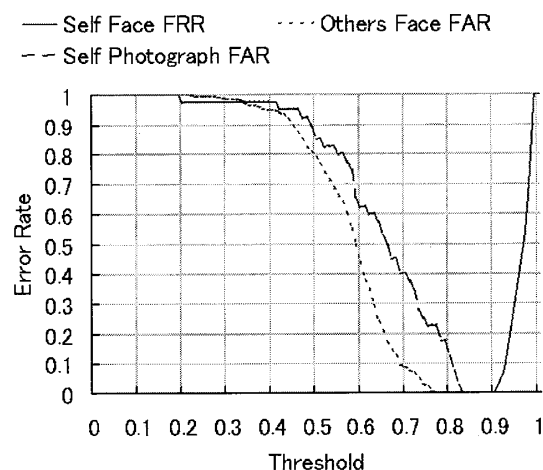


図7 第2正準角を使った場合の FAR-FRR
Fig.7 FAR-FRR curve using the 2nd smallest canonical angle.

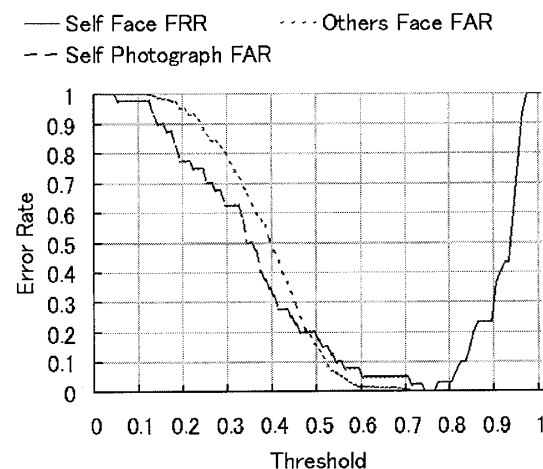


図8 第3正準角を使った場合の FAR-FRR
Fig.8 FAR-FRR curve using the 3rd smallest canonical angle.

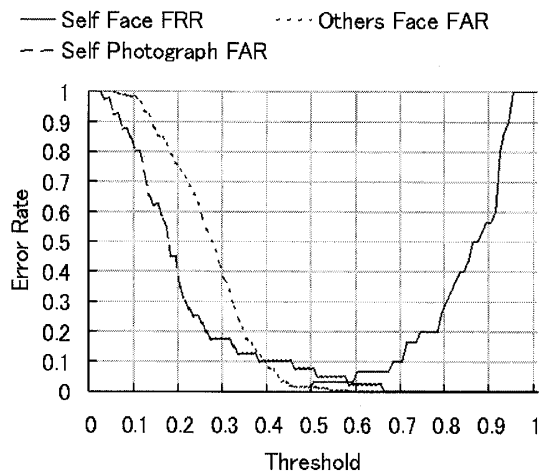


図9 第4正準角を使った場合の FAR-FRR
Fig.9 FAR-FRR curve using the 4th smallest canonical angle.

表7 平均を使った場合の類似度

Table 7 The similarity values using the mean.

本人/他人	実物/写真	最高/最低	類似度
本人	実物	最低値	0.904
本人	写真	最高値	0.828
他人	実物	最高値	0.770

表8 積を使った場合の類似度

Table 8 The similarity values using the product.

本人/他人	実物/写真	最高/最低	類似度
本人	実物	最低値	0.730
本人	写真	最高値	0.537
他人	実物	最高値	0.455

あるが、FRRとFARの間の谷間は、第2正準角が第3正準角より小さいことも分かる(図7, 図8参照)。

3.2 正準角の組合せによる効果

前節では、最小から第4正準角までの単独の照合能力を実験によって検証した。他の可能性として正準角の組合せを利用することが考えられる。ここでは、そのうち、最小から第3正準角までに対応する固有値の平均と、最小から第3正準角までに対応する固有値の積とを類似度として利用した場合を、実験的に検証する。

表7, 表8は、それぞれの場合の類似度を示したものである。

平均を使った場合と積を使った場合、いずれも、本人の実物(Self Face)に対する最低類似度が、本人の写真(Self Photograph)及び他人の実物(Others Face)に対する最高類似度より大きいことから、これらの類似度が有効であることが分かる。

