アンサンブル学習を用いた多重制約相互部分空間法による顔認識

西山 正志 山口 修 福井 和広

† (株) 東芝 研究開発センター 〒 212-8582 川崎市幸区小向東芝町 1

E-mail: †masashi.nishiyama@toshiba.co.jp

あらまし 多重制約相互部分空間法は,複数の制約部分空間へ入力部分空間と辞書部分空間を射影することで,識別に有効な特徴抽出を行い,高い識別性能を実現している.従来の多重制約相互部分空間法では,複数の制約部分空間の学習について十分な検討がなされておらず,効率的な学習方法が必要とされていた.そこで,本稿では,複数の識別器を効率よく学習し,その組み合わせで安定な認識系を構成できるアンサンブル学習の考え方に着目し,複数の制約部分空間を学習する方法について提案する.提案手法の有効性を,照明変動が生じる環境で撮影された顔画像を用いた認識実験で確認した.

キーワード 顔認識,アンサンブル学習,部分空間法,多重制約相互部分空間法

Face Recognition using Multiple Constrained Mutual Subspace Method with Ensemble Learning

Masashi NISHIYAMA[†], Osamu YAMAGUCHI[†], and Kazuhiro FUKUI[†]

† TOSHIBA Corporate Research and Development Center 1, Komukai-toshiba-cho, Saiwai-ku, Kawasaki, 212-8582, JAPAN

E-mail: †masashi.nishiyama@toshiba.co.jp

Abstract Multiple Constrained Mutual Subspace Method (MCMSM) can effectively extract features for recognition by projecting input and reference subspaces to multiple constraint subspaces. In previous paper, learning method of multiple constraint subspaces was not adequately discussed. In this paper, we propose a new method of learning multiple constraint subspaces in a framework of the ensemble learning. Through experimental results, we show the effectiveness of this method for face recognition.

Key words Face Recognition, Ensemble Learning, Multiple Constrained Mutual Subspace Method

1. はじめに

顔のパターンの変動は,照明条件,顔向き,表情,眼鏡などの装飾品,髪型などに依存し極めて多様であり,誤認識を引き起こす要因となる.実環境の中で安定な顔認識を行うためには,変動に影響されなく識別に有効な特徴をパターンから抽出する必要がある.識別に有効な特徴抽出を行うことができる認識手法の一つとして,我々は制約相互部分空間法 [1] を提案している.制約相互部分空間法では,入力を単一のベクトルではなく複数のベクトルから求めた部分空間とする.入力部分空間と辞書部分空間を,特徴抽出を行う制約部分空間へ直交射影し,部分空間同士のなす角度を類似度として求める(図 1).制約部分空間は,顔向き,表情,照明など識別に不用と考えられる個人の変動をできるだけ含まず,識別に有効と考えられる人物間における変動の差分を強調する空間である [1], [2], [11].

実環境で不特定多数の人物を精度良く識別できる制約部分空間は,様々な人物の様々な変動を含む学習パターンを用いて学習される.制約相互部分空間法の識別精度は,制約部分空間の学習パターンが増加すればするほど様々な人物や様々な照明条件の変化に対して全体的に向上するが,個々の人物や照明条件ごとに分けて評価すると逆に低下する場合がみられる.これは,特徴空間上で学習パターンが非線形性をもって多様に分布しているため,単一の制約部分空間のみでは,全ての人物について変動の差分を表すことは難しいからであると考える.そこで我々は,学習パターンを複数のクラスタに分割し,それぞれのクラスタに属するパターンから学習した複数の制約部分空間を用いて認識を行う多重制約相互部分空間法[3]を提案している.多重制約相互部分空間法は,入力部分空間と辞書部分空間を複数の制約部分空間上へ射影し,それぞれ類似度を求め結合することで類似度を決定する(図2).これにより,従来の制約

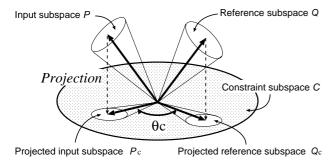
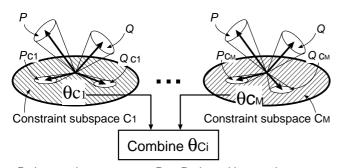


図 1 制約相互部分空間法の概念図



P: Input subspace Pci: Projected input subspace Q: Reference subspace Qci: Projected reference subspace

図 2 多重制約相互部分空間法の概念図

相互部分空間法に比べ, さらに安定な特徴抽出ができ, 識別精度を向上させることができる.

