

## 研究テーマ 顔画像認識に関する研究

研究者	安本 護	財団法人ソフトピアジャパン	雇用研究員
	本郷仁志	財団法人ソフトピアジャパン	雇用研究員
	興水大和	中京大学	兼業研究員
	伊藤英則	名古屋工業大学	兼業研究員
	橋本恭貴	メディアドライブ株式会社	共同研究員
	鎌田一男	メディアドライブ株式会社	共同研究員

### フェーズ I

#### 1 研究の概要

個人適応型マンマシンインタフェースを実現するためには、ユーザが誰であるかを知る必要がある。そのため、入力されるユーザの顔画像を用いて個人認識を行う。個人認識を行う研究では、正面顔画像を用いたものが多く、主成分分析等を用いることで高い認識率が得られることが知られている。しかし、実環境での実装を考えた場合、ユーザを拘束しないように正面顔だけでなく、多様な方向を向いた多方向顔画像に対しても顔認識の信頼性を向上させることが重要である。

本研究では、多方向から取得した顔画像に対して顔認識を試みる。顔認識の手法として、アピアランススペースの特徴量である四方向面特徴と統計的判別器である線形判別分析を用いる。10人の被験者を対象とした実験では、1台のカメラで296°の顔向き範囲に対し、92.9%の認識率が得られた。

本研究は、マルチカメラを用いた「顔向き推定に関する研究」と密接に関係しており、その報告に顔画像認識に関しても記述されているため、本節では詳細を割愛する。

### フェーズ II

#### 1 研究の概要

フェーズ I では、四方向面特徴と線形判別分析を用いて 10 人の被験者に対する顔画像認識が可能であることを示した。顔向きに応じた顔認識手法は、識別率を低下させず、顔向きの識別範囲を拡張することが出来る有効な手法である。しかし、顔向きの推定結果に依存するため、顔向きの推定を誤ると人物識別に大きく影響する欠点がある。顔向きは顔形状の個人差により画一的に推定することは困難であり、顔向きは異なっても、見かけ上の顔が似ている人物同士も存在することが予想される。従って、階層的な手法において、絞り込み方法とその手段が重要になる。

これらの問題を解決するために、統計的手法により段階的にカテゴリを絞り込み、その中から最適に分類する判別空間を適宜作成しパターンを分類する適応的判別空間手法を提案する。具体的には、大まかな顔向き推定と顔認識で候補者を絞り込み、その候補者各人の顔向き推定判別空間を用いて、詳細な顔向きを推定する。その推定結果に基づいた学習用顔画像データから新たに顔識別空間を構築し、その判別空間により人物を識別する。つまり、大分類で候補者を選出し、その候補者らを最適に識別する判別空間を生成させることで識別能力を高める。実験では、150名、35方向（水平 $\pm 45^\circ$ 、垂直 $\pm 30^\circ$ ）の顔画像データに対して、新しく生成した判別空間を用いることで識別率を 97.6% から 98.8% に向上させることが出来た。

#### 2 研究の目標

フェーズ I で提案した手法は、有効性は確認できたが、被験者が 10 人であった。そのため、より多くの被験者を用いた個人識別実験を行うことで拡張性を検討する。また、フェーズ I では顔向

き推定で誤った場合には、個人識別に大きく影響する問題があった。そのため、最終結果を得るまでの階層化をより深くし、適応的に判別空間を作成することで、絞り込みと判別空間の最適化を行う提案手法の有効性を検証する。

### 3 実施内容

#### 3.1 四方向面特徴による顔画像認識

フェーズIでは、四方向面特徴と線形判別分析を用いた顔画像認識手法を提案している。本手法は、顔領域から抽出した四方向面特徴を用いて、線形判別分析により得た判別空間で距離が最小となるクラスを識別結果とするアピランススペースの手法である。そのため、学習データの集合を調整することにより識別カテゴリを容易に変更することができる。本手法は、顔画像認識以外にも、顔向き推定、手サイン認識などにも有効である。

##### 3.1.1 顔領域抽出

顔領域抽出には色情報を用いる。色情報により顔領域を抽出する手法は数多く提案されているが、個人差、照明の変化にロバストな肌色領域抽出は大きな課題である。本研究では、顔認識に着目するため、画像データは単純背景とした。そのため、肌色基準値手法を用いる。

