

局所部分空間集合を用いたアンサンブル識別に基づく3次元物体認識

呉 嘉寧[†] 福井 和広[†]

[†] 筑波大学大学院

システム情報工学研究科コンピュータサイエンス専攻

〒305-8577 茨城県つくば市天王台 1-1-1

E-mail: †lacarte@viplab.is.tsukuba.ac.jp, ††kfukui@cs.tsukuba.ac.jp

あらまし 3次元物体認識に複数視点を導入することで識別性能を向上させることが可能である。一般に複数視点から得られた特徴ベクトルの分布が非線形性を持つため、カーネルトリックを用いた非線形識別手法がよく用いられる。この非線形識別手法は特徴ベクトルを高い次元の空間に写像し非線形性を弱めることで多くの応用で有効性が示されているが、その計算量は膨大である。そこで本論文ではカーネルを用いた手法と同等の性能を非線形写像なしに達成する手法を提案する。具体的にはまず非線形分布を複数の局所部分空間で近似する。次に局所部分空間の集合をクラスタリング条件と各部分空間の次元の様々な組み合わせに対して生成し、アンサンブル識別を行う。提案法の有効性は公開データセットによる識別実験によって検証する。

キーワード アンサンブル識別, 相互部分空間法, 平均学習部分空間法, 物体認識

3D Object Classification Based on Ensemble Classification with Local Subspaces

Jianing WU[†] and Kazuhiro FUKUI[†]

[†] Department of Computer Science, Graduate School of Systems and Information Engineering,
University of Tsukuba

Tennodai 1-1-1, Tsukuba-shi, Ibaraki, 305-8577 Japan

E-mail: †lacarte@viplab.is.tsukuba.ac.jp, ††kfukui@cs.tsukuba.ac.jp

Abstract Classification performance of 3D object classification can be improved by multiple view points. Kernel-based methods are often introduced to handle the nonlinearly distributed feature vectors obtained from multiple view, by transforming the distribution to a higher dimensional space. However, this nonlinear mapping makes their computation to be complex. We aim to construct a comparable method with the kernel-based methods without using nonlinear mapping. Firstly we approximate a distribution of feature vectors with multiple local subspaces. Secondly we consider these local subspaces as weak classifiers and apply ensemble classification algorithm. We will evaluate the proposed method by classification experiments using a public data set.

Key words Ensemble Classification, Mutual Subspace Method, Averaged Learning Subspace Method, Object Classification

1. はじめに

画像パターンに基づくビューベース物体認識を行うに当たって、単一方向の画像だけではなく複数方向から見た画像を採用することは自然である。このような画像は複数視点カメラの採用や単一カメラで回転などの運動を行う物体を撮影することで得られる。対象画像を特徴抽出、あるいは単に画像を輝度値の2次元配列とみなし、その行を展開してベクトル化する。複数

の入力画像から得られる特徴ベクトル集合を用いた識別手法として、部分空間法の拡張である相互部分空間法 [1] が注目されている。

部分空間法は学習において識別対象となる物体クラスごとに複数の特徴ベクトルを取得し、クラスごとに主成分分析 (KL展開) を行い、各クラスの部分空間を生成する。識別では一本の入力特徴ベクトルと各クラスの部分空間の角度を算出する。入力ベクトルはもっとも角度の小さいクラスに識別される。相

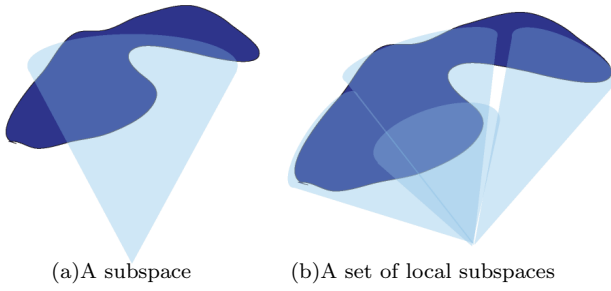


図 1 単一の線形部分空間による近似と複数の局所部分空間による近似の比較
Fig.1 Comparison between approximation using a linear subspace and a set of local subspaces .

