

## 相互部分空間法による三次元パターンマッチングにおける正準角の性質

前田 賢一 山口 修

(株) 東芝 研究開発センター マルチメディアラボラトリー  
 〒 212-8582 川崎市幸区小向東芝町 1  
 E-mail:{ken.maeda, osamu1.yamaguchi}@toshiba.co.jp

あらまし 相互部分空間法を利用して顔と顔写真とを識別する場合、最小の正準角以外を使う必要がある。どの正準角を使うべきかという問題に答えるためには、それぞれの正準角の機能・性能を知ることが必要である。ここでは、筒状に凸に湾曲させた写真を用いることによって、それを明らかにする。

### Properties of Principal Angles in 3-D Pattern Matching Using the Mutual Subspace Method

Ken-ichi MAEDA and Osamu YAMAGUCHI

Multimedia Laboratory, Corporate R & D Center, Toshiba Corporation  
 1, Komukai-Toshiba-Cho, Saiwai-ku, Kawasaki 212-8582, Japan  
 E-mail:{ken.maeda, osamu1.yamaguchi}@toshiba.co.jp

**Abstract** It is necessary to use the canonical angles other than the smallest one when the Mutual Subspace Method is applied to differentiate the face and its photograph. In order to find which one to use, it is necessary to understand properties of the canonical angles. We analyze them using curved photographs in cylindrical shape.

## 1 はじめに

前田ら [1, 2, 3] は、「相互部分空間法」を利用して、顔と顔写真との区別をすることを提案した。そのためには、静止画ではなく、顔の縦横の回転 (図 1 参照) を含む動画が利用され、パターンの変形分布は部分空間で近似された。

相互部分空間法とは、前田ら [4] によって提案されたもので、辞書と入力との両方のパターン分布を部分空間で近似し、部分空間の間の正準角 [5, 6] の  $\cos^2$  を類似度とする方式である。この値は、

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

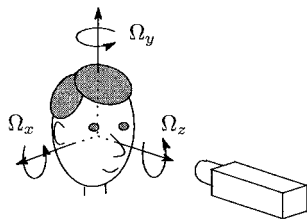


図 1: 顔の回転  
 Fig.1 Rotations of a Face.

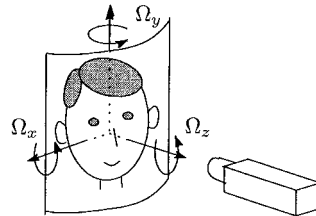


図 2: 湾曲させた写真の回転  
 Fig.2 Rotation of a Curved Photograph.

あるいは、

$$x_{ij} = \sum_{n=1}^N (\phi_i, \psi_n) (\psi_n, \phi_j) \quad (2)$$

という行列の固有値として求めることができる。(ただし、 $\{\phi_m\}_{m=1}^M$  および  $\{\psi_n\}_{n=1}^N$  は、部分空間の正規直交基底である。)

従来、最小正準角以外の正準角 (2 番目や 3 番目に小さい正準角) を利用することが述べられていたが、それぞれの機能や性能に関しては、十分に明確であるとはいえなかった。ここでは、先の文献 [3] で示唆された、より困難な対象である筒状に凸に湾曲させた写真 (図 2 参照) を使うことによって、これらを明確にする。

また、先行研究で明確にできなかった理由を明らかにし、同時に、それを避ける方法を示す。

## 2 問題と理論的背景

前田らの最初の提案 [1] では、第3正準角（以後、 $n$  番目に小さな正準角を第  $n$  正準角と呼ぶ）を利用することによって顔と顔写真とを区別できるとした。その理由は、顔の縦横の微小回転によって生成されるパターンの分布が3次元部分空間で近似されるためと主張されていた。実際、先の文献 [3] では、その様子が図3のように示されていた。

しかし、この文献によれば、第2正準角も第3正準角も、わずかな差はあるものの、同程度に有効であるように見えた。第2正準角と第3正準角との優劣が不明確である理由として考えられるのは、立体的な顔と平面の写真とは、回転変形させた際に第2正準角が第3正準角と同様に離れているためではないかということである。この仮説を検証するには、筒状に凸に湾曲させた写真を使ってみるのが有効であると思われる。その理由は、 $\Omega_y$  の回転に対して、顔と写真との差が少なくなることが期待され、それによって第2正準角と第3正準角との間の違いが顕在化するとと思われるためである。