肌色基準値手法は、CIE- $L^*u^*v^*$ 表色系の $u^*$ 、 $v^*$ 値を用い、入力画像より肌色の取り得る $u^*$ 、 $v^*$ 値の範囲内で最も画素数の多い $u^*$ 、 $v^*$ 値を肌色基準値とする。図1にCIE- $L^*u^*v^*$ 色空間での肌色基準値と顔画像における色分布の一例を示す。次に、入力画像の各画素の $u^*$ 、 $v^*$ 値と肌色基準値との $u^*v^*$ 平面上での距離を算出する。図2(b)に肌色基準値からの距離を濃淡値で表した画像を示す。次に $u^*v^*$ 平面上での距離ヒストグラムを作成し、判別分析法を用いて分割し閾値以下を肌色領域とする。図2(c)に本手法による抽出結果を示す。個人差や照明などの変化により肌色基準値が変動するが、肌色基準値からの色差を用いることで安定した肌色領域抽出が可能である。

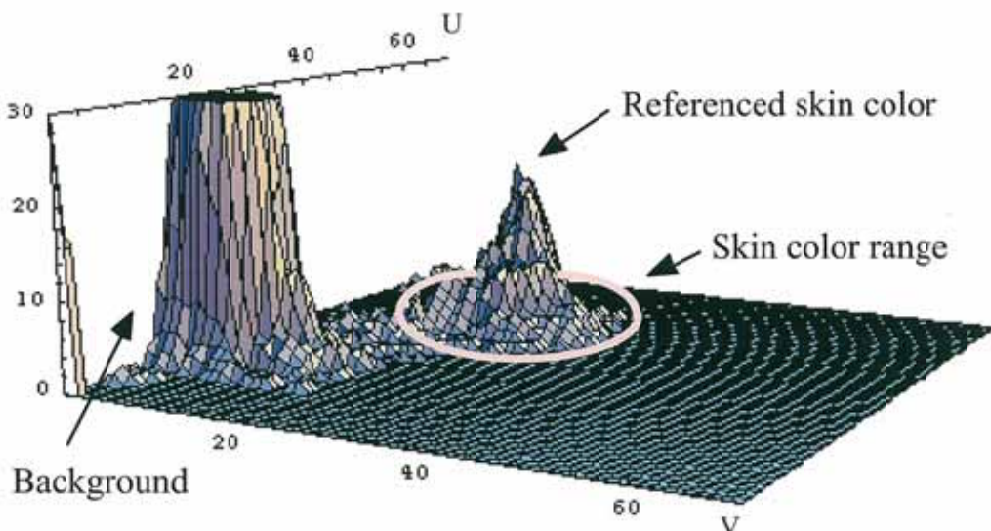


図1 肌色基準値と顔画像色の分布例



図2 肌色領域抽出結果

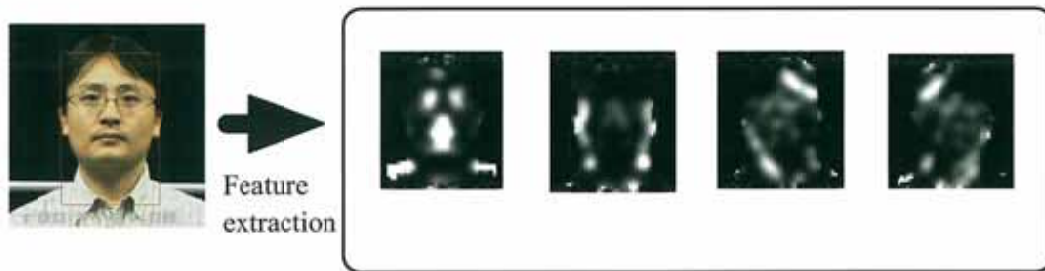


図3 四方向面特徴

### 3.1.2 四方向面特徴

四方向面特徴とは、各画素のエッジ勾配を四つの方向面に割り当てて表現したものである。本研究では、肌色基準値手法で抽出した顔領域の濃淡値から、Prewitt オペレータで求めた各画素のエッジ勾配を、エッジ方向に基づき、四方向（水平、垂直、右斜め、左斜め）に分解した四方向面画像をガウシアンフィルタに通して低解像度化した特徴量である。四方向面特徴は解像度を低下させてもエッジ方向情報を保持しているため、エッジの形状変動に対してロバストになる。さらに計算コストを削減することができ高速化が図れる。図3に低解像度化する前の四方向面画像の一例を示す。

本稿で用いる四方向面特徴は、肌色領域から四方向面画像を生成し、解像度を  $8 \times 8$  に正規化した  $256(8 \times 8 \times 4)$  次元の特徴量とする。

### 3.1.3 線形判別分析

線形判別分析は、クラス内分散を最小に、クラス間分散を最大にするような特徴空間を生成する写像行列を求める。つまり、学習パターンを識別するのに最適な判別空間を生成する。本手法では、変形にロバストな四方向面特徴を入力として、線形判別分析で得た写像行列により判別空間へ投影し、判別クラスの平均ベクトルとの距離が最小となるクラスを識別結果とする。

