

修士論文

# 濃度こう配特徴を用いた 顔画像の照合と認識



平成 21 年度修了  
三重大学大学院工学研究科  
博士前期課程 情報工学専攻

小牧 優士

# はじめに

従来の個人認証の技術として、カードなどの所有物やパスワードなどの秘密情報を用いた方法がある。しかし、カードの盗難、紛失、パスワードの漏洩などの問題があり、他人によるなりすましに対して十分なセキュリティが確保されているとはいえない状況である。これらの問題を解決するために生体情報による個人認証が注目されている。この技術はバイオメトリクスと呼ばれ、個人ごとに異なる体の部位の特徴を用いて個人認証を行う技術である。バイオメトリクスに利用される特徴は、指紋、虹彩、声紋、顔などがある。本研究では、生体情報の一つである顔を用いた個人認証を取り扱う。顔画像を用いた個人認証は、入退室管理、不審者の照合などへの応用が期待され、盛んに研究されている [1]。

顔画像認証の基礎技術として顔画像認識がある。顔画像認識とは、入力画像と事前取得され登録されている複数個人のテンプレート画像とを比較し、入力画像中の人物がどの人物であるかを特定することである。この処理は、入力画像中の人物が登録されているテンプレート中に含まれているという仮定の下で行われる。しかし、顔画像認識の一般的な利用場面においては、事前に登録されていない部外者が入力として与えられることが考えられる。この場合、部外者を棄却する必要がある。入力として与えられた画像中の人物が部内者か部外者かを判定し、部外者を棄却する処理を本研究では照合と呼ぶ。

顔画像認証の特徴としては

- 動画を用いることが可能である。
- 指紋認証における、センサーに指を押しつけるといった照合動作が不要である。
- 人間が個人を識別する方法と同じであるため直感的でわかりやすい。

などがあげられる。しかし、顔画像認証には、画像中の顔の向きや顔への照明の当たり方、眼鏡や鬚、時間の経過による顔そのものの变化などによって簡単に大きく見え方が変化するため認証が困難であるという問題点がある。

本研究では、Beveridge らが開発した顔認識システム CSU Face Identification Evaluation System[2] を使用する。このシステムでは顔画像の画素値を特徴量とする濃淡特徴を用いて認識を行っており、高次元の特徴ベクトルを処理する必要がある。本研究では、特徴量として濃度こう配特徴を導入し、顔画像の照合、認識の性能向上を目的とする。濃度こう

配特徴とは画像の濃度こう配の強度と向きとを用いた特徴量である。

実験では FERET データベース [3] を用いて、学習データにおける 1 人あたりの画像数を変化させ、照合、認識性能に与える影響を調べた。照合において、特徴量に濃度こう配特徴を使用し、次元削減に主成分分析（Principal Component Analysis : PCA）と線形判別分析（Linear Discriminant Analysis : LDA）とを組み合わせると、1 人あたりの学習画像数を増加させたときの等価エラー率（Equal Error Rate : EER）が低下していくことがわかった。学習に 1 人あたり平均 4 枚の画像を用いた場合と、1 人あたり 1 枚の画像を用いた場合とを比較すると EER が 28.64% から 15.93% まで低下した。また、1 人あたりの画像数が平均 4 枚の学習データを用いて、次元削減に PCA と LDA を組み合わせると適用した場合について、濃淡特徴を用いた場合と濃度こう配特徴を用いた場合との照合性能を比較した。Total では EER が 27.88% から 15.93% まで低下した。表情変化では 21.08% から 10.36% まで、光源変化では 29.44% から 18.75% まで、経年変化では 34.76% から 26.02% まで EER がそれぞれ低下した。

次に、部内者のうちどの人物であるかを特定する認識の性能評価を行った。学習データにおける 1 人あたりの画像数を増加させた場合、特徴量に濃度こう配特徴を用いて、次元削減に PCA と LDA を組み合わせると適用すると、認識エラー率が低下していくことがわかった。学習に 1 人あたり平均 4 枚の画像を用いた場合、1 人あたり 1 枚の画像を用いた場合と比較すると、認識エラー率が 48.11% から 20.27% まで低下した。また、学習に 1 人あたり平均 4 枚の画像を用いた場合に、特徴ベクトルの次元削減に LDA を適用し、濃淡特徴を用いた場合と濃度こう配特徴を用いた場合との認識性能を比較した。濃度こう配特徴を用いた方が Total の認識エラー率が 34.37% から 20.27% まで低下した。表情変化では 21.31% から 11.4% まで、光源変化では 29.85% から 23.88% まで、経年変化では 64.97% から 39.21% まで認識エラー率がそれぞれ低下した。

これらの実験結果より、顔画像の照合、認識において、特徴量に濃度こう配特徴を使用し、次元削減に PCA と LDA を組み合わせると適用した場合、学習データの 1 人あたりの画像数を増加させると照合、認識の性能が向上することがわかった。また、今回の実験では評価用画像の人物は学習データには含まれない。つまり、別の人物の画像を用いて学習を行っても照合、認識の性能が向上することがわかった。