互部分空間法 [1] では識別における入力特徴ベクトルを複数にし、これらを主成分分析して入力部分空間する．さらに学習した各クラスの部分空間と入力部分空間の類似度 (正準角, 後述) を算出する．入力部分空間は類似度のもっとも高いクラスに識別される．相互部分空間法は複数視点の入力に基づく識別であり, 3次元物体の識別に有効な枠組みである．

しかし, 複数視点画像では物体の見え方が大きく変化するため, 得られる特徴ベクトルのパターン分布は複雑で非線形的な広がりを持つ可能性が高い．特徴ベクトルのパターン分布を主成分分析によって単一の線形部分空間で近似する相互部分空間法では図 1 の (a) に示すように表現できず, 識別性能が低下する．

非線形的なパターン分布の識別において有効な手法として, 核非線形相互部分空間法 [2] が知られている．この手法はパターン分布を非線形写像を用いて高い次元の特徴空間に写像し, 写像先の特徴空間において線形分離可能なパターンに変換して識別するものであり, 唇の認識などで高い性能を示している [2] . しかし, 学習データ数 n の増加に伴い n^2 の計算量が生じ, 実時間システムへの応用が難しい．

本論文では 3次元物体の識別に伴う非線形的なパターン分布に有効で, かつ非線形手法より少ない計算量で実現できる識別手法を提案する．提案法の主なアイデアは以下の二つである．

まず, 単一の線形部分空間での近似が困難なパターン分布に対して, 図 1 の (b) に示すようにクラスタリングでこれを分割し, それぞれのクラスに対して主成分分析を行って部分空間近似を行い, パターン分布を線形部分空間 (以後局所部分空間と呼ぶ) の集合で近似する．入力部分空間とあらかじめ学習した各クラスの局所部分空間集合の正準角に基づく類似度を算出することで識別を行う．この際, 分割数や次元などのパラメータを最適に定めることが難しいため, 図 2 に示すように分割数と局所部分空間の次元を変えながら, それぞれの場合を弱識別器としてアンサンブル学習を行うことでこのパラメータ選定の難しさに対応する．

また, 前述の弱識別器を構成する過程ではクラス内の分布を再現できるよう構成されているが, クラス間の分離について考慮したとは言えない．そこで, 学習部分空間法 [3] のフレームワークに基づき, 学習フェーズにおいて各局所部分空間にクラ

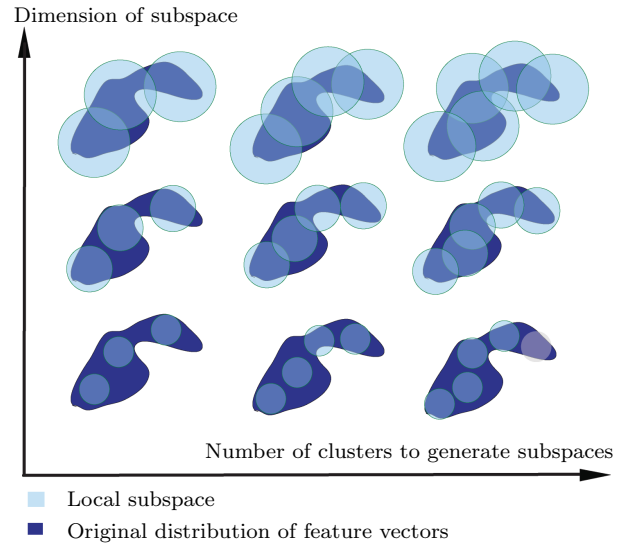


図 2 クラスタ数と局所部分空間の次元の関係
Fig.2 The relation between the number of clusters and the dimension of the local subspaces.

ス間分離を考慮した補正を適用する．

以下本論文ではまず第 2 章で従来手法について説明し, 次に第 3 章で提案手法のアルゴリズムや識別の流れを説明する．第 4 章では公開データベースを用いた従来法, 提案手法および非線形手法の比較実験を行い, 本手法の有効性を示す．第 5 章でまとめと今後の課題について述べる．

2. 従来手法

2.1 部分空間の正準角に基づく類似度

本研究の基礎となる相互部分空間法 [1] では部分空間同士の類似度として, 2つの部分空間の成す角度である正準角を用いる． N 次元の部分空間 P と M 次元の部分空間 Q の間には N 個の正準角が定義できる ($N \leq M$) . P と Q が完全に一致している場合はすべての正準角が 0 度であり, 完全に直交している場合はすべての正準角が 90 度となる．本論文では最小正準角を類似度計算に用いる． P と Q の最小正準角を θ として, その類似度 $Ang(P, Q) = \cos^2 \theta$ は式 (1) で定義する．

$$\cos^2 \theta = \max_{\substack{u \in P, v \in Q \\ \|u\| \neq 0, \|v\| \neq 0}} \frac{|(u, v)|^2}{\|u\|^2 \|v\|^2} \quad (1)$$