各クラスの顔画像を顔向き別に分類すれば、顔向きを推定する判別空間が生成され、人物別に分類することで人物を識別する判別空間が生成される。従って、生成した判別空間を変えるだけで、同一の特徴量から顔向き推定と顔認識を行うことが可能となる。

## 3.2 適応的判別空間手法

フェーズIで提案した階層的人物識別手法では、顔向きを推定した後、顔向きに応じて顔認識をおこなっている。この手法では、顔向きのバリエーションを限定することで認識率の向上が予想できる。しかし、階層的人物識別手法の問題点として、顔向きの推定を誤るとその後の顔認識に影響を与えることになる。顔向きは、顔形状の個人差により画一的に表現できないため、汎用的な辞書による詳細な顔向き推定は困難である。そのため、顔向きによる絞り込み段階で正解者（本人）を排除する危険がある。また、顔向きは異なるが、見かけ上の顔が似ている人物同士も存在することが予想される。従って、顔向き推定精度と顔向き限定顔認識が密接な関係となり実験的調整が必要となる。その他の問題点として、絞り込んだ人物の中から、詳細に人物を識別する新たな基準が必要となる。

そこで、適応的判別空間手法は、階層的人物識別手法を大分類と位置付け、段階的にカテゴリを絞り込み候補者を選出し、その中から最適に顔画像を識別する判別空間を適宜作成し人物を識別す

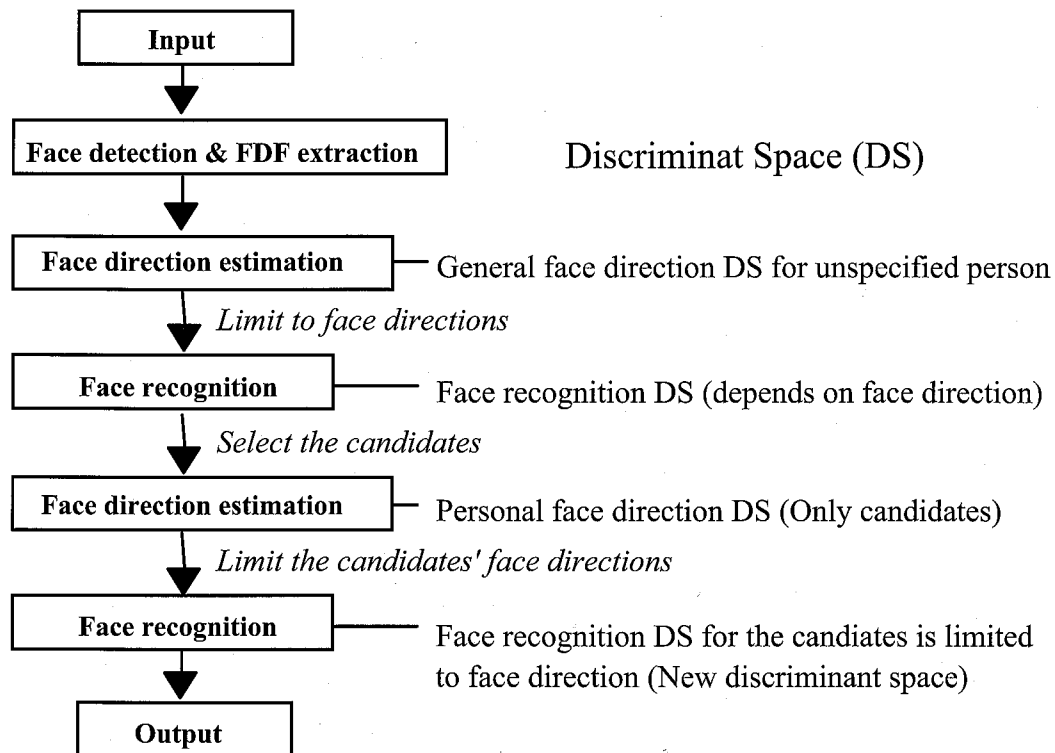


図4 適応的判別空間手法の流れ

る。

処理の流れを図4に示す。まず、大まかな顔向き推定と顔認識で候補者を絞り込む。この時、顔認識する判別空間を顔向き推定結果より拡張した範囲を学習する。これにより顔向き推定の誤りをカバーすることができる。次に、候補者の顔向きを候補者各人の顔向き推定判別空間を用いて、詳細な顔向きを推定する。その推定結果に基づいた学習用顔画像データから新たに顔識別空間を構築し、その判別空間により個人を識別する。従って、大分類で候補者を選出し、その候補者らを最適に識別する判別空間を生成させることにより識別能力を高める。その際に、個人別に顔向きを推定することで、個人差を打ち消すことができる。各候補者で推定された顔向きの学習用顔画像から、再度判別空間を生成することで、候補者間で最適に識別する判別空間を使用することができ、識別率の向上が期待できる。つまり、候補者とその顔向きを絞り込むことで、微妙な識別も可能となる識別空間を適宜作成することになる。