しかし,文献[3]では,複数の制約部分空間をどのように学習するかについては十分な検討がなされていなかった.多重制約相互部分空間法の識別精度をさらに高めるためには,複数の制約部分空間を効率的に学習する方法が必要である.学習パターンをうまく利用することで複数の識別器を効率よく学習し,その組み合わせで安定な認識系が構成できるアンサンブル学習[4],[5]が知られている.アンサンブル学習による認識系は,線形の識別器では誤識別が生じるような非線形なパターン分布の境界面を,複数の線形の識別器の組み合わせで近似的に表すことができるといわれている.アンサンブル学習の考え方を用いることで,学習パターンが非線形性をもって多様に分布する場合にも,識別に有効な複数の制約部分空間が学習できると考えられる.

本稿では、多重制約相互部分空間法において、アンサンブル学習の考え方を利用し、複数の制約部分空間を学習する方法を提案する。これにより、学習パターンを複数のクラスタに分割して、それぞれのクラスタに属する学習パターンから制約部分空間を学習する場合[3]に比べて識別に有効な制約部分空間を学習することができ、さらなる識別精度の向上が期待できる。

まず、2. において、多重制約相互部分空間法を顔認識に適用する方法について概要を述べる、次に、3. において、複数の制約部分空間をアンサンブル学習する方法について提案する、最後に、4. で、照明変動が生じる環境で撮影された顔画像を用いた認識実験により提案手法の有効性を確認する.

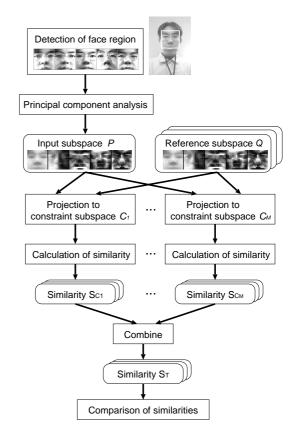


図 3 多重制約相互部分空間法による顔認識の流れ

2. 多重制約相互部分空間法による顔認識

2.1 顔認識の流れ

多重制約相互部分空間法を用いた顔画像認識の流れを図3に示す.まず,入力された画像から瞳と鼻孔を基準として顔領域のパターンを切り出し入力ベクトルに変換する[9].時系列に獲得された複数の入力ベクトルから,主成分分析により,入力部分空間を求める.次に,入力部分空間と予め登録されている辞書部分空間を複数の制約部分空間に射影し,それぞれの制約部分空間上で類似度を算出する.得られた複数の類似度を結合し入力部分空間と辞書部分空間の類似度を決定する.この類似度が最大になる辞書部分空間に対応する人物が入力された人物と判定する.2.2節より,それぞれの処理について詳細を述べる.

2.2 制約部分空間への射影

入力部分空間 P と辞書部分空間 Q から識別に有効な特徴を抽出するために,図 2 のように M 個の制約部分空間 $C_i(i=1,\dots,M)$ を用いる.制約部分空間は,個人のパターンの変動を抑えつつ人物間の違いを強調する空間であり,詳細な学習方法については 3. で述べる.ここでは,入力部分空間を制約部分空間に射影する方法について述べる.