$\cos^2 \theta$ は以下の行列 X の最大固有値となる．

$$X = (x_{mn}) \quad (m, n = 1 \dots M) \quad (2)$$

$$x_{mn} = \sum_{l=1}^N (\psi_m, \phi_l)(\phi_l, \psi_n) \quad (3)$$

ここで ψ_m, ϕ_l は部分空間 P と Q の第 m, l 基底ベクトル, (ψ_m, ϕ_l) は ψ_m と ϕ_l の内積を表す．

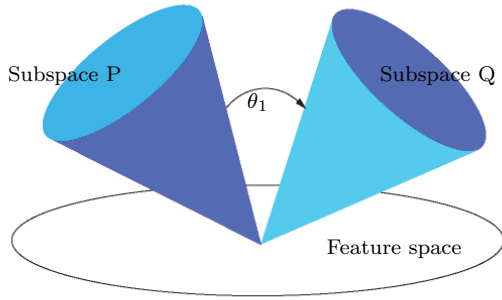


図 3 相互部分空間法の概念図

Fig.3 Concept of mutual subspace method .

2.2 相互部分空間法の非線形拡張

核非線形相互部分空間法 [2] は相互部分空間法の非線形拡張である。パターン分布を高次元の特徴空間へ写像し、特徴空間における正準角に基づいて識別を行う。非線形写像 $\Psi: R^N \rightarrow F$ を定義すると、特徴ベクトル u を特徴空間 F に写像した $\Psi(u)$ が得られる。 F は非常に高次元もしくは無限次元であるため、直接 $\Psi(u)$ と $\Psi(v)$ の内積を求めることは困難であるが、カーネルトリックを用いると F における内積が求まる。これにより、特徴空間 F における部分空間はカーネル PCA [2] を用いて求まる。クラス P' と Q' の学習ベクトルを u と v とすると、 P' と Q' にカーネル PCA を適用して得られる基底 V と W の内積は式 (4) で求められる。

$$(V, W) = \sum_{i=1}^M \sum_{j=1}^{M'} a_i a_j k(u_i, v_j) \quad (u_i \in P', v_j \in Q') \quad (4)$$

a_i, a_j は係数であり、 $k(\cdot, \cdot)$ はカーネル関数である。式 (4) を式 (3) に代入すると F における正準角が計算できる。

3. 提案手法

3.1 局所部分空間集合によるパターン分布の近似

相互部分空間法では特徴ベクトルのパターン分布を単一の線形部分空間で近似したが、第 1 章で述べたように複数視点から得られるパターン分布は非線形的である可能性が高く、単一の線形部分空間では精度の高い近似ができない。そこで提案法では対象パターン分布をユークリッド距離に基づいて複数のクラスに分割し、各クラスをそれぞれ線形部分空間で近似する。

具体的にはまずパターン分布に対し k-means 法でクラスタリングを行う。そして得られた各クラスに対して PCA を適用して各局所部分空間の基底ベクトルを求める。

3.2 局所部分空間集合による類似度の定義

入力パターン分布を近似した部分空間 A とクラス c のパターン分布を近似した局所部分空間の集合 B^c, B^c に含まれる k 個の局所部分空間を $B_i^c (i = 1, \dots, k)$ とすると、 A と B_i^c の類似度 $Ang(A, B_i^c)$ は式 (1) から求まる。

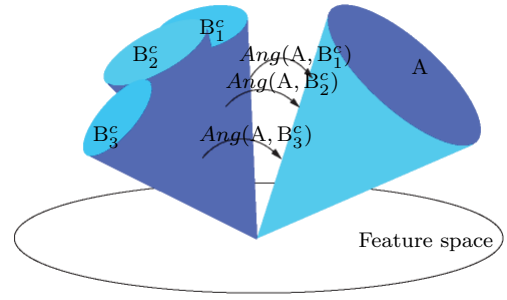


図 4 提案法の模式図

Fig.4 Concept of the proposed method .