## 4 結果

本研究で提案する適応的判別空間手法の有効性を確認するため、他の四方向面特徴による顔画像認識手法と比較する。まず、多方向から取得された顔画像データを用いて、線形判別分析によって全方向の画像データを学習する認識手法を評価する。次に、階層的認識手法と適応的判別空間による認識手法を評価する。

### 4.1 顔画像データ

多方向の顔画像を収集するために、同期をとった任意のカメラ映像をメインPCからの操作で同時にデータ収集できる多視点画像収集システムを構築した。本システム構成を簡単に説明する。

シンクジェネレータで同期をとった16台の3CCDビデオカメラ(DXC-9000、SONY)で撮影した映像を、各々タイムコード・インポーズ(VG-50、HORITA)を通し、ビデオキャプチャボード(カノーパス社製、Power Capture Pro)を搭載したPC(Pen-III 600 MHz、Windows2000、以後、カメラPCと呼ぶ)に入力する。各カメラPCは、ビデオレートのカラー画像をモーションJPEG圧縮でハードディスクに格納する。このとき、VITCタイムコードを画像データに記録する。つまり、





図 5 顔画像データ例

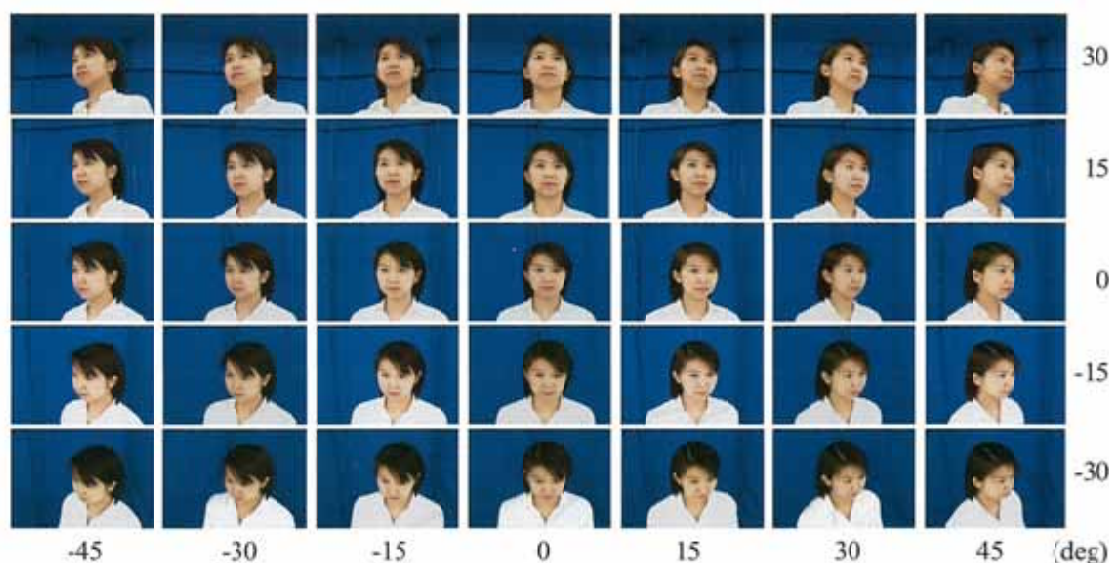


図 6 多方向画像データ例

画像上にスーパーインポーズされたタイムコード情報を読み取ることで、収集した画像データの同期を確認することができる。

顔画像データは、14 台の CCD カメラを被験者に対して垂直・水平ともに 15 度間隔で 7 台 2 列に設置し、多方向から収集した。画像サイズは、640×480 ピクセル、24 ビットカラー（RGB 各 8 ビット）である。被験者の背景はブルーバックとした。照明は、被験者の顔および周辺が一樣な明るさになるように調整した。被験者は、15 歳から 64 歳迄を年代別・性別を均等な人数に振分け、計 300 名を収集した。被験者の画像例を図 5 に示す。なお、被験者に対する制限は、サングラスや色付きレンズの眼鏡を着用しないことのみで、髪型、服装、化粧などは自由とした。