今後の課題として、顔画像の照合、認識の更なる性能向上、顔検出処理と統合し、全自動の顔画像の照合、認識の評価などが挙げられる。

# 目次

|  |           |
|--|-----------|
| はじめに   | i         |
| <b>第 1 章 緒言</b>  | <b>1</b>  |
| 1.1 研究の背景  | 1         |
| 1.2 関連する従来研究   | 2         |
| 1.2.1 画像全体を用いる手法 (Holistic matching method)                      | 2         |
| 1.2.2 特徴に基づく構造マッチング手法 (Feature-based structural matching method) | 5         |
| 1.2.3 画像全体を用いる手法と特徴に基づく手法とを組み合わせた手法 (Hybrid Approachs)           | 6         |
| 1.3 研究の目的  | 7         |
| 1.4 本論文の構成   | 8         |
| <b>第 2 章 本研究に関連する理論と技術</b>                                       | <b>9</b>  |
| 2.1 濃度こう配特徴 (gradient features)                                  | 9         |
| 2.1.1 濃度こう配特徴とは  | 9         |
| 2.1.2 濃度こう配特徴の抽出方法   | 9         |
| 2.2 特徴選択 (次元削減)  | 11        |
| 2.2.1 主成分分析 (Principal Component Analysis : PCA)                 | 11        |
| 2.2.2 線形判別分析 (Linear Discriminant Analysis : LDA)                | 13        |
| <b>第 3 章 顔画像の照合と認識</b>   | <b>15</b> |
| 3.1 前処理  | 15        |
| 3.2 特徴抽出   | 16        |
| 3.3 特徴選択 (次元削減)  | 16        |
| 3.4 距離計算   | 17        |
| 3.5 分類   | 17        |

|              |                           |           |
|--------------|---------------------------|-----------|
| <b>第 4 章</b> | <b>実験</b>                 | <b>18</b> |
| 4.1          | 実験データ . . . . .           | 18        |
| 4.1.1        | 学習データ . . . . .           | 18        |
| 4.1.2        | 評価用画像 . . . . .           | 19        |
| 4.2          | 評価方法 . . . . .            | 20        |
| 4.3          | 実験 . . . . .              | 21        |
| 4.4          | 結果と考察 . . . . .           | 21        |
| <b>第 5 章</b> | <b>結言</b>                 | <b>26</b> |
| 5.1          | まとめ . . . . .             | 26        |
| 5.2          | 今後の課題 . . . . .           | 27        |
| 付録 A         | 一般化固有値問題から標準固有値問題への変換     | 28        |
| 付録 B         | 研究データ                     | 31        |
| B.1          | プログラムソースファイルリスト . . . . . | 31        |
| 謝辞           |                           | 34        |

# 第 1 章

## 緒言

### 1.1 研究の背景

個人を認証することは、PC へのログイン、銀行 ATM など多くの場面で必要となっている。個人情報の漏洩やなりすましによる被害が増加しており、それらを防ぐために個人認証の重要性が高まっている。

従来の個人認証の技術として、カードなどの所有物やパスワードなどの秘密情報を用いた方法がある。しかし、カードの盗難、紛失、パスワードの漏洩などの問題があり、他人によるなりすましに対して十分なセキュリティが確保されているとはいえない状況である。これらの問題を解決するために生体情報による個人認証が注目されている。この技術はバイオメトリクスと呼ばれ、個人ごとに異なる体の部位の特徴を用いて個人認証を行う技術である。バイオメトリクスに利用される特徴は、指紋、虹彩、声紋、顔などがある。本研究では、生体情報の一つである顔を用いた個人認証を取り扱う。顔画像を用いた個人認証は、入退室管理、不審者の照合などへの応用が期待され、盛んに研究が行われている [1]。

顔画像認証の基礎技術として顔画像認識がある。顔画像認識とは、入力画像と事前を取得され登録されている複数個人のテンプレート画像とを比較し、入力画像中の人物がどの人物であるかを特定することである。この処理は、入力画像中の人物が登録されているテンプレート中に含まれているという仮定の下で行われる。しかし、顔画像認識の一般的な利用場面においては、事前に登録されていない部外者が入力として与えられることが考えられる。この場合、部外者を棄却する必要がある。入力として与えられた画像中の人物が部内者か部外者かを判定し、部外者を棄却する処理を本研究では照合と呼ぶ。

顔画像認証の特徴としては

- 動画を用いることが可能である。

- 指紋認証における、センサーに指を押しつけるといった照合動作が不要である。
- 人間が個人を識別する方法と同じであるため直感的でわかりやすい。

などが挙げられる。しかし、顔画像認証には、画像中の顔の向きや顔への照明の当たり方、眼鏡や鬚、時間の経過などによる顔そのものの变化などによって簡単に大きく見え方が変化するため認証が困難であるという問題点がある。

## 1.2 関連する従来研究

自動顔画像認識は、建物の入退室管理、銀行 ATM などでの個人識別、不審者の犯罪者リストとの照合など様々な分野への利用が期待され、盛んに研究が行われている。

自動で顔画像認識を行うには、以下の 3 つの処理を全て実現する必要がある。

1. 顔を含んだ画像全体から顔領域を検出し切り出す（顔検出）。
2. 切り出された顔領域から顔の特徴を抽出する（特徴抽出）。
3. 得られた特徴をもとに、入力画像中の人物が事前に登録されている画像中のどの人物であるかを特定する（認識）。

認識処理は全自動顔画像認識における最終ステップであり、この処理の精度低下は全自動顔画像認識そのものの精度低下に直結する。そのため、認識処理の高精度化が非常に重要となる。ここでは従来の顔画像認識に関する研究について概説する。

これまで、多くの顔画像認識の手法が提案されてきたが、それらは以下の 3 つのカテゴリーに大別できる。

1. 画像全体を用いてマッチングを行う手法（Holistic matching method）
2. 特徴に基づいてマッチングを行う手法（Feature-based matching method）
3. 画像全体と特徴との両方を用いる手法（Hybrid method）

本研究で使用した濃度こう配特徴と濃淡特徴は、画像全体を用いてマッチングを行う手法に属する。

それぞれのカテゴリーの特徴と、それぞれのカテゴリーに属する手法について次に述べる。

### 1.2.1 画像全体を用いる手法（Holistic matching method）

認識処理部に入力画像をそのまま処理させる手法で、顔領域全体を利用する。顔領域の表現方法として広く利用されているのが、主成分分析（Principal Component Analysis : PCA）に基づく固有顔（eigenfaces または eigenpictures）[4][5] である。