- (i) 入力部分空間を張る N 個の基底ベクトルを制約部分空間上へ射影し射影ベクトルを求める .
 - (ii) 各射影ベクトルの長さを正規化する.
- (iii) N 個の正規化ベクトルに対して Gram-Schmidt の直 交化を施す .
 - (iv) 直交化された N 個の正規化ベクトルが射影された入

力部分空間の基底ベクトルとなる.

この手続きを用意した制約部分空間の個数 M だけ繰りかえす . 辞書部分空間も同様の手続きで制約部分空間へ射影することができる .

2.3 部分空間同士の類似度の算出

制約部分空間 C_i に射影された入力部分空間を P_{C_i} とし、辞書部分空間を Q_{C_i} とする . P_{C_i} と Q_{C_i} 間の類似度 S_{C_i} は第 1 正準角と呼ばれる二つの部分空間がなす角度 θ_{C_i} により式 (1) で決定される [10] .

$$S_{C_i} = \cos^2 \theta_{C_i} \tag{1}$$

部分空間同士が完全に一致していれば $heta_{C_i}=0$ である $\cos^2 heta_{C_i}$ は,以下の行列 ${f X}$ の最大固有値となる.

$$\mathbf{X}\mathbf{a} = \lambda \mathbf{a} \tag{2}$$

$$\mathbf{X} = (x_{jk}), (j, k = 1 \sim N) \tag{3}$$

$$(x_{jk}) = \sum_{l=1}^{N} (\psi_j, \phi_l)(\phi_l, \psi_k)$$
 (4)

ここで, ψ_j , ϕ_k は部分空間 P_{C_i} , Q_{C_i} の j,k 番目の正規直交基底ベクトル,N は部分空間の次元数を表す.実際には,正準角は N 個求まるが,以下では第 1 正準角のみを取り扱う.

2.4 類似度の結合

複数の制約部分空間上で得られた類似度を結合するために,それぞれの類似度に対して重みを考慮する.全ての類似度に対して重みを一定とする場合,平均値,最大値,最小値,中央値を用いるといったことで結合できると考えられる.平均値で結合する場合,多重制約相互部分空間法の類似度 S_T は式 (5) で決定される.

$$S_T = \frac{1}{M} \sum_{i=1}^{M} S_{C_i} \tag{5}$$

ここで,M は射影する制約部分空間の個数, S_{C_i} は制約部分空間 C_i 上での入力部分空間 P_{C_i} と辞書部分空間 Q_{C_i} の類似度を表す.また,入力部分空間に合わせて逐次的に重みを変更することもできる.入力部分空間の識別に有効と考えられる制約部分空間上での類似度に対して,重みを付けることで結合する [3] .

3. 複数の制約部分空間に対するアンサンブル 学習

3.1 単一の制約部分空間の学習

複数の制約部分空間の学習方法を述べる前に,単一の制約部分空間の学習方法について説明する.制約部分空間を学習するために,図 4 のように,L 人の人物について,それぞれ顔のパターン変動をまとめて主成分分析を適用し学習用部分空間を用意する.パターン変動には,照明条件,顔向き,表情,装飾品,髪型などによる変動が含まれる.次に,学習用部分空間同士をなるべく類似させないように制約部分空間を求める.単一

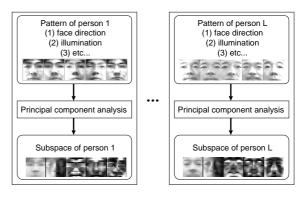


図 4 学習用部分空間の生成

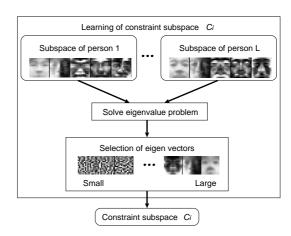


図 5 単一の制約部分空間の学習の流れ

の制約部分空間の学習の流れを図5に示す。制約部分空間は,L個の学習用部分空間が形成する和空間からこれらの主成分部分空間を取り除いた一般化差分部分空間を求めることで学習される[11]。学習用部分空間を張る基底ベクトルを各人物について N_B 本用いて求めた行列の固有値問題を解き,固有値が小さい順から選んだ N_C 個の基底ベクトルで張られる空間を一般化差分部分空間とする[11]。得られた一般化差分部分空間を制約部分空間として用いる。なお, N_B,N_C は実験的に定める。