実験では、垂直±30 度、水平方向±45 度の範囲を対象として、肌色基準値手法を用いて顔領域が正しく抽出できた 150 名の顔画像を用いた。2 回収集した内の 1 回目を学習用、2 回目をテスト用とした。各方向に対して 1 秒間撮影した画像から 10 枚サンプリングし、総画像枚数は 10,500 枚（150 名×35 方向×20 枚）となる。図 6 に多方向顔画像の一例を示す。

#### 4.2 全方向の画像データを学習する手法

収集した顔画像データを用いて、四方向面特徴による顔識別能力の評価実験を行った。ここでは、顔向きの変異と人数に対する顔認識率によって評価を行った。顔向きの変異として、正面顔を含む縦横±15度の範囲の9方向(D09)、縦横±30度の範囲の25方向(D25)、全方向の35方向(D35)の3種類で比較した。

実験結果を図7に示す。全体的に人数が増えるにつれ認識率が低下し、特に顔向きの変異が増えるにつれ認識率の低下が顕著となった。D09は人数が150名と増えても認識率があまり低下せず高い識別率を保持することができた。このことは、正面顔近辺であるD09の範囲での顔画像の変化が四方向面特徴により十分に吸収されたことが示唆される。一方、D35のように左右45度の顔を単一クラスに含めた場合は、四方向面特徴空間においてもばらつきが大きく、そのため判別空間が適切に生成されず、認識率の低下を招いたと考えられる。これらより顔向きの変異を制限することで、顔識別率を向上させ得ることが示唆される。

#### 4.3 階層的顔画像認識手法

まず顔画像データに対して階層的人物識別手法の最適な組合せを探る。最適な組合せによる階層的認識手法を探ることによって、適応的判別空間手法の効果を検討する。

##### 4.3.1 顔向き推定実験

3種類の方向分解能、即ち、3方向、7方向、35方向で評価した。詳細を図8に示す。3方向、7方向の推定には、水平方向の顔向き別に分類した。このように分類した理由は、予備実験より水平方向より垂直方向に変化した顔画像の方が四方向面特徴空間においてばらつきが小さかったためである。

顔向き推定は、未知の人物に対しても機能しなければならない。そこで、本実験では50人を学習人物とし、別の50人を未知人物として任意に選出した。なお、判別空間の生成には学習人物の学習データのみを用いた。

実験結果を表1に示す。3方向の推定が最も推定率が高く、学習人物、未知人物問わず95%以上の識別率を得た。一方、35方向を推定した場合、識別率が73.6%、未知人物に対しては65.3%と識別率が低下した。3方向、7方向共に学習人物と未知人物で識別率に差がなく、本手法では少なくとも50人学習すれば汎化性があることが示唆された。

個人用顔向き推定判別空間を用いて、顔向きを推定した平均正答率を表1の括弧内に示す。個人用判別空間を用いることで、35方向においても95%以上の推定率を得ることができた。