### 固有顔 (eigenfaces or eigenpictures)

この手法は、PCAにより顔画像を低次元の特徴で再構成する方法である。PCAを用いて顔画像を低次元の特徴で表す手順は後述(2.2.1)する。原画像は高い冗長性を含むことが知られている[6]。顔画像のようにスケール、回転などが正規化され、限られたクラス数のオブジェクトにおいては冗長性が増加する[7][8]。PCAは画像全体を表現する方法で、基底の非相関化を行う。この表現方法を用いることで冗長性を低下させるとともに、顔の部分的な隠れへの敏感さを低減させるメリットがある。

機械による顔画像認識の最初の成功例はこの固有顔[9]を用いた手法である。学習データの特徴ベクトルを用いてPCAを行い、固有顔と呼ばれる固有ベクトルを事前に求める。入力顔画像はそれぞれの固有顔に対する重み係数からなる特徴ベクトルで表せる。これにより、ある未知の顔画像に対し、その特徴ベクトルを生成し、顔画像データベースのそれぞれの特徴ベクトルとのユークリッド距離を求める。その中から、距離値が最小となる画像の人物に認識される。

多くの固有顔に基づくシステムにおいて、画像のぼやけ、部分的な隠れ、背景の変化がある場合に性能が高いことが証明されている。また学習データセットに鏡像を加えることで、性能が向上することが示された[4]。

### ベイズアプローチ (Bayesian Approach)

固有顔による個人認識では、認識処理部でユークリッド距離を用いてデータベース中のどの人物かを特定していた。ユークリッド距離ではなく、類似度の確率的な尺度を導入することで標準的な固有顔のアプローチはベイズアプローチ[10]へと拡張された。このベイズアプローチの欠点は、各クラスにつき非常に限られた数の学習サンプルから高次元空間における確率分布を推定する必要があることである。この問題を避けるために、多クラス分類問題はベイズ分析 (Bayesian Analysis) に基づく類似性評価を用いてよりシンプルな2クラス分類問題へと変換される。

### 線形判別分析 (Linear Discriminant Analysis : LDA) を用いる手法

LDA[11]を用いた顔画像認識も高い性能が得られることが知られている[12][13][14][15]。LDAによる学習は散乱行列分析によって行われる[16]。LDAを用いて顔画像を低次元の特徴で表す手順は後述(2.2.2)する。

固有特徴 (eigenfeatures) の判別分析[14]は、顔か非顔かのクラスを決定するだけでな



く、顔クラス内の個人も決定するために画像検索システムに適用された。

### 進化論的追跡 (Evolution Pursuit : EP)

EP に基づく、順応性のある表現とその顔画像認識への応用が発表されている [17]。この手法は射影追跡法 (projection pursuit method) と類似しており、データ圧縮とパターン分類のために、学習により最適基底を求める。EP の課題は、未知画像に対して、学習における経験的リスクを低減させることと、保証付きリスクの信頼区間を狭めることに伴う経験的リスクを低減させることとのバランスを取り、学習機械としての汎化能力を向上させることである。この目的のために、EP は遺伝的アルゴリズム (Genetic Algorithm : GA) の特徴が実装され、最適基底を決定するための解を探索する。EP はオリジナルのデータを低次元のホワイトニングされた PCA 空間へ投影することから始まる。この空間中でランダムな向きを持つ基底ベクトルを探す。このとき、性能とクラス分離とによって定義される適応度関数で進化が行われる。

### 独立成分分析 (Independent Component Analysis : ICA)

顔画像認識に ICA を用いる手法が提案されている [19]。ICA は PCA の一般化であり、どちらも非相関化を行うが、PCA が 2 次モーメントの非相関化を行うのに対し、ICA は 2 次以上の高次モーメントの非相関化を行う。また、これに関連した ICA の特徴として、

- 非直交な成分分離
- 基底の空間的局所性

が挙げられる。

### ニューラルネットワークを用いた手法

ニューラルネットワークを用いた全自動顔検出 / 認識システムが報告されている [20]。提案されたシステムは Probabilistic Decision-Based Neural Network (PDBNN) [21] に基づいており、以下の 3 つのモジュールで構成される。

- 顔検出モジュール
- 目位置特定モジュール
- 顔認識モジュール

他の多くの手法とは異なり、顔領域に含まれるのは両眉毛、両目、鼻のみで、口は含まれない [4]。顔の上部のみを用いるのは、口周辺の変化による表情変化などの影響を除外し、頑健なシステムを構築するためである。さらに頑健にするために、顔の上部領域の画像の解像度を  $14 \times 10$  画素に低下させ、正規化された輝度とエッジの特徴を生成する。これらの特徴の値は、2 つの PDBNN にそれぞれ与えられ、最終的な認識結果は 2 つの PDBNN から出力された値を融合して求められる。

### その他の手法

固有顔に基づく手法は、特徴空間における入力画像と参照用テンプレートとの距離を用いた最近傍法による認識である。このような点と点の距離を用いるのではなく、入力画像と 2 枚の参照用テンプレート間を結ぶ線分との距離を用いる手法が提案されている [22]。