この部分的な類似度 $Ang(A, B_i^c)$ を用いて、入力部分空間 A とクラス c の局所部分空間集合との類似度 $Sim(A, B^c)$ は式 (5) のように定義できる。

$$Sim(A, B^c) = \frac{1}{k} \sum_{i=1}^k Ang(A, B_i^c) \quad (5)$$

入力パターン分布 A と全クラスの局所部分空間集合の類似度を算出し、最終的に入力パターン分布 A は最も高い類似度を示したクラスに識別される。

3.3 アンサンブル学習

パターン分布を分割する分割数 k と各局所部分空間の次元、クラスタリングの初期配置などのパラメータを事前に最適決定することが困難である。そこで本論文では図 2 に示すように分割数 k と局所部分空間の次元 j を変えながら、それぞれの場合を弱識別器としてとらえて、これらを統合するアンサンブル学習を用いることで性能向上を図る。

分割数 k と局所部分空間の次元 j を変えながら類似度 $Sim(A, B^c)$ を算出し、それらの総和を次式のように定義し、最終的な類似度とする。

$$Sim_{total}(A, B^c) = \sum_j \sum_k Sim(A, B^c) \quad (6)$$

3.4 局所部分空間の補正

局所部分空間は該当クラスのパターンを最適に表現できても、クラス間の関係は考慮されていない。そこで、各局所部分空間にクラス間分離を考慮した変換を行うことで性能向上を試みる。クラス間分離を逐次的に補正する手法として、平均学習部分空間法 [3] が知られており、文字認識などにおいて有効性が確かめられている。

平均学習部分空間法では識別に先立ち、学習に使用した特徴ベクトルで識別を行い、その識別結果に基づいて各クラス部分空間の関係を補正する。補正は次のような規則で行う。各部分空間について誤識別を起こした場合を (1) ベクトルと部分空間が同クラスであるにも関わらず識別結果が他クラスとなった場

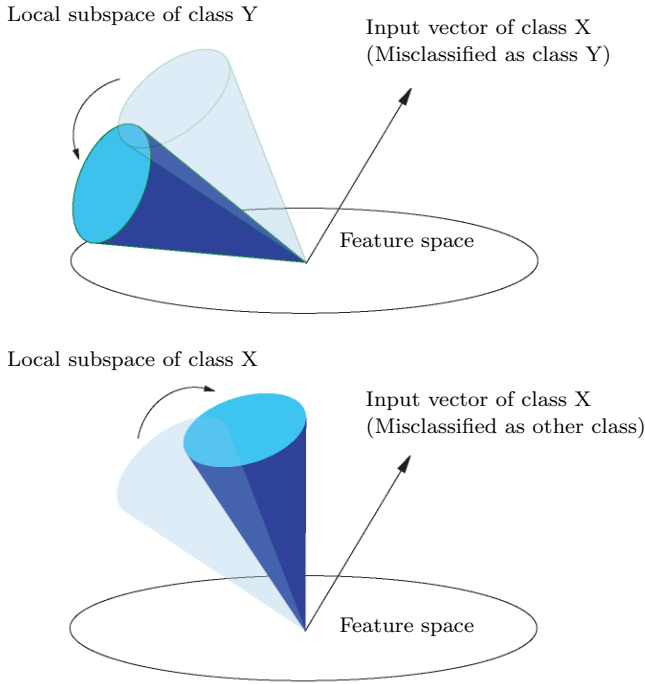


図 5 局所部分空間の補正 .

Fig. 5 Modification of local subspace.

合, (2) 逆にベクトルと部分空間が異なるクラスであるにも関わらず識別結果が自クラスとなった場合に分ける. 該当する特徴ベクトルについて, どの場合で誤識別したかをマーキングする. すべてのベクトルについて識別が終了した後, 各部分空間について, その部分空間が誤識別した特徴ベクトルのマーキングに応じて, 図 5 に示すように (1) の場合は該当部分空間をベクトルに近づけるように回転し, (2) の場合は該当部分空間をベクトルから離すように回転する. 提案手法では平均学習部分空間法と同様に誤識別を起こした局所部分空間を補正する.