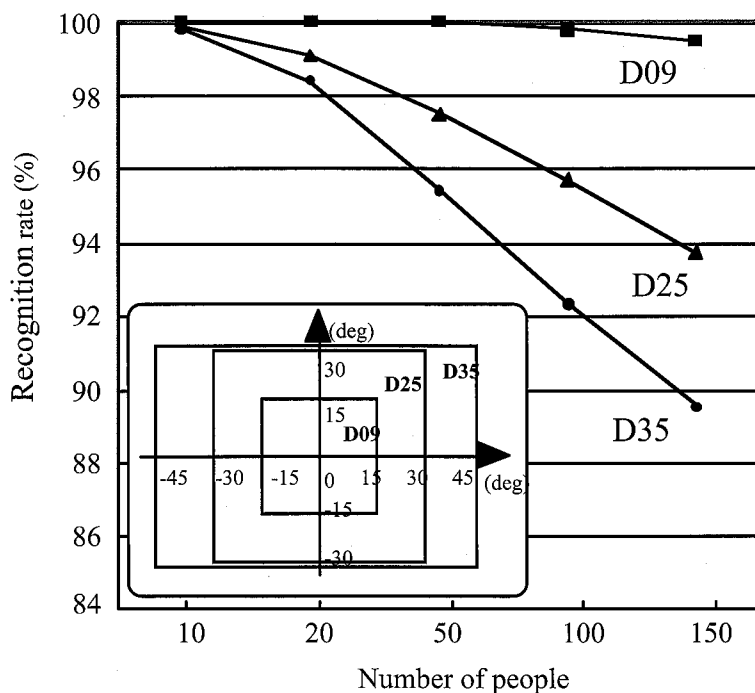


図7 人数と顔向きに対する認識率

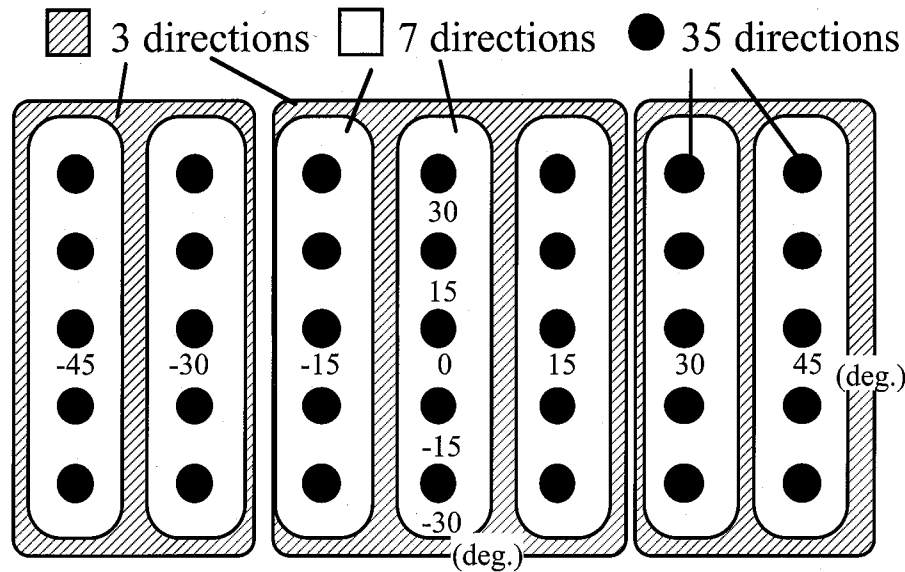


図 8 顔向き分類

表 1 顔方向推定結果 (%)

顔向き推定分解能	3方向	7方向	35方向
学習人物 (個人用判別空間)	95.0 (98.8)	86.0 (96.5)	73.6 (95.3)
未知人物	95.8	84.0	65.3

以上の結果より、未知の人物に対しては大まかに顔向きを推定し、個人が特定できれば、その人物の顔向き推定判別空間を用いることで詳細な顔向き推定ができることが示唆された。

#### 4.3.2 認識結果

前述したように、階層的人物識別では絞り込み構造が問題となる。つまり、顔向き推定の分解能と顔識別が受け持つ範囲の組合せが重要となる。特に、この過程を大分類として位置付けているため、対象人物が候補者として残るようになる必要がある。そこで、顔向き推定実験より推定精度が不十分なことから、顔向き推定を補うため、顔認識の判別空間として、顔向き推定の範囲より拡張した範囲を学習する階層構造を提案する。これにより、顔向き推定の誤りを顔認識でカバーすることができると思われる。

顔向き推定に3方向と7方向を選び、顔識別をその方向を示す顔画像のみを学習した基本範囲辞書と一回り拡張した範囲を学習した拡張範囲辞書との比較実験を行った。学習範囲について図9を用いて説明する。7方向の基本範囲辞書(a)の場合は、顔向き推定で分割された領域のみを学習する。例えば、正面顔(0度)と推定されれば、FD4の範囲を学習した顔認識判別空間で顔識別を行う。一方、拡張範囲辞書(b)では、FD4にFD3、FD5を加えた領域を対象として学習する。3方向においても同様に、顔向き推定範囲FD2の基本範囲辞書(c)に対して、拡張範囲辞書(d)は、一回り広げた範囲となる。実験データは、150名、35方向の顔画像データを用いた。顔向き推定には、先の顔向き推定実験で求めた判別空間を用いた。

実験結果を表2に示す。全ての組合せにおいて、図7に示した階層化していない手法から得た89.5%を上回った。さらに拡張範囲辞書を用いることで7方向の認識率を93.6%から97.6%に向上させることができた。7方向の顔向き推定率は3方向より低かったが、最終的な顔識別率は上回る結果が得られた。表2の括弧内に示した識別率は、顔向き推定結果が正しく推定されたときのシミュレーション結果である。7方向-拡張範囲辞書においては0.7%の差しか現れなく、拡張範囲辞書の効果を確認することができた。