また分類器にサポートベクタマシン (Support Vector Machine : SVM) を用いた手法も提案されている [23]。

### 1.2.2 特徴に基づく構造マッチング手法 (Feature-based structural matching method)

この手法では目、鼻、口のような局所的な特徴が抽出され、それらの配置、幾何学的な位置や見えかたの統計量などが分類器に与えられ、認識が行われる。

この手法グループに属する代表的な手法には、局所特徴の形状に基づく手法 [24][25]、1 次元隠れマルコフモデル (Hidden Markov Model : HMM) の手法 [26]、疑似 2 次元 HMM [27] などがある。Elastic Bunch Graph Matching (EBGM) [28][29] は最も成功したシステムの 1 つである。そのシステムはダイナミックリンクアーキテクチャ (Dynamic Link Architecture : DLA) [30][31] に基づいている。ウェーブレット、特にガボールウェーブレットは顔の表現のためにこれらのグラフマッチング手法において、ビルや建物のような役割を果たす。局所特徴表現は、jet と呼ばれるウェーブレット特徴に基づく異なるスケール、異なる回転角のウェーブレット係数で構成される。これらの局所的に推定されるウェーブレット係数は、光源変化、位置のずれ、ゆがみ、回転、スケーリングに対して頑健である。

### Elastic Bunch Graph Matching (EBGM)

この手法は DLA を拡張した手法である [29]。姿勢変動の問題を解決するために、顔の姿勢は事前のクラス情報を用いて最初に決定され [32]、姿勢変動に応じて jet 変換が学

習される [33]. EBGM アプローチに基づくシステムは、顔検出、顔抽出、姿勢の推定、性別分類、スケッチ画像に基づく認識、一般的な物体認識などに応用されている。EBGM の成功は、人間の視覚システムに類似していることが理由として挙げられる。

### 1.2.3 画像全体を用いる手法と特徴に基づく手法とを組み合わせた手法 (Hybrid Approachs)

この手法では、全体的な特徴と局所的な特徴との両方を用いる。例として、modular eigenfaces[34] では全体の固有顔と局所的な固有顔の両方を用いる。

eigenfaces の概念は eigeneyes, eigenmouth などのような固有特徴へと拡張された。低次元空間においては、eigenfaces よりも固有特徴の方が性能が高い。

#### PCA と局所特徴分析 (Local Feature Analysis : LFA) とのハイブリッド

実際のシステムには PCA と LFA とのハイブリッドの手法を用いるべきであると主張されてきた [7]。大きい固有値を持つ場合に PCA の性能は高くなるが、高次元の場合は LFA を用いる方が性能が高くなる。主な eigenpictures のシステムは、全体的、集約的であるためノイズを抑えるのに有効な平滑化フィルタであり、一方、高次元のモードはさざ波、すなわちノイズを増幅させる分離フィルタであると論じられた [7]。

#### フレキシブルな見え方モデルに基づく手法

この手法では、顔を特定するために、形状と画素値の情報とをモデル化して用いる [35]。この形状モデルは Active Shape Model (ASM) と呼ばれ、画像の形状に適するように変形する、オブジェクトの統計的な形状モデルである。統計的形状モデルは PCA を用いて学習される。このときの変数は形状モデルの座標である。判別分析法によって、クラス間変動による形状変動はクラス内変動による形状変動とは分離され、分類が行われる。

平均的な形状モデルに基づき、形状を考慮しない全体的なモデルが PCA を用いて生成される。隠れなどの局所的な見え方の変化に対して頑健になるように、局所モデルが形状モデル上に作成される。モデルの輪郭線に直交する方向における画素値の分布を手がかりとする。最後、入力画像から抽出された、形状パラメータ、形状を考慮しない画像のパラメータ、局所的な輪郭線の 3 つの情報をを用いて、マハラノビス距離を計算する。

### 顔部品に基づく手法

顔部品を用いた、顔検出と顔認識のシステムが提案されている [36][37]. 顔部品に基づく手法の基本的な考え方は、顔を口や目などの顔部品へと分解することである. 口や目などの顔部品は、柔軟な幾何学的モデルによって相互に連結される. この手法は先述の EBGM と類似しているが、ガボールウェーブレットではなくグレイスケールの顔部品を用いる点で異なる. 顔部品を用いる理由は、頭部の姿勢の変化により顔部品の位置も変化してしまうが、顔部品の位置は幾何学的モデルの柔軟性により位置が特定できるからである. しかし、この手法の大きな欠点は、異なる視点、異なる照明条件で撮影された大量の学習用画像が必要となることである.

### 3次元モーフィングモデル

顔部品に基づく手法における大量の学習用画像が必要となる問題を解決するために、3次元モーフィングモデル (3D morphable models) [38] が用いられている. このモデルは、様々な姿勢、照明条件の任意の合成画像を生成する. 3次元の顔モデルを生成するために、正面、斜め、横のそれぞれから1枚ずつ撮影した合計3枚の画像しか使用しない. 一度3次元モデルを生成すれば、検出、分類に必要な合成画像が生成できる.

## 1.3 研究の目的

本研究では、Beveridge らが開発した顔認識システム CSU Face Identification Evaluation System[2] を使用する. このシステムでは、顔画像の画素値を特徴量とする濃淡特徴を用いて処理を行っており、高次元の特徴ベクトルを処理する必要がある. 濃淡特徴は、画像の画素値をそのまま要素として特徴ベクトルを生成するため光源の変化に対して弱い. また、特徴ベクトルの次元数が高次元になるため、照合、認識に有効でない特徴が含まれることや計算コストの面でも望ましくない. 本研究では、特徴量として濃度こう配特徴を導入し、顔画像の照合、認識の性能向上を目的とする. 濃度こう配特徴とは、画像の濃度こう配の強度と向きとを用いた特徴量であり、手書き文字認識では高い精度が得られることが知られている. 濃度こう配特徴を抽出して得られる特徴ベクトルの次元数は、パラメータにより変化するが、濃淡特徴より一般的に小さい. 濃度こう配特徴についての詳細は、2.1 で詳しく述べる.