3.2 並列的な学習

複数の制約部分空間を学習するために、アンサンブル学習の一つの手法である Bagging [7] を適用する方法について述べる.Bagging は単純なアルゴリズムであるが複数の識別器の組み合わせで頑健な認識系を構成できるといわれている.Bagging は学習パターンからランダムにサンプリングを繰りかえすことで異なる識別器を生成する.

学習用部分空間の個数を L 人分とすると以下の手続きで M 個の制約部分空間が学習できる .

- (i) L 個の学習用部分空間からランダムに R 個の学習用部分空間を選択する .
- (ii) R 個の学習用部分空間の基底ベクトルを用いて 3.1 節 の手続きで 1 個の制約部分空間を生成する .
- (iii) M 個の制約部分空間が生成されるまで (i) に戻る . 図 6 に ,複数の制約部分空間を並列的に学習する流れを示す . なお ,R は実験的に定める .

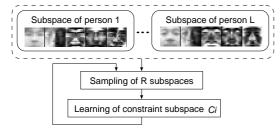


図 6 並列的な学習の流れ

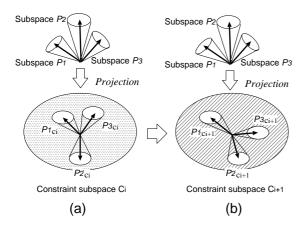


図 7 異なる制約部分空間での分布変化の概念図

3.3 逐次的な学習

3.3.1 学習に用いる重みの決め方

複数の識別器を逐次的に学習する手法の一つである AdaBoost [8] の考え方を,複数の制約部分空間の学習に適用する.AdaBoost は,直前に学習した識別器で誤った学習パターンの重みを増して新たな識別器の学習を行う.学習パターンに対する重みを逐次的に更新することで,パターン数が少なくとも様々な識別器が学習でき,安定な認識系が構成できるといわれている.

逐次的に複数の制約部分空間を学習する場合,何に対してどのように重みを与えるかを考える必要がある.ここで,制約部分空間の学習に用いた学習用部分空間を識別する場合を考える.制約部分空間上で学習用部分空間同士のなす角度が小さい場合,これらは類似しており識別誤りが発生する可能性が増える.よって,制約部分空間 C_i 上で他の学習用部分空間となす角度が小さい学習用部分空間の重みを増やすことで,次の制約部分空間 C_{i+1} を生成する.これにより,図 7(a) のように,制約部分空間 C_i 上では角度の小さかった学習用部分空間が,(b) のように,制約部分空間 C_i 上では角度の小さかった学習用部分空間が,(c) のように,制約部分空間で識別誤りが発生しやすかった部分空間に対しては識別性能を向上させることができると考える.

3.3.2 逐次的な学習のアルゴリズム

まず,L 人分の学習用部分空間 $P_j(j=1\dots L)$ を 3.1 節と同様の方法で用意する.制約部分空間 C_i を学習する際の P_j に対する重みを $W_i(j)$ とすると,以下の手続きで M 個の制約部分空間が学習できる.

- (i) 重み $W_1(j)$ に初期値を与える.
- (ii) 学習用部分空間 P_i の基底ベクトルに対して重み $W_i(j)$

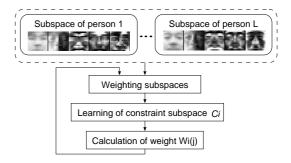


図 8 逐次的な学習の流れ

表 1 実験条件

評価用の画像の撮影人数 25 人制約部分空間の学習用の画像の撮影人数 25 人照明条件(1人につき) 10 条件各照明条件での試行回数(1人につき) 7 回1 回の試行に用いる入力パターンの枚数 10 枚



図 9 撮影画像の例

をかける .