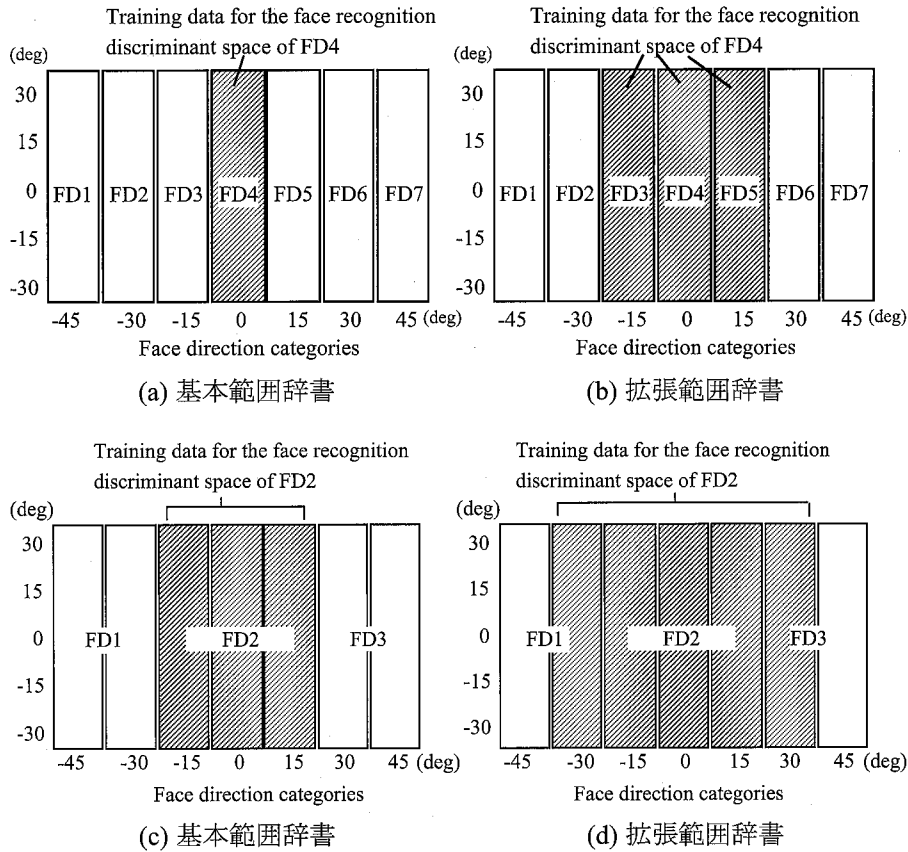


図9 顔向き別認識学習範囲例

表2 認識率 (%)

	顔向き推定	
	3方向 / (シミュレーション)	7方向 / (シミュレーション)
基本範囲辞書	95.6 / (97.7)	93.6 / (98.6)
拡張範囲辞書	95.6 / 96.5)	97.6 / (98.3)

表3 正解率と提案手法による認識率

Ranking	1	3	5	10	20	50	ADS
Rate (%)	97.6	98.9	99.3	99.6	99.8	100	98.8

ADS: 適応的判別空間手法

以上より、顔認識範囲を拡張することで、最終的な人物識別率を向上させることができ、本手法の有効性を確認した。

#### 4.4 適応的判別空間手法

先の階層的認識実験において、個人差による顔向きのばらつきから顔識別の学習範囲を拡張させることは、やはり顔識別率の低下を招くことが示された。

顔向きの分解能と顔識別辞書の範囲を調整することである程度までの認識率向上は期待できるが、顔画像データに強く依存することになる。そこで、本提案手法を評価するために、最も識別率が高かった7方向-拡張範囲辞書の結果に対して、適応的判別空間手法を適用した。

表3に7方向-拡張範囲辞書で選出された候補者の中に正解が含まれる正答率を示す。1位の97.6%は、階層的人物識別の実験結果を意味する。50位まで候補者を増やすと100%含まれることに



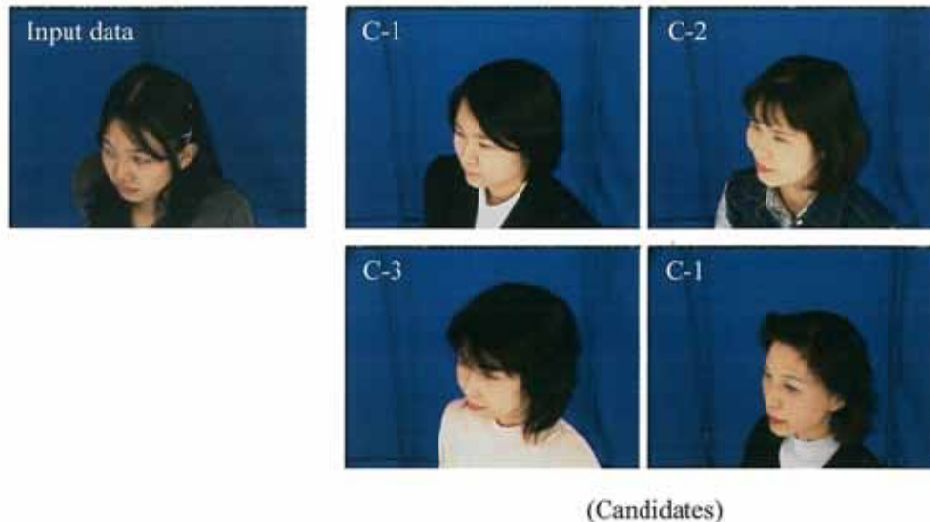


図 10 候補者例

なるが、候補者全体の 1/3 となり以降の識別に負担がかかる。今回は 99.6%の正答率に達する上位 10 名を候補者とした。次に、各候補者の顔向きを各人の顔向き推定判別空間により推定する。ここでは、35 方向の推定を行う。その推定結果に基づいて、人物識別に用いる顔画像を選出し、人物識別空間を作成する。つまり、顔向きが限定された 10 名の顔画像から選出することになる。