## 3 湾曲させた写真による実験

実験は、過去の実験 [3] で使われたのと同じものを利用した。すなわち、10名（男：7名、女：3名）の被験者から、画面に表示したターゲットを顔全体で追従するように指示して、縦、横、さまざまな方向に分けて、複数回採取したもので

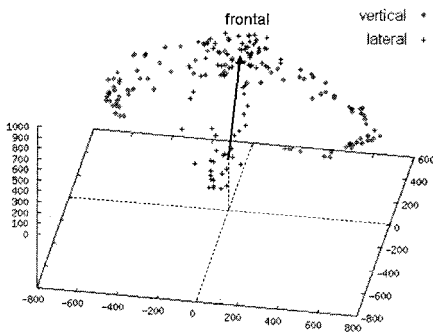


図3: 顔の横と縦の回転による多様体  
Fig.3 The Manifold Created by Lateral and Vertical Rotations.

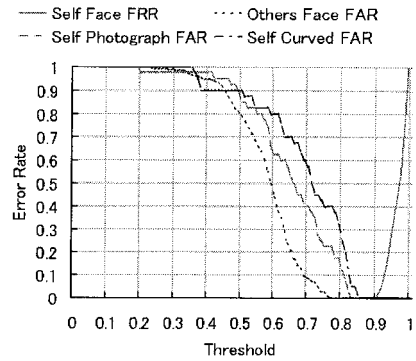


図4: 第2正準角によるFRR/FARカーブ  
Fig.4 FRR/FAR Curves Using the 2nd CA.

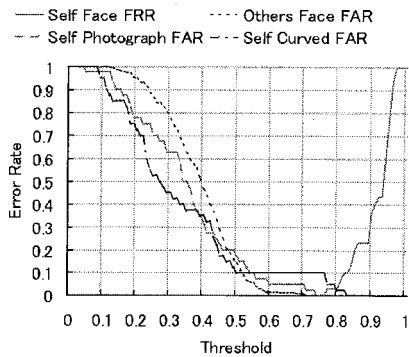


図5: 第3正準角によるFRR/FARカーブ  
Fig.5 FRR/FAR Curves Using the 3rd CA.

ある。正規化後の画像サイズは、縦横が  $14 \times 15$  の210次元である。また、辞書と入力の部分空間の次元数を、それぞれ7次元とした。

図4, 5に、それぞれ、第2, 第3正準角を使って、実際に湾曲させた写真を識別した実験の結果を示す。

まず、図4からわかるのは、湾曲させた本人の写真 (Self Curved) に対するFAR (Self Curved FAR) が平らな写真 (Self Photograph) に対するFAR (Self Photograph FAR) より右側にシフトしているということである。したがって、自分の顔 (Self Face) に対するFRR (Self Face FRR) との差が、湾曲によって小さくなっているということになる。

表 1: 写真に対する類似度  
Table 1 Similarities for Photographs.