前節と同様に、FARとFRRもプロットしてみる。

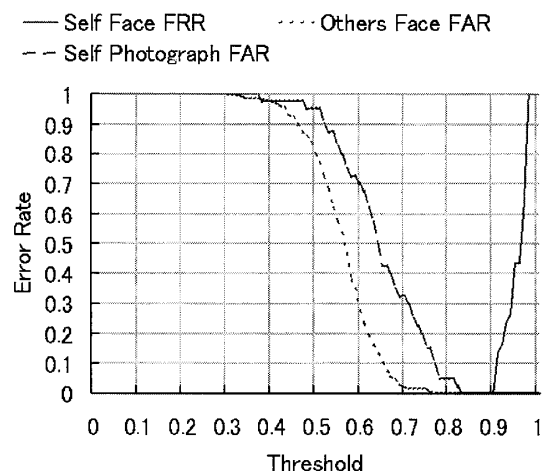


図10 平均を使った場合の FAR-FRR
Fig.10 FAR-FRR curve using the mean.

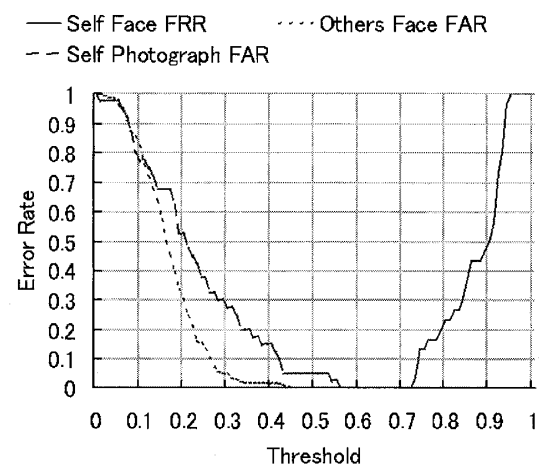


図11 積を使った場合の FAR-FRR
Fig.11 FAR-FRR curve using the product.

結果を、図10, 図11に示す。

これらの図によっても、いずれも有効であることが分かる。図10は、他の図と比較した場合に、図7に形状が似ている。細かく比較してみると、平均を使った場合の方が、第2正準角を使った場合よりFRRとFARとの間の差がわずかに大きいことが分かる。これは、平均を使う方が照合性能がわずかに優れていることを示す。

一方、図11は、他の図よりは、図8に類似している。しかし、それは比較の問題であり、双方が類似しているといえるほどではない。照合性能に関しては、FRRとFARとの間に大きな差があるため、優れていることが期待できる。

4. 考察

4.1 類似度の定義の意味

ここで仮定したのは、次の3点である。

- 観測方向の変化が微小であるため部分空間で近似できること。

- 観測方向の変化は横と縦の回転（図 1, 図 2 の Ω_x と Ω_y ）であること。

- もとのパターンと、この回転とによって部分空間の次元が近似的に 3 次元となること（付録 2. 参照）。この仮定のもとで、2 個の部分空間があるとき、部分空間同士の類似度をどのように定義するかということが、最初の問題である。

相互部分空間法で、最小の正準角を利用したということは、それぞれの部分空間に属する 2 本のベクトルのなす最小の角度を測定したことに相当する [3]。これに対して、第 3 正準角を利用するという事は、最も遠い方向に角度を測定したことになる。

次のアナロジーを考えてみる。2 個の物体があり、それらが同じかどうかを比較する場面を想定する。その際、最も似たところを見るという考え方がある。実際に物体を観測することができず、いろいろな角度から写した比較的少数の静止画だけが与えられたとする。どの方向から写したかという情報がないケースを考えると、ほぼ一致したものがあれば、同じ物体とみなしてもよいという考え方があるはずである。これは、最小の正準角を使うことのアナロジーである。

同じく 2 個の物体を比較するのに、ほぼ同じ方向から写した比較的多数の静止画が与えられたとしてみよう。同じと見られる静止画の組を取り除いて、最後に残った静止画の組（これも、ほぼ同じと見られる）を比較して、例えば、片方にきずが写っている場合には違くとみなすという考え方もあるはずである。これは、最大正準角を使うことのアナロジーである。

後者の場合に、更に、同じとみなした静止画の組も、完全に同じはずはないので、どの程度似ているかを類似度で表し、それを平均するという考え方もあり得る。これは、正準角に対応する固有値の平均を使うことのアナロジーである。類似性の代わりに、似ていないことをバッドマークとして積算することも考えられる。これが、固有値の積を使うことのアナロジーである。