- (iii) 重み付けされた L 個の学習用部分空間の基底ベクトルを用いて 3.1 節の手続きで 1 個の制約部分空間 C_i を学習する ...
- $({
 m iv})$ 制約部分空間 C_i を用いて以下の式 (6) によりあらたな重み $W_{i+1}(j)$ を決定する .
- (v) M 個の制約部分空間が生成されるまで (ii) に戻る . 図 8 に制約部分空間の逐次的な学習の流れを示す . あらたな重み $W_{i+1}(j)$ は , 以下の式で求めることができる .

$$W_{i+1}(j) = \frac{S'_j}{\sum_{j=1}^L S'_j} \tag{6}$$

$$S'_{j} = \sum_{k=1, j \neq k}^{L} \cos^{2} \theta_{C_{i(jk)}} \tag{7}$$

ここで, $\theta_{C_{i(jk)}}$ は,学習用部分空間 P_j,P_k を制約部分空間 C_i へ射影した後のなす角度を表す.部分空間がなす角度は 2.3 節の方法で求まる.また,式 (7) にしきい値 T 以上の $\cos^2\theta_{C_{i(jk)}}$ から総和をとるという条件を設けることで,類似性の高い学習部分空間のみから学習を行うことができる.なお,T は実験的に決定する.

4. 実 験

本稿で述べた手法の有効性を確認するために顔画像を用いた



図 10 獲得した顔パターンの例

表 2 実験結果

	手法	$\mathrm{ER}(\%)$	EER(%)
(a)	最近傍決定則	4.6	23.9
(b)	部分空間法	4.6	12.9
(c)	制約相互部分空間法	4.6	5.0
(d)	多重制約相互部分空間法 (分割)	3.3	7.5
(e)	多重制約相互部分空間法 (並列)	1.8	4.0
(f)	多重制約相互部分空間法 (逐次)	1.4	3.9

認識実験を行った. 図9のように,照明条件を変化させて取得 された画像を用い,実験の条件は表1の設定とした.1人の人 物につき 10 種類の照明条件 $I_i(i=1,\ldots,10)$ の画像を取得し た.50人について撮影を行い,25人の顔のパターンを制約部 分空間の学習用とし,残りの25人を評価用のパターンとした. それぞれの人について,照明条件毎に辞書パターンと入力パ ターンを用意し,入力部分空間と辞書部分空間を生成した.1 個の入力部分空間は 10 枚のパターンから生成し,1 個の辞書部 分空間は 70 枚のパターンから生成した. 入力部分空間と辞書 部分空間の生成に用いたパターンは異なる. 入力部分空間と辞 書部分空間の照明条件の組み合わせは 100 種類である. 各照明 条件で1人につき7回の試行を行っており,25人で合計17500 回の試行を行った.顔のパターンは,画像中から[9]に述べら れている方法で,図10のような15×15pixelsの大きさで抽 出した.パターンに対して,ヒストグラム平坦化と縦方向の微 分を行った後に 210 次元のベクトルに変換した. 得られたベク トルを用いて,入力部分空間と辞書部分空間を主成分分析によ り生成した.主成分分析を適用する際に,平均ベクトルを引か ずに求めた自己相関行列[6]を用いた.

評価は,(a) 最近傍決定則,(b) 部分空間法,(c) 制約相互部分空間法,(d) 従来の学習方法[3] を用いた多重制約相互部分空間法,(e) 並列的な学習を用いた多重制約相互部分空間法,(f) 逐次的な学習を用いた多重制約相互部分空間法について行った.それぞれの手法のパラメータは以下のように設定した.

(a) 最近傍決定則

辞書パターンを全て記憶しておき,入力パターンと最も類似する辞書パターンとの距離を類似度として識別を行った.(c)-(f)と条件を揃えるため 1 回の試行で 10 枚の入力パターンを用いた.