CA	2nd		3rd	
	Flat	Curved	Flat	Curved
photo01-1	0.711	0.676	0.183	0.128
photo02-1	0.484	0.620	0.186	0.204
photo03-1	0.516	0.507	0.175	0.085
photo04-1	0.587	0.742	0.332	0.271
photo05-1	0.724	0.812	0.453	0.389
photo06-1	0.795	0.809	0.701	<b>0.791</b>
photo07-1	0.736	0.849	0.285	0.436
photo08-1	0.546	0.687	0.178	0.200
photo09-1	0.820	0.791	0.444	0.450
photo10-1	0.413	0.353	0.244	0.223
photo01-2	0.726	0.720	0.337	0.241
photo02-2	0.663	0.583	0.337	0.173
photo03-2	0.805	0.538	0.526	0.088
photo04-2	0.690	0.708	0.268	0.239
photo05-2	0.779	0.792	0.365	0.423
photo06-2	0.744	0.841	0.740	<b>0.824</b>
photo07-2	0.809	0.778	0.378	0.406
photo08-2	0.593	0.699	0.322	0.226
photo09-2	0.663	0.784	0.365	0.461
photo10-2	0.830	0.368	0.533	0.230
photo01-3	0.584	0.735	0.241	0.113
photo02-3	0.584	0.621	0.370	0.167
photo03-3	0.465	0.617	0.402	0.175
photo04-3	0.588	0.714	0.383	0.301
photo05-3	0.500	0.803	0.160	0.449
photo06-3	0.795	0.814	0.559	<b>0.768</b>
photo07-3	0.687	0.818	0.509	0.348
photo08-3	0.579	0.674	0.433	0.227
photo09-3	0.653	0.791	0.327	0.490
photo10-3	0.616	0.373	0.283	0.281
photo01-4	0.635	0.702	0.128	0.101
photo02-4	0.722	0.627	0.360	0.175
photo03-4	0.816	0.536	0.395	0.101
photo04-4	0.649	0.708	0.217	0.328
photo05-4	0.493	0.787	0.140	0.426
photo06-4	0.709	0.821	0.596	<b>0.764</b>
photo07-4	0.200	0.830	0.042	0.421
photo08-4	0.650	0.660	0.407	0.238
photo09-4	0.771	0.814	0.497	0.492
photo10-4	0.561	0.379	0.121	0.261

一方、図 5 では、湾曲させた写真は、平らな写真とほぼ同じ FAR を示している。ただし、湾曲させた写真の FAR が顔の FRR とクロスしている。これは、適当な閾値によって湾曲した写真と顔とを区別することができないものが存在することを示している。

しかし、エラー率が 0.1 であり、データの人数が 10 名であることを考えると、これは特定の 1 名が関係したものではないかと疑われるものである。

そこで、第 2 および第 3 正準角に対して、写真を入力した場合の類似度を調べてみた (表 1 参照)。

ここで、湾曲した写真の自分自身の辞書に対

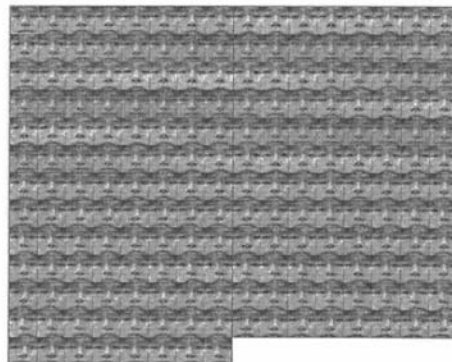


図 6: 人物#06 の顔データ  
Fig.6 Acquired Face Data of Person #06.

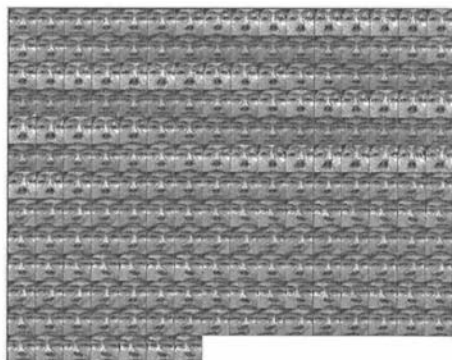


図 7: 人物#03 の顔データ  
Fig.7 Acquired Face Data of Person #03.

する第 3 正準角の類似度 (3rd, Curved) を見ると、photo06-n の値だけが大きいことがわかる。これは、人物#06 の辞書が特殊であることを表わしていると考えられる。そこで、人物#06 の辞書を作成したデータを調査した。図 6 にデータを表示する。また比較のために人物#03 のデータを図 7 に示す。

これからわかるのは、人物#06 がほとんど顔を動かしていないということである (データ収集に際しては、画面に表示されたマークを顔全体で追いかけるように指示されていた)。したがって、人物#06 の辞書に、三次元物体としてのさまざまな方向からの画像が含まれていないことがわかった。そこで、人物#06 を除いて、再度湾曲した顔による実験を実施することにした。

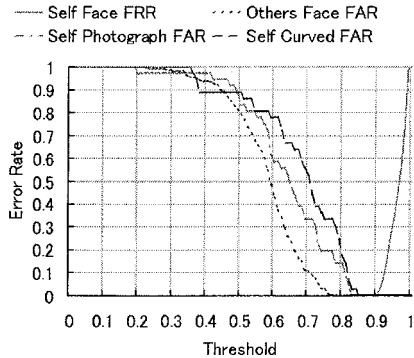


図 8: 第 2 正準角による FRR/FAR カーブ (#6 を除く)

Fig.8 FRR/FAR Curves Using the 2nd CA (without #6).