テスト入力ベクトル x と局所部分空間 B_i^c の部分空間法による識別が誤ったとき B_i^c を補正する. $S_{B_i^c}(k)$ は局所部分空間 B_i^c の共分散行列であり, k 回目の補正は下式のように行われる.

$$S_{B_i^c}(k+1) = S_{B_i^c}(k) + \alpha \sum_{x_l \in C_{B_i^c}} x_l x_l^T - \beta \sum_{x_l \in D_{B_i^c}} x_l x_l^T \quad (7)$$

$S_{B_i^c}(0)$ は補正前の B_i^c の共分散行列である. ベクトルの集合 $C_{B_i^c}$ にはクラス c であるにも関わらず他のクラスに識別されたベクトルが含まれ, ベクトルの集合 $D_{B_i^c}$ には他クラスであるにも関わらずクラス c に識別されたベクトルが含まれる. また, α と β は補正の度合いを調整するパラメータであり, 予備実験に定める.

3.5 識別の流れ

提案手法は図 6 に示すように大きく学習フェーズと識別フェーズに分かれる.

まず入力画像をラスタ操作してそのまま特徴ベクトルと見

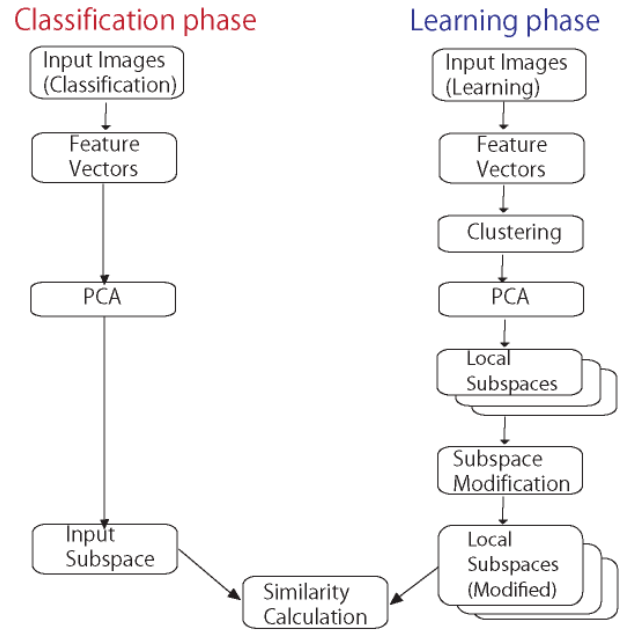


図 6 識別の流れ

Fig. 6 Flow chart of classification process.

なす. この過程は両フェーズに共通である. 学習フェーズでは分割数を多段階に変化させて, 入力される各クラスの特徴ベクトルを k-means 法によりクラスタリングする. 各クラスに PCA を適用し, 局所部分空間で近似する. 次に学習に使用した特徴ベクトルをテスト入力として各局所部分空間と部分空間法を行い, 誤った局所部分空間を補正する.

識別フェーズでは入力される特徴ベクトルを主成分分析して部分空間とし, 次元を変えながら各クラスの局所部分空間集合と式 (7) で定義した類似度を算出し, 識別結果とする.

4. 識別実験

4.1 比較実験

従来手法である MSM, 核非線形相互部分空間法 (KMSM) および提案手法の性能と計算量を評価する. さらに, 提案手法で採用したアンサンブル学習および局所部分空間補正の有効性を確かめる. 評価には 41 視点から撮影した物体からなる公開データセット ETH 80 Image Set [11] を用いる. このデータセットには図 7 に示すように, 8 クラス, 各クラスについて 10 個体の物体の背景を除去した画像が含まれる. 各物体については 41 視点から撮影された画像があり, 図 8 にその一例を示す. 画像はグレースケールの 320×320 ピクセルであるが, 評価実験ではこれを 16×16 ピクセルに縮小し, ラスタ操作して 256 次元の特徴ベクトルとする.

評価実験を次のように行った. 各クラスからランダムに 5 物体を選び, そのうち学習データとして 4 物体の画像を用いて, テスト用として残り 1 物体の画像を用いる. テスト用とした物体の画像から連続した 10 フレーム (視点) を入力とし, 1 フ

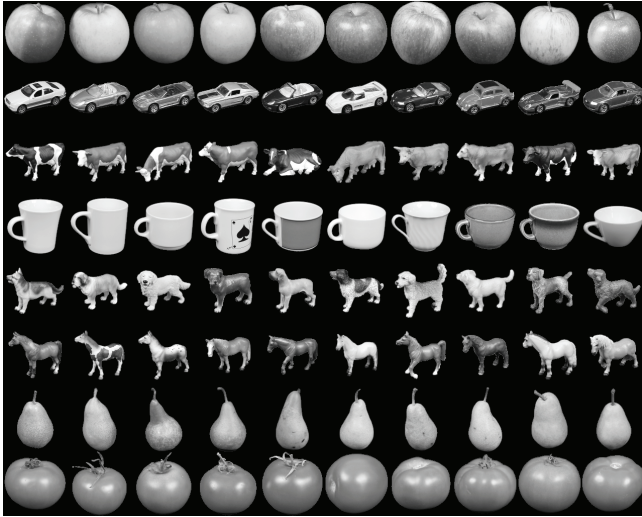


図 7 The ETH-80 Image Set.
Fig. 7 The ETH-80 Image Set.