(b) 部分空間法

入力パターンと辞書部分空間との射影長を類似度として識別を行った.辞書部分空間の次元数は 40 とした.(c)-(f) と条件を揃えるため 1 回の試行で 10 枚の入力パターンを用いた.そ

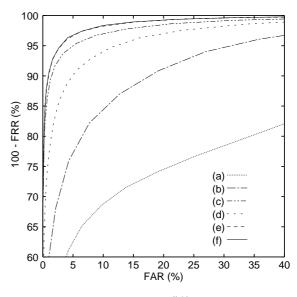


図 11 ROC 曲線

れぞれで類似度を算出し平均値を求めた.

(c) 制約相互部分空間法

入力部分空間と辞書部分空間の次元数 N はともに 7 とした.制約部分空間は,学習用部分空間の L=25 人分全てから 1 個を学習し,その次元数 N_C は 170 とした.学習用部分空間の基底ベクトルの本数 N_B は 30 とした.

(d) 多重制約相互部分空間法 (分割)

制約部分空間の学習用の 25 人のパターンを , それぞれの照明条件 I_1 から I_{10} に分割し合計 M=10 個の制約部分空間を学習した . 類似度の結合方法は式 (5) の平均値を用いた . N_B,N_C は (c) と同じ値を用いた .

(e) 多重制約相互部分空間法 (並列)

L=25 人分の学習用部分空間からランダムに R=8 個を選択し,合計 M=10 個の制約部分空間を学習した.類似度の結合方法は式 (5) の平均値を用いた.この学習方法は R 個の学習用部分空間の選択にランダム性があるために,一連の手続きを10 回行い,そのつどに後述する ER と EER を算出し平均値を求めた. N_B,N_C は (c) と同じ値を用いた.

(f) 多重制約相互部分空間法 (逐次)

L=25 人分の学習用部分空間に対して逐次的に重みを変化させながら,合計 M=10 個の制約部分空間を学習した.重みの初期値は $W_1(j)=1/25(j=1\dots 25)$ とし,式 (7) にしきい値 T 以上の類似度から総和をとる条件を用いた.しきい値は $T=3.5\sigma_i$ とし, σ_i は制約部分空間 C_i を用いての算出した学習用部分空間同士の類似度の標準偏差とした.類似度の結合方法は式 (5) の平均値を用いた. N_B,N_C は (c) と同じ値を用いた.

認識性能の評価には以下の2つの基準を用いた.

(i) エラー率 (ER:Error Rate)

入力部分空間と辞書部分空間の類似度が,本人同士の値よりも他人との値の方が高くなる割合.

(ii) 等価エラー率 (EER:Equal Error Rate)

FAR(他人受理誤り率)と FRR(本人排除誤り率) が等しい時の

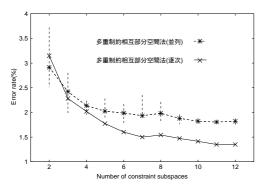


図 12 制約部分空間の個数を増加させた時の ER

割合.顔認識システムの場合,登録されていない未知の人物に対応する必要がある. EER が低いとそのシステムは信頼性が高いといえる. FAR は以下の式で求まる.

$$FAR = rac{他人の類似度がしきい値より大の試行数}{(登録人数×試行数) - 本人の試行数} (8)$$

一方, FRR は以下の式で求まる.

それぞれの基準についての評価結果を表 2 に示す.制約部分空間により特徴抽出を行った (c),(d),(e),(f) は,行わなかった (a),(b) に比べて,大きく EER が改善されている.中でも,提案手法の結果 (e) と (f) は,EER がさらに改善されており,ER も改善されいてた.図 11 の ROC 曲線でも認識性能が改善されていることが確認された.また,(c) の EER に比べて (d) の EER が悪くなっており,従来の分割による学習では条件により認識性能が低下した.これは,学習パターンを照明条件で分割したため,学習用部分空間には照明によるパターン変動が含まれておらず,個人のパターン変動を抑える制約部分空間が学習できなかったためと考えられる.