## 1.4 本論文の構成

第2章では，本研究で取り扱う濃度こう配特徴，特徴選択（次元削減）手法について述べる．第3章では，本研究における処理の流れと各処理の詳細について説明する．第4章では，実験に使用した顔画像の例，実験の条件，結果と考察について述べる．第5章では，本研究のまとめと今後の課題について述べる．

## 第 2 章

# 本研究に関連する理論と技術

この章では、画像中の顔の見え方を濃度こう配特徴を用いて表現する方法と、特徴選択により特徴ベクトルの次元数を削減する方法について述べる。

### 2.1 濃度こう配特徴 (gradient features)

#### 2.1.1 濃度こう配特徴とは

濃度こう配特徴とは画像の濃度のこう配をその向きごとにヒストグラム化した特徴量である。濃度こう配特徴は手書き文字認識においては高い精度が得られることが知られている [39]。画像の濃度こう配を用いる特徴量として、近年では、濃度こう配特徴と同様の手法によって算出される HOG 特徴 [40]、スケーリングや回転に対して不変な Scale-Invariant Feature Transform (SIFT) [41]、SIFT を高速化した Speeded Up Robust Features (SURF) [42] など提案されている。本研究では、濃度こう配特徴をグレースケールの顔画像から抽出し、顔の見かけを表現する。

#### 2.1.2 濃度こう配特徴の抽出方法

濃度こう配特徴の一般的な抽出方法は以下の通りである。

1. 入力画像  $I = I(x, y)$  の画素  $(x, y)$  に対して以下のガウシアンフィルタを適用する。

$$I'(x, y) = \frac{\sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} f(i, j) I(x + i, y + j)}{\sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} f(i, j)}, \quad (2.1)$$

$$f(i, j) = \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right). \quad (2.2)$$

ここで,  $N$ ,  $\sigma$  はそれぞれ, 5, 1.0 である.

2. 入力画像に対し Sobel オペレータを適用し, エッジ検出処理を行う. これにより, 画素  $(x, y)$  における濃度こう配の水平成分  $G_x(x, y)$  と垂直成分  $G_y(x, y)$  とを求める. 算出された  $G_x(x, y)$ ,  $G_y(x, y)$  を用いて, 次式により濃度こう配の強度  $G(x, y)$  と向き  $\theta(x, y)$  とを算出する.

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \quad (2.3)$$

$$\theta(x, y) = \tan^{-1} \left\{ \frac{G_y(x, y)}{G_x(x, y)} \right\}. \quad (2.4)$$

算出されたこう配の向きを  $L$  段階に量子化する. 入力画像の横, 縦それぞれの画素数を  $X$ ,  $Y$  とすると, この時点の特徴の数 (特徴ベクトルの要素数) は  $X \times Y \times L$  となる.

3. 画像を  $n \times m$  の小領域に分割する. それぞれの小領域内で, こう配強度を向きごとに累積してヒストグラム化する. この処理により特徴の数は  $n \times m \times L$  に削減される.
4. ヒストグラム化されたこう配強度を, 分割された任意の小領域  $(2i, 2j)$  ( $i = 0, 1, \dots, \frac{n}{2} - 1$ ,  $j = 0, 1, \dots, \frac{m}{2} - 1$ ) を中心とした  $5 \times 5$  の領域について, 向きごとにガウシアンフィルタにより平滑化する. この処理により, 領域数を削減する.
5. ヒストグラム化した濃度こう配の向きを削減する.  $L$  段階に量子化された向きに対しひとつおきに窓の中心を設定し,  $[1 \ 4 \ 6 \ 4 \ 1]$  の重みを用いて加重平均して向きを  $L/2$  に削減する.

手順 3. のガウシアンフィルタによる平滑化により, 画像中の顔の位置のずれを吸収することができると考えられる. また手順 4. の濃度こう配の向きに対する加重平均により, 連続値で得られたこう配の向きを  $L$  段階に量子化する際の量子化誤差を低減できると考えられる. 以上の処理により抽出した濃度こう配画像の例を図 2.1 に示す. 図はこう配の強度, 向きをそれぞれ明度, 色相によって表現している. このとき, 生成される特徴ベクトルの次元数は以下の式により求めることができる.

$$D = \frac{(n+1)(m+1)L}{8}. \quad (2.5)$$

$n, m, L$  それぞれの値は, 特徴ベクトルの次元数を変動させ照合, 認識の性能に影響を与える.