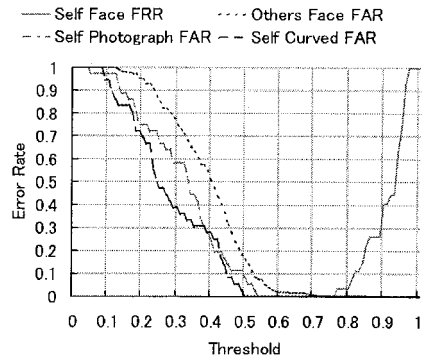


図 9: 第 3 正準角による FRR/FAR カーブ (#6 を除く)

Fig.9 FRR/FAR Curves Using the 3rd CA (without #6).

#### 4 湾曲した顔による再実験

人物#06 を除いた湾曲した顔による最実験の結果を、図 8、図 9 に示す。

性能の基準として、エラー率が 0 のときの FRR と FAR とのギャップの大きさを採用することになると、写真に対して（平面的なものも含めて）第 3 正準角の類似度が効果的であるということがわかる。一方、本人と他人との識別には第 2 正準角の方が効果があることもわかる。

面白い現象として、個人識別において正準角の Boosting を行なうと、第 2 正準角の重みが第 1 正準角より大きくなるという現象が Kim ら [7] によって報告されている。参考のため、ここで利用したデータで第 1 正準角の性能を図 10 に示す。本人の FAR と他人の FRR との間のギャップが、わずかに狭いことが見て取れる。

さらに、複数の正準角の組合せの効果を調べるため、類似度の平均 [3, 8] と積 [3] に関する結果をプロットした（図 11, 図 12 参照）。

両方とも効果があるが、類似度の積を用いるのが最も効果的であることがわかる。ただし、本人に対する類似度が小さくなるので、識別の閾値を適当に小さくする必要がある。

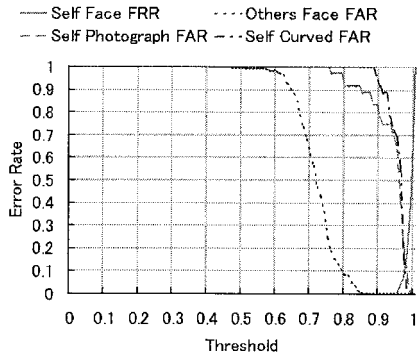


図 10: 第 1 正準角による FRR/FAR カーブ (#6 を除く)

Fig.10 FRR/FAR Curves Using the 1st CA (without #6).

#### 5 考察

前節で述べたように、再実験により第 2 正準角、第 3 正準角の性質や組合せの性質が明確になった。しかし、不適切な辞書データを発見するのにデータの写真を目で見るのはたいへんである。そこで、佐藤 [9] の示唆により、辞書部分空間を作る際の K-L 展開の固有値を調べることにした。人物#03 と人物#06 とに対する固有値を図 13 に示す。

これを見ると、人物#06 の固有値分布が第 1

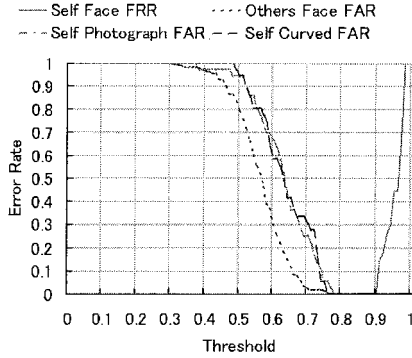


図 11: 類似度の平均による FRR/FAR カーブ (#6 を除く)

Fig.11 FRR/FAR Curves Using the Mean (without #6).