次に,制約部分空間の個数 M を 2 個から 12 個まで増加させたときの ER の変化を図 12 に,EER の変化を図 13 に示す.並列的な学習では前の実験と同様に ER や EER を 10 回算出しその平均値を点線でつないでおり,点線に付属する誤差棒は ER や EER の最大値と最小値を結んだものである.制約部分空間の個数が少ない場合,並列的な学習では認識性能に大きなばらつきが生じることが分かった.並列的な学習と逐次的な学習のどちらでも,個数が増加するにつれて ER と EER が改善され認識性能が向上する結果が得られた.

さらに制約部分空間の個数が増加した場合について,並列的な学習のみであるが実験を行った. L,R,N_B,N_C は (\mathbf{e}) と同じ値を用いた.M=100 のとき ER は 1.7% で EER は 3.7% となった.表 $2(\mathbf{e})$ の M=10 のときと比べて若干ではあるが認識性能が改善されていた.制約部分空間の個数を極端に増加させることで大幅な認識性能の改善は得られなかったが,逆に性能が悪化することもなかった.

以上の結果より,本稿で提案した複数の制約部分空間をアンサンブル学習し多重制約相互部分空間法に用いることの有効性を示した.

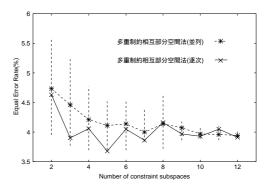


図 13 制約部分空間の個数を増加させた時の EER

5. おわりに

本稿では、多重制約相互部分空間法において、学習パターンが非線形に分布する場合にも識別に有効な複数の制約部分空間を学習するために、アンサンブル学習の考え方を取り入れた方法を提案した、アンサンブル学習には並列的に学習するか逐次的に学習するかの二つの方法があり、これらの方法の考え方を複数の制約部分空間の学習に適用した、並列的な学習では、学習用部分空間のサンプリングを繰り返すことで複数の制約部分空間を学習した、一方、逐次的な学習では、直前に学習した制約部分空間で類似する学習用部分空間の重みを増して新たな制約部分空間の学習を行った、これらの学習を用いることで、多重制約相互部分空間法の認識性能が向上することを、照明変動が生じる環境で撮影された顔画像を用いた実験により確認した、

今後の課題として,制約部分空間の個数を少なくしても安定な認識が行えるような学習方法の検討や,認識性能が改善される原因の詳細な解析などが必要である.

文 献

- [1] 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画像認識 照明変動の影響を抑える制約相互部分空間の学習-, 信学論 D-II Vol. J82-D-II, No.4, 613-620, 1999
- [2] K.Fukui, O.Yamaguchi, Face Recognition Using Multiviewpoint Patterns for Robot Vision, 11th International Symposium of Robotics Research, 2003
- [3] 西山 正志, 山口 修, 福井 和広, 多重制約相互部分空間法による 顔画像認識, 信学技報 PRMU2003-145 pp.31-36, 2003
- [4] 麻生 英樹, 津田 宏治, 村田 昇, パターン認識と学習の統計学 新 しい概念と手法, 岩波書店, 2003
- [5] R.O.Duda, P.E.Hart and D.G.Stork, Pattern Classification, John Wiley & Sons, 2000
- [6] 石井 健一郎, 上田 修功, 前田 英作, 村瀬 洋, わかりやすいパターン認識、オーム社、1998.
- [7] L.Breiman, Bagging Predictors, Machine Learning, Vol.24, No.2, pp.123-140, 1996
- [8] Y.Freund, R.E.Schapire, A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, JCSS 55(1), pp.119-139, 1997
- [9] 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", 信学論 D-II Vol. J84-D-II, No.6, 1045-1052, 2001
- [10] 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, 345-352, 1985
- [11] 福井 和広、部分空間表現に基づく顔画像認識、信学技報 PRMU2002-206 pp.33-40, 2003