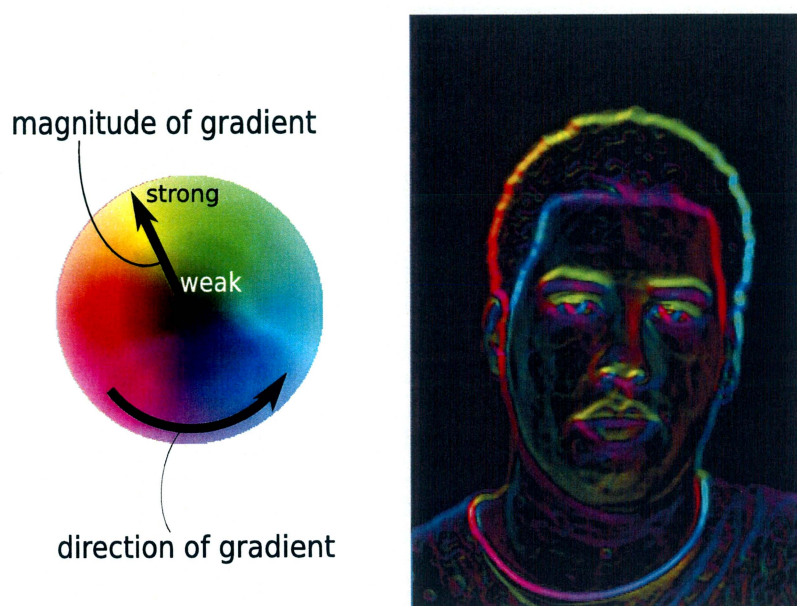


図 2.1: 濃度こう配画像の例

## 2.2 特徴選択（次元削減）

濃度こう配特徴、濃淡特徴を抽出して生成される原特徴ベクトルはいずれも高次元となる。高次元の原特徴ベクトルには相関の高い要素の組が含まれ、冗長性を含んだ特徴ベクトルになる [6]。また、高次元の特徴ベクトルを照合、認識に用いるとエラー率や計算コストが増加する。そのため、特徴抽出により得られた特徴ベクトルに対し、特徴選択により次元削減を行う。これにより照合、認識に有効であると考えられる特徴が選択され、精度の向上が期待できる。本研究では主成分分析と線形判別分析の 2 手法を用いる。

### 2.2.1 主成分分析（Principal Component Analysis : PCA）

PCA は多次元空間上の特徴点を分散の大きい少数の低次元の直交部分空間に線形射影する手法である。多次元空間上の特徴点をより見やすくあるいは扱いやすくするために、固有空間を利用して少ない次元で表現する手法である。多次元の特徴量を低次元化することにより、照合、認識に有効な特徴を選択することができる。例として、2 次元の特徴ベクトルで表現された 2 クラスのサンプルを 1 次元に削減する場合を図 2.2 に示す。図 2.2 において、全サンプル分布の分散が最大となる方向に主軸  $y_1$  をとる。サンプルを主軸  $y_1$  に投影することで、1 次元データでも 2 クラスの分離が可能となる。

$n$  次元の原特徴ベクトルを PCA により  $n'$  次元に削減する手順を以下に示す。



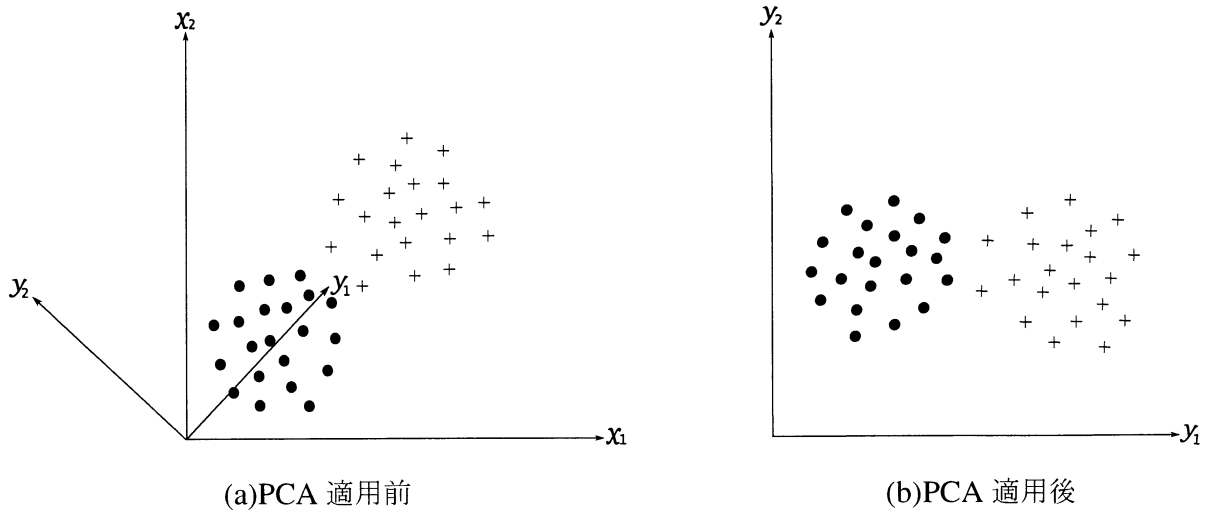


図 2.2: PCA による次元削減の例

学習データに含まれる画像  $j$  ( $j = 1, 2, \dots, m$ ) から抽出した特徴ベクトルを

$$\mathbf{I}_j = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{pmatrix} \quad (2.6)$$

と表す.

学習データの全平均ベクトル  $\mathbf{M}$ , 散乱行列  $\mathbf{S}_t$  をそれぞれ以下の式により計算する.

$$\mathbf{M} = \frac{1}{m} \sum_{j=1}^m \mathbf{I}_j, \quad (2.7)$$

$$\mathbf{S}_t = \sum_{j=1}^m (\mathbf{I}_j - \mathbf{M})(\mathbf{I}_j - \mathbf{M})^T. \quad (2.8)$$

以下の式を満たす固有値行列  $\mathbf{\Lambda}$ , 固有ベクトル行列  $\mathbf{\Phi}$  をそれぞれ求める.

$$\mathbf{S}_t \mathbf{\Phi} = \mathbf{\Phi} \mathbf{\Lambda}. \quad (2.9)$$

固有値を大きい順にソートし, 各固有値に対応する  $d$  個の固有ベクトルを用いて, 以下の式により原特徴ベクトルを  $d$  次元まで削減する ( $d \leq n$ ). ここで  $\mathbf{x}$  は原特徴ベクトル,  $\mathbf{y}$  は次元削減後の特徴ベクトルである.