実験結果から認識率は、98.8%を得た（表 3）。99.6%を越えることはできないが、階層的人物識別手法の 97.6%を向上させることができた。向上した要因のひとつとして、個人別顔向き推定により正しく顔向きが推定されるようになったことが挙げられる。

図 10 は、階層的人物識別手法において顔向きを誤っていたために、隣に示した候補者が選出された例である。これが個人用顔向き推定を用いることで入力画像の正しい顔向きを推定することができ、その結果、新たに生成された判別空間により正しく人物識別されたと考えられる。

本実験において、階層的人物識別手法では 1 位で識別されていたにも拘わらず、適応的判別空間手法を用いることで逆に誤った結果が 17 件、全体の 0.03%観察された。非常に少数のデータではあるが、顔画像を 35 方向まで絞り込むことで識別クラスのばらつきが減る反面、学習データによっては非常に偏っている場合も考えられる。この問題は、新たな判別空間を生成する際に、近傍の顔向き顔画像を加えるなどして回避できると思われる。

今回、処理時間として Octane(SGI, R12000, 300MHz)上で、一枚の画像を最終判定まで処理するのに約 13 秒要した。ほとんどの時間が、新しい判別空間を生成するために要した時間であった。つまり、10 クラスの学習に要した時間である。今回は候補者を上位 10 としたが、大分類された時点ですでに、候補者の 1 位と 2 位との間で判別空間上の距離が離れている場合が多く観察されている。判別空間上の距離値を用いて、候補者を 10 名以下に絞り込む手法やその時点で人物を確定する手法などによって時間短縮も可能と思われる。

本手法は、顔向き推定範囲を拡張することで認識可能な顔方向を拡張できる。また、階層構造を工夫することにより、大量の顔画像データから類似した人物を認識することが可能と思われる。

### フェーズ III

顔認識の応用例として、「動画像検索機能付き次世代画像ファイリングソフト」、「動画像からの人物/シーン検索機能」、「顔画像情報による操作画面の自動選択ソフト」のプロトタイプ開発を行った。その概要は以下の通りである。

#### (1) 動画像検索機能付き次世代画像ファイリングソフト

デジタルカメラやスキャナ等の普及により、個人が扱う画像データは多量になってきている。こ

のような多量の画像データを扱う際に、顔画像認識を用いて画像データを自動的に個人毎にグループ分けすることができれば非常に有用である。また、顔認識以外にも、各画像に対して属性やコメント等の情報を付加することにより、多様な検索手法が実現できる。本ソフトは、多量の画像群から顔画像や検索条件の入力によって様々な検索を実現する。また動画像からも、顔画像検出を行い、同様の処理を行うことで画像データのファイリングを効率的かつ容易にサポートすることが出来る。

## (2) 動画像からの人物／シーン検索機能

動画像を見る際に、特定の人物やシーンが登場する場面を検索したい場合がある。本機能は、入力画像を用いた顔画像認識により、動画像全体を走査し、目的の人物を探索する。検出されたシーンをサムネイル化によって表示することで、ユーザは目的のシーンを簡単に探し出すことができる。本機能によって、動画像中の目的場面の頭出しや、場面の選択が容易となるため、見るだけでなく、編集作業もサポートすることが期待できる。

これらのプロトタイプの開発から以下の改善点が挙げられた。

- ・ 顔画像認識の精度を向上し、検索ヒット率の向上
- ・ 画像の写りによる違い（顔向き、照明条件等）に対する頑健性の向上
- ・ 少数の学習画像からの学習手法の検討

## 今後の取り組み

顔画像認識は、個人適応型のマンマシンインタフェースや個人認証システムにおいて重要な要素技術である。また、プロトタイプを開発したような応用も、今後のデジタルメディアの増大によって有用となることが予想される。

今後の課題としては、応用システムのプロトタイプ開発で挙げられた点（顔認識の高精度化、頑健性の向上、少数の画像データによる学習）の改善である。これらの技術を高めることで、実用的なアプリケーションシステムが期待できる。