それぞれは、立場の違いであり、理論的には、いずれももっともらしく思える。実際に応用するに際しては、現実の問題がどの立場をとるべきかを考察する必要がある。例えば、辞書の作成や認識に際して、十分なバリエーションが収集できない場合には、最小正準角（あるいは、それに近い第 2 正準角）を使うのが良さそうである。

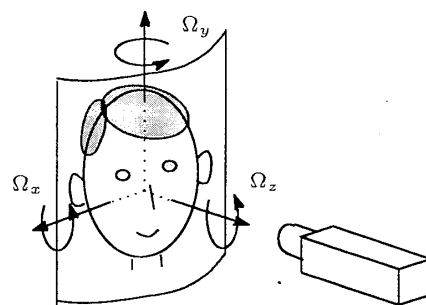


図 12 湾曲させた写真
Fig. 12 Inflected photograph.

4.2 第 2 正準角と第 3 正準角の比較

物体の回転に対して、形状と照明とによって見え方が変化するが、それが、部分空間の正準角にどのような影響を与えるかということは、明確に推定することが困難である。

実験結果から単純に判断すると、単独の正準角を用いる場合には、第 2 正準角の方が第 3 正準角より性能が良さそうである。現象として理解できる理由は、第 2 正準角の方が値が大きく、FRR の立下りが急しゅんであることである。また、FAR（他人、写真ともに）の立下りも急しゅんであり、FAR と FRR の間に、両方とも 0 である領域が広い。

しかし、写真を使って顔認証システムをだますことを考えた場合、写真を傾げるだけでなく、湾曲させる変化も試みられる可能性がある（図 12）。

第 2 正準角と第 3 正準角とは、直接の変形を反映するものではなく、単に二つの部分空間の似ている方向が第 2 正準角に反映しているにすぎない。前記の湾曲によるだましを考えた場合、必ずしも第 2 正準角が優れているかどうかは不明である。ただし、これを検証する場合には、パターンの見え方と照明条件のいずれが分布により大きな影響を与えているかということを明らかにする必要があると考えられる。現時点では、両方が混在したままであり、たまたまある照明条件のもとで実験を行っているにすぎない。したがって、本論文の実験条件としては、写真の変形は、湾曲させずに微小回転変形のみ限定したもの（図 2 参照）であるということに言及する必要がある。

4.3 平均と積の比較

平均と積とを比較した場合、類似度の値そのものは、当然のことながら、平均の方が大きい。これは、FRR の急しゅんな立下りに反映されている。一方で、他人や写真に対する類似度も、積と比較して大きい。平均を使った特性は、第 2 正準角を使った場合に似ている

(図 10 と図 7 の類似性より). しかし, 多くの場合, 積を使った方が, FAR の曲線と FRR の曲線との間に大きな差があることにも注目する必要がある.

ここで, 前述のアナロジーではなく, 数学的に固有値の平均が意味するものを考えてみる. 二つの部分空間の次元が 3 次元であったとしてみる. この場合, 式 (1) に示す行列 X は 3 次元である. この行列の 3 個の固有値の平均は, この行列のトレースの $1/3$ に等しい. 式 (2), または式 (3) に注目し, x_{ii} の平均を求めると,

$$\sum_{i=1}^3 x_{ii}/3 = \sum_{i=1}^3 \sum_{m=1}^3 (\psi_i, \phi_m)^2/3 \quad (8)$$

となる. この式が意味するところは, 二つの部分空間の正規直交基底の単純類似度 (ベクトルのなす角の余弦の 2 乗) の平均 (正確には, その 3 倍) である.

次に, 固有値の積が意味するものを考えてみる. ここでも, 部分空間の次元が 3 次元であるとする. 第 1 の部分空間中に, 辺の長さが 1 である立方体を考えると, その体積は 1 となる. その立方体の各辺が, 正準角を測る方向に一致するようにおく. これを第 2 の部分空間に射影し, それを再度, 第 1 の部分空間に射影したとする. そうすると, 各辺は, $\cos^2 \theta_k$ ($k = 1, 2, 3$) 倍になる (図 13). この立方体の体積は, 固有値の積になっている.

以上が, 固有値の平均と積に対する幾何的な解釈である.