$$\mathbf{y} = \mathbf{\Phi}^T \mathbf{x}. \quad (2.10)$$

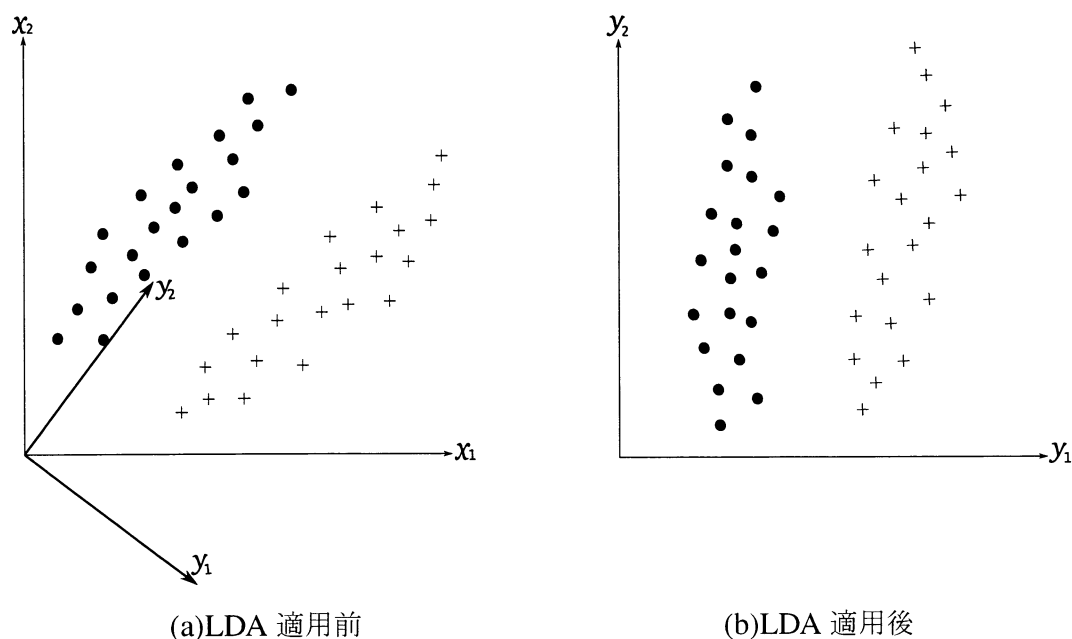


図 2.3: LDA による次元削減の例

### 2.2.2 線形判別分析（Linear Discriminant Analysis : LDA）

LDA ではフィッシャーの線形判別を用いる。フィッシャーの線形判別は、クラス内変動に対するクラス間変動の比を最大にすることでデータの分離性を高め、優れた低次元部分空間を得ることができる。次元削減後の次元数は（学習サンプル数 - 1）まで削減できる。例として、2次元の特徴ベクトルで表現された2クラスのサンプルをLDAを用いて1次元に削減する場合を図2.3に示す。図2.3において、クラス内変動に対するクラス間変動の比を最大にする主軸  $y_1$  を決定する。サンプルを主軸  $y_1$  に投影することで、1次元データでも2クラスの分離が可能となる。

$n$ 次元の原特徴ベクトルをLDAにより  $n'$ 次元に削減する手順を以下に示す。

学習データに含まれる画像  $j$  ( $j = 1, 2, \dots, m$ ) から抽出した特徴ベクトルを式 (2.6) で表す。

以下の式により級内散乱行列  $S_w$ 、級間散乱行列  $S_b$  を計算する。ここで、 $\mathbf{X}_k$ ,  $\mathbf{M}_k$ ,  $\mathbf{M}$  はそれぞれクラス  $k$  に属する画像から抽出された特徴ベクトル、クラス  $k$  に属する画像から抽出された特徴ベクトルの平均、全ての画像から抽出された特徴ベクトルの平均で

ある.

$$\mathbf{S}_w = \sum_{k=1}^c (\mathbf{X}_k - \mathbf{M}_k)(\mathbf{X}_k - \mathbf{M}_k)^T, \quad (2.11)$$

$$\mathbf{S}_b = \sum_{k=1}^c (\mathbf{M}_k - \mathbf{M})(\mathbf{M}_k - \mathbf{M})^T. \quad (2.12)$$

$\mathbf{S}_w$ ,  $\mathbf{S}_b$  を用いてフィッシャー比

$$J(\Phi) = \frac{|\Phi \mathbf{S}_b \Phi^T|}{|\Phi \mathbf{S}_w \Phi^T|} \quad (2.13)$$

を最大にする  $\Phi$  を求める. これは, 以下の一般化固有値問題を解くことで求められる.

$$\mathbf{S}_b \Phi = \mathbf{S}_w \Phi \Lambda. \quad (2.14)$$

固有値を大きい順にソートし, 各固有値に対応する  $d$  個の固有ベクトルを用いて, 以下の式により原特徴ベクトルを  $d$  次元まで削減する ( $d \leq n$ ). ここで  $\mathbf{x}$  は原特徴ベクトル,  $\mathbf{y}$  は次元削減後の特徴ベクトルである.

$$\mathbf{y} = \Phi^T \mathbf{x}. \quad (2.15)$$

## 第 3 章

# 顔画像の照合と認識

この章では、顔画像の照合、認識の処理の流れと各処理の詳細について述べる。  
本研究では、以下の手順により顔画像の照合、認識を行う。

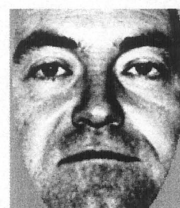
1. 前処理として、入力画像から顔領域を切り出し、輝度を正規化した画像を生成する。
2. 前処理後の画像から特徴を抽出し、特徴ベクトルを生成する。
3. 得られた特徴ベクトルを特徴選択により次元削減する。
4. 特徴選択で得られた特徴ベクトルをもとに、入力画像と複数個人の参照用画像それぞれの特徴間の距離を計算する。
5. 距離値をもとに、入力画像の人物を事前に登録されている部内者か、登録されていない部外者かに分類する（照合）。部内者であればどの人物であるかを特定する（認識）。