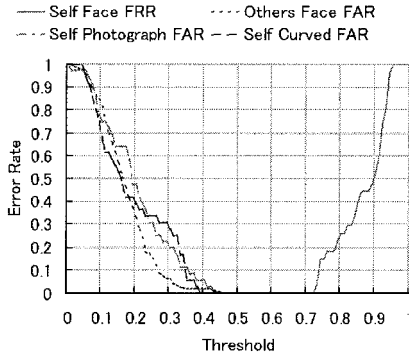


図 12: 類似度の積による FRR/FAR カーブ (#6 を除く)

Fig.12 FRR/FAR Curves Using the Product (without #6).

固有値に集中していることがわかる。これが意味するのは、バリエーションが少ないということである。従って、辞書を作る際の固有値を図のようにプロットすることにより、容易に不適切な辞書データを見分けられることがわかった。

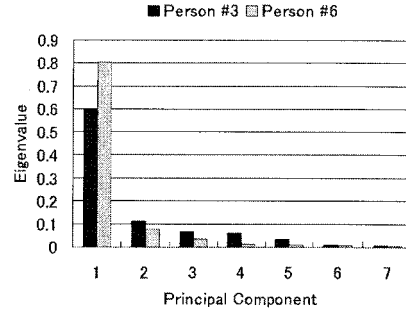


図 13: 辞書部分空間の固有値

Fig.13 Eigenvalues of Reference Subspaces.

## 6 まとめ

先行研究 [3] で、顔と写真との識別に関して、第 2 正準角と第 3 正準角との優劣が明確でなかったのは、平面的な写真のみしか扱わなかったことに起因していた。また、湾曲した写真を使った実験でも優劣がつかなかったのは、辞書データの中に顔を十分に動かしていないもの (人物 #06) があったためであった。これを除いた再実験の結果、顔と写真との識別には単独では第 3 正準角が最も有効であることがわかった。ただし、本人と他人との識別には第 2 正準角の方が有効であった。また、類似度の組合せとしては、積を用いるのが最も効果的であることがわかった。さらに、不適切な辞書データを見分けるには、辞書作成に使われた K-L 展開の固有値をプロットすることが有効であることを示した。

今後の課題として、十分な量のデータで効果を検証する必要がある。

**謝辞** 不適切な辞書データの見分けに K-L 展開の固有値を見ることを示唆していただいた、東京工業大学の佐藤 誠 教授に感謝いたします。日頃の有益な議論、データや解析ツールの提供をしていただいた社内共同研究者の各位と筑波大学の福井 和広 准教授に感謝いたします。

## 参考文献

- [1] K. Maeda, O. Yamaguchi, and K. Fukui: "Towards 3-dimensional pattern recognition," *Proc. SPR04*, pp. 1061-1068, Aug. 2004.

- [2] 前田賢一, 山口修, 福井和広: “部分空間の正準角を利用した3次元パターンマッチング法,” 信学技報, PRMU2004-70, pp. 129–136, Sep. 2004. (情処学研報, 2004-CVIM-145-17, pp. 129–136, Sep. 4004.)
- [3] 前田賢一, 山口修, 福井和広: “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 —,” 信学論 (D), vol. J89-D, no. 6, pp. 1288–1296, Jun. 2006.
- [4] 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, Mar. 1984.
- [5] F. Chatlin: *Valeurs Propres de Matrices*, Masson, 1998. (W. Ledermann, *Eigenvalues of Matrices*, John-Wiley and Sons, 1993.)
- [6] H. Hotelling: “Relations between two sets of variates,” *Biometrika*, 28, pp. 321–372, 1936.
- [7] T. Kim, O. Arandjelovic and R. Cipolla: “Learning over Sets Using Boosted Manifold Principal Angles (BoMPA),” *BMVC*, Sep. 2005.
- [8] K. Fukui and O. Yamaguchi: “Face Recognition Using Multi-viewpoint Patterns for Robot Vision,” *Proc. of ISRR03*, 2003.
- [9] 佐藤誠, 前田賢一: 口頭での議論, Nov. 2006.