どれが良いかというのは, 実験的に優劣をつけることも可能であるが, 先に述べたように, あくまで, どのように評価したいかということに依存していると考えられる.

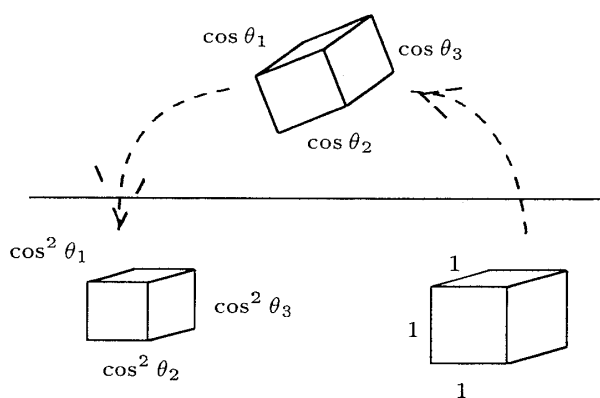


図 13 立方体の射影

Fig. 13 Projection of a cube.

5. む す び

部分空間の間の正準角を利用することにより, 三次元パターン間の類似度を定義するという方式を検討し, また, 実験によって, 以下の点を明らかにした.

- 三次元物体と写真とを区別するのに, 最小正準角だけでは不十分である.
 - 正準角を単独で用いる場合, 少なくとも平面的な写真に対しては, 第 2 正準角が, 最も有効で, 次は第 3 正準角である.
 - 第 4 正準角は, 第 2 正準角, 第 3 正準角と比較して有効性が少ない.
 - 第 3 正準角までに相当する固有値の平均は, 第 2 正準角単独の結果と類似する.
 - 第 3 正準角までの固有値の積は, 小さな値になるものの, FAR と FRR の曲線間の差が大きい.
- 以下の点は, 本論文の範囲を越えており, 今後に残された問題である.
- カメラの光軸回りの回転を考慮した場合 (図 1, 図 2 における Ω_z を考慮した場合).
 - 円筒状に曲げた写真への対応 (これによる類似度の変化と対策).
 - 制約空間を利用した場合 (文献 [8] で提案された拡張に対しても, 本論文で提案した方式が有効であるかどうか).
 - 形状, 照明変化が部分空間の正準角に与える影響の明確化.
 - 顔と顔写真の区別という限られた問題から, 一般的な三次元物体認識への拡張.

謝辞 本研究を進めるにあたり, 有益な議論, データ収集, 解析ツール作成などの協力をして頂いた関係者に感謝します.

文 献

- [1] K.W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches to three-dimensional face recognition," Proc. ICPR04, vol.1, pp.358-361, 2004.
- [2] 村瀬 洋, シュリー ナイヤー, "2 次元照合による 3 次元物体認識," 信学論 (D-II), vol.J77-D-II, no.11, pp.2179-2187, Nov. 1994.
- [3] 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," 信学論 (D), vol.J68-D, no.3, pp.345-352, March 1984.
- [4] K. Fukui and O. Yamaguchi, "Face recognition using multi-viewpoint patterns for robot vision," Proc. ISRR03, pp.192-201, 2003.
- [5] 福井和広, 山口 修, "一般化差分部分空間に基づく制

約相互部分空間法,” 信学論 (D-II), vol.J87-D-II, no.8, pp.1622-1631, Aug. 2004.

- [6] K. Maeda, K. Fukui, and O. Yamaguchi, “Towards 3-dimensional pattern recognition,” Proc. SPR04, pp.1061-1068, 2004.
- [7] 前田賢一, 山口 修, 福井和広, “部分空間の正準角による3次元パターンのマッチング法,” 信学技報, PRMU2004-70, 2004 (情処学研報, 2004-CVIM-145-17, pp.129-136, 2004).
- [8] 福井和広, 山口 修, 鈴木 薫, 前田賢一, “制約相互部分空間法を用いた環境変動にロバストな顔認識—照明変動を抑える制約部分空間の学習,” 信学論 (D-II), vol.J82-D-II, no.4, pp.613-620, April 1999.
- [9] エルッキ オヤ (著), 小川英光, 佐藤 誠 (訳), パターン認識と部分空間法, 産業図書, 1986.
- [10] H. Hotelling, “Relations between two sets of variates,” Biometrika, vol.28, pp.321-372, 1936.
- [11] F. Chatelin (著), 伊理正夫, 伊理由美 (訳), 行列の固有値, シュプリンガー・フェアラーク東京, 1993.
- [12] 山口 修, 福井和広, 前田賢一, “動画像を用いた顔認識システム,” 信学技報, PRMU97-70, 1997.
- [13] E. Oja and J. Parkkinen, “On subspace clustering,” Proc. ICPR '84, pp.692-695, 1984.