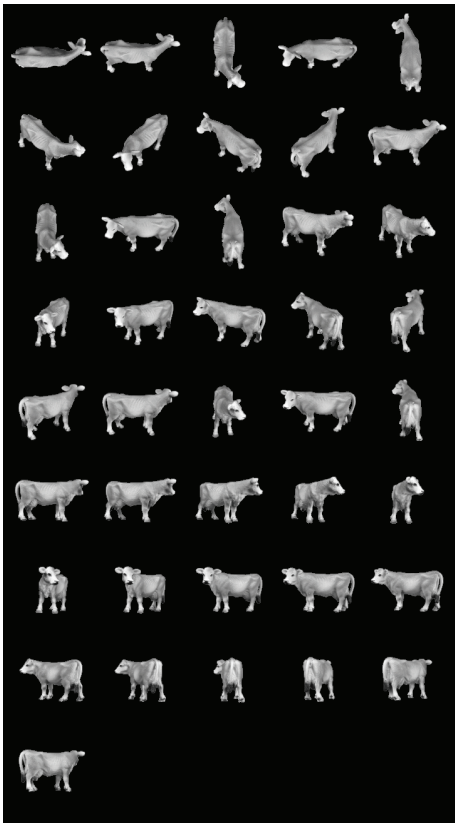


図 8 データセットに含まれる画像のサンプル．対象物体を 41 視点から撮影した画像である．

Fig. 8 Sample of images contained in the data set. These are images from 41 viewpoints of an object contained in the data set.

フレームずつずらしながら 41 回の識別実験を行う．8 クラス \times 41 回 = 328 回の実験を行った後，学習とテスト用のデータを入れ替えながら識別実験を $328 \times 5 = 1640$ 回繰り返す．

提案法の分割数は 2 から 5 とし，局所部分空間の補正に使用するパラメータ α と β はそれぞれ予備実験から 3 と 1 とした．MSM は入力，辞書ともに予備実験より 3 次元に固定する．提

表 1 各手法の識別結果

Table 1 Result of each method.

手法	識別率 (%)	分離度	EER(%)
MSM	69.5	0.34	20
KMSM	87.2	0.41	15
提案法 補正前	<u>86.5</u>	<u>0.44</u>	<u>14</u>
提案法 補正後	<u>91.0</u>	<u>0.51</u>	<u>10</u>

表 2 各手法の 1 入力 (10 フレーム) あたりの計算時間

Table 2 Calculation time per input (10 images) of each method.

手法	計算時間 (秒)
MSM	0.1
提案法	0.4
KMSM	3.1

表 3 各弱識別器と提案法の性能比較

Table 3 Performance comparison of weak classifiers and the proposed method.

分割数 k	次元	識別率 (%)	分離度	EER(%)
2	dim 1	69.5	0.37	15
2	dim 2	75.6	0.39	12
2	dim 3	70.6	0.43	14
3	dim 1	68.3	0.37	18
3	dim 2	72.1	0.37	15
3	dim 3	67.5	0.41	14
4	dim 1	74.6	0.40	15
4	dim 2	73.1	0.39	15
4	dim 3	76.2	0.43	13
5	dim 1	72.1	0.39	15
5	dim 2	74.6	0.40	13
5	dim 3	70.8	0.40	14
2 ~ 5	dim1 ~ dim3	<u>86.5</u>	<u>0.44</u>	<u>14</u>

案法は入力と辞書の次元は寄与率によって変動する．アンサンブル学習に使用した組み合わせは累積寄与率 98% を達成する次元を dim1, dim1 より 1 次元減らした次元を dim2, dim2 より 1 次元減らした次元を dim3 とした．