付 録

1. 部分空間の正規直交基底の作成法

あるカテゴリーに属するパターン (切出し, 正規化などの処理は, 例えば, 文献[8]のように行われた後のもの) の集合を $\{f_\alpha\}$ とする. 最初に相関行列 (特性核とも呼ばれる) K を式 (A.1) のように作成する.

$$K = \sum_{\alpha} w_{\alpha} \langle f_{\alpha}, f_{\alpha} \rangle \quad (\text{A.1})$$

ただし, w_{α} は, 総和が1になるような重みで, 通常はパターンの個数の逆数が用いられる. 次に, K の固有値問題,

$$K\phi_m = \lambda_m\phi_m \quad (\text{A.2})$$

を解く. λ_m は, 大きい方からソートされているとすると, 上から M 個をとることにより, $\{\phi_m\}_{m=1}^M$ を得ることができる.

2. 顔の微少回転による部分空間の次元数

顔を横と縦に回転した場合, そのパターンの分布は, パターンの存在する空間 (本論文の実験の場合は 210 次元) 中の多様体 (*manifold*) 上に乗っている. 回転が微少である場合には, これを正面パターンの位置で接する接平面で近似することができる. 実際に回転したパターンの分布を図 A.1 に示す.

照明が α 倍強くなったとすると, パターンを表すベ

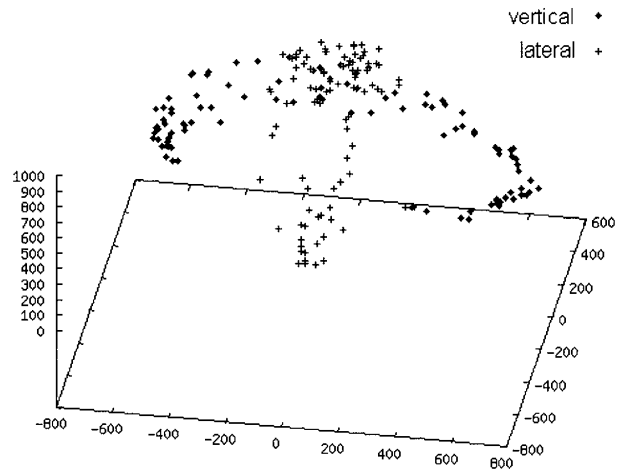
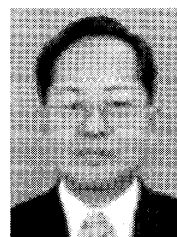


図 A.1 顔の横と縦の回転による多様体

Fig. A.1 The manifold created by lateral and vertical rotations.

クトルのノルムも α 倍になるため, パターンの分布を 3 次元で近似するのが適当である.

(平成 17 年 10 月 4 日受付, 18 年 1 月 13 日再受付)



前田 賢一 (正員)

1974 東工大・工・電子物理卒. 1976 同大学院理工学研究科修士課程了. 1976 (株) 東芝入社. 現在, 同社研究開発センター技監. パターン認識, コンピュータアーキテクチャの研究に従事. 1989~1990 エジンバラ大学 AIAI 客員研究員. 1980 本会学術奨励賞受賞. IEEE Senior Member. 情報処理学会会員.



山口 修 (正員)

1992 岡山大・工・卒. 1994 同大学院修士課程了. 1994 (株) 東芝入社. 現在, 同社研究開発センター研究主務. コンピュータビジョン・顔画像処理の研究に従事. 1996 情報処理学会全国大会優秀賞, 2002 本会論文賞, 2003 情報処理学会山下記念賞受賞. 情報処理学会会員.



福井 和広 (正員)

1986 九大・工・機械卒. 1988 同大学院修士課程了. 1988 (株) 東芝入社. 同社研究開発センター主任研究員を経て, 2004 筑波大学大学院システム情報工学研究科助教授. パターン認識・コンピュータビジョンの理論と応用に関する研究に従事. 工博. 2002 本会論文賞受賞. IEEE Member. 情報処理学会会員.