MSM, 提案手法, KMSM の識別結果は表 1 に示すとおりとなった．提案法 補正後の実験結果はすべての弱識別器を補正した場合のものである．分離度は 1 に規化されており，値が高いほど高い識別性能を示す．EER(Equal Error Rate) は FAR(False Accept Rate) 曲線と FRR(False Reject Rate) 曲線の交点であり，低いほど高い性能を示す指標である．

MSM, 提案手法, KMSM の識別フェーズにおける計算時間が表 2 である．

アンサンブル学習の効果を確かめるため表 3 に補正前の分割数 k や次元を固定した各弱識別器の識別結果とそれらをアンサンブル学習した補正前の提案法の識別結果を分割数 2~5, 次元 dim1~3 として示した．

4.2 考察

表 1 に見られるように，提案手法は従来の MSM より高い識別性能を示しており，17% の識別率向上を示している．非線形

手法である KMSM と同等の性能が得られる．また，提案法の局所部分空間補正の適用前と適用後の結果を比較すると，適用後がすべての指標において高い性能を示しており，局所部分空間の補正は有効であるといえる．

表 2 では MSM，提案法と KMSM の計算量を比較している．提案手法は MSM より計算量が増えているものの，非線形手法である KMSM の 13% 程度で識別を行うことができる．また，学習データの増加に対しては提案法と MSM がともにオーダー n の計算量増加であり，KMSM はオーダー n^2 である．したがって学習データの増加に伴って提案法は KMSM と比べて計算量の観点でさらに有利になると考えられる．以上の結果によって，提案法が非線形手法より少ない計算量で同程度の識別性能を示していると言える．

アンサンブル学習に採用した各弱識別器の識別結果が表 3 である．いずれの場合もアンサンブル学習を行った場合より低い性能であり，アンサンブル学習が識別性能の向上に寄与していることがわかる．

5. む す び

本論文では局所部分空間集合によって非線形パターン分布を近似し，アンサンブル学習を適用した 3 次元物体認識手法を提案した．各クラスのパターン分布を分割して主成分分析を適用し，それぞれ局所部分空間とする．さらに様々な条件で生成した局所部分空間集合を用いた各識別器を弱識別器と捉えアンサンブル学習を用いることで識別性能を向上させる．比較実験を通して提案法が従来手法より性能向上し，非線形手法と同程度の識別性能がより少ない計算量で得られた．また，局所部分空間の補正によって識別性能が向上することを確かめた．

今後の課題として，識別対象に応じた特徴抽出を採用し，提案手法と組み合わせることで更に識別性能の向上を図ることが挙げられる．

文 献

- [1] O. Yamaguchi, K. Fukui and K. Maeda: Face recognition using temporal image sequence, Proc. IEEE Third International Conference on Automatic Face and Gesture Recognition, pp.318-323, 1998.
- [2] H. Sakano, N. Mukawa: Kernel mutual subspace method for robust facial image recognition, Proc. Fourth International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies, Vol.1, pp.245-248, 2000.
- [3] J. Laaksonen, E. Oja: Subspace dimension selection and averaged learning subspace method in handwritten digit classification, Proceedings of the 1996 International Conference on Artificial Neural Networks, p.227-232, 1996.
- [4] L. Wolf, A. Shashua: Learning over sets using kernel principal angles, Journal of Machine Learning Research, 4(10), pp.913-931, 2003.
- [5] Y. Freund: Boosting a weak learning algorithm by majority, Proc. Third Annual Workshop on Computational Learning Theory, 1990.
- [6] Y. Freund, R. E. Schapire: A decision-theoretic generalization of on-line learning and an application to boosting, Journal of Computer and System Sciences, No.55, 1997.
- [7] 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", 電子情報通信学会論文誌 (D), Vol.J68-D, No.3,

- pp.345-352, 1985.
- [8] E. Oja: Subspace methods of pattern recognition, Research Studies Press, 1983.
 - [9] 黒沢 由明: 球面ガウス分布から導出される部分空間法, Trans. IEICE, Vol.J81-D-2, No.6, pp.1205-1212, 1998.
 - [10] 杉山 善明, 有木 康雄: 多重部分空間法に基づくテレビスポーツニュース映像の自動分類, Trans. IEICE, Vol.J81-D-2, No.9, pp.2112-2119, 1998.
 - [11] B. Leibe and B. Schiele: Analyzing appearance and contour based methods for object categorization, Proc. IEEE Conference on Computer Vision and Pattern Recognition, Vol.2, pp.409-415, 2003.