### 3.1 前処理

人物の髪型、衣服などは変化しやすく、それらを含んだ画像を用いて個人を識別することは望ましくない。そのため、図 3.1 (a) に示されるような原画像から髪の毛、衣服を除いた顔領域を切り出す。この処理は顔検出と呼ばれる。切り出しには原画像中の人物の両目、鼻、口の座標を与える必要がある。本研究ではこれらの座標は事前に取得され、顔検出が正しく行われたと仮定して処理を行う。切り出された画像の輝度と大きさを正規化し、前処理後の画像とする。前処理後の画像の例を図 3.1 (b) に示す。前処理後の画像の大きさは縦、横それぞれ 150 画素、130 画素である。



(a) 原画像



(b) 前処理後の画像

図 3.1: 本研究で用いる画像

## 3.2 特徴抽出

特徴抽出処理では、顔画像の照合、認識に用いる特徴ベクトルを前処理後の画像から抽出する。本研究では 2.1 で述べた濃度こう配特徴 (gradient features) と、画像の濃度値をそのまま要素とする濃淡特徴 (grayscale features) のうちいずれかの特徴量を抽出し、それぞれを用いた場合の照合、認識性能を比較する。

濃淡特徴を抽出して得られる特徴ベクトルの次元数は  $150 \times 130 = 19500$  である。

## 3.3 特徴選択 (次元削減)

特徴抽出で得られた特徴ベクトルに対して特徴選択により次元削減を行う。これにより、分類に有効であると考えられる特徴が選択され、精度の向上が期待できる。特徴選択に必要なパラメータは、学習用画像から抽出された特徴ベクトルを用いて算出される。特徴選択は 2.2 で述べた 2 手法を用いて、PCA のみを適用する場合と、PCA と LDA を組み合わせて適用する場合のうちいずれかにより特徴選択を行い、それぞれを適用した場合の照合、認識性能を比較する。

### 3.4 距離計算

特徴選択により得られた特徴ベクトルをもとに，入力画像から抽出された特徴ベクトル  $\mathbf{q} = (q_1, q_2, \dots, q_W)^T$  と参照用画像データベース中の人物  $j$  に対応する特徴ベクトル  $\mathbf{p}^{(j)} = (p_1, p_2, \dots, p_W)^T$  との距離  $D(\mathbf{q}, \mathbf{p}^{(j)})$  を計算する．距離関数にはユークリッド距離を使用し，距離値は以下の式で計算できる．

$$D(\mathbf{q}, \mathbf{p}^{(j)}) = \sqrt{\sum_i (q_i - p_i^{(j)})^2}. \quad (3.1)$$

### 3.5 分類

本研究における分類は，入力画像中の人物が事前に登録されていない部外者 ( $\omega_{\text{outsider}}$ ) であるか，事前に登録された部内者のうちどの人物であるか ( $\omega_j$ ) を決定する．

距離計算により得られた，入力画像から抽出された特徴ベクトル  $\mathbf{q}$  と参照用画像データベース中の人物  $j$  に対応する特徴ベクトル  $\mathbf{p}^{(j)}$  との距離  $D(\mathbf{q}, \mathbf{p}^{(j)})$  の  $i$  に関する最小値を  $D_{\min}$  とする．この  $D_{\min}$  は，入力画像中の人物が部内者の場合，本人との距離値であることが期待される．一方，入力画像中の人物が部外者の場合，得られた距離値は全て他人との距離値であるため部内者の場合に比べて大きくなることが考えられる．よって  $D_{\min}$  の値としきい値  $T$  を用いて次式により部外者を棄却し，部内者の人物をどの人物であるか特定する．

$$\mathbf{q} \in \begin{cases} \omega_{\text{outsider}}, & (D_{\min} \geq T) \\ \omega_j, & (D_{\min} < T) \end{cases}. \quad (3.2)$$

## 第 4 章

# 実験

### 4.1 実験データ

本研究では、顔画像データベースとして、The Facial Recognition Technology (FERET) データベース [3] を用いる。FERET データベースは、顔画像認識アルゴリズムの評価を目的に収集されたデータベースである。FERET データベースは、1000 人を超える個人の顔を収集した大規模なデータベースである点が特徴で、顔画像認識手法の評価では標準的に用いられている。本研究では、FERET データベースに含まれる正面顔画像のうち、fa, fb, fc, dup1 の 4 グループを使用する。fa は基準の画像であり、fb, fc, dup1 はそれぞれ、表情変化、光源変化、経年変化のある画像である。それぞれのグループに含まれる正面顔画像の例を図 4.1 に示す。

入力として与えられた評価用画像との比較を行うための参照用テンプレートには fa の画像を用いる。fa の画像は、各人物につき 1 枚ずつ存在する。評価用画像には、fb, fc, dup1 の画像をそれぞれ独立に用いる。また、参照用テンプレート、評価用画像とは独立して学習データを作成する。学習データには、fa, fb, fc, dup1 の画像が含まれる。学習データ、評価用画像の詳細はそれぞれ、4.1.1, 4.1.2 で述べる。

#### 4.1.1 学習データ

学習データは、次元削減に必要となるパラメータを算出するために用いられる。学習データに含まれる各人物 1 人あたりの画像数を変化させた場合に、照合、認識性能にどのような変化があるかを調べた。学習データに含まれる人数は 125 人で固定とし、各人物 1 人あたりの平均画像数が 1 枚、2 枚、3 枚、4 枚の 4 種類の学習データをデータ作成して、それぞれの学習データを用いた場合の照合、認識性能を比較した。また、学習データに含まれる人物は評価用画像には含まれない。そのため、評価対象の人物とは全く別の人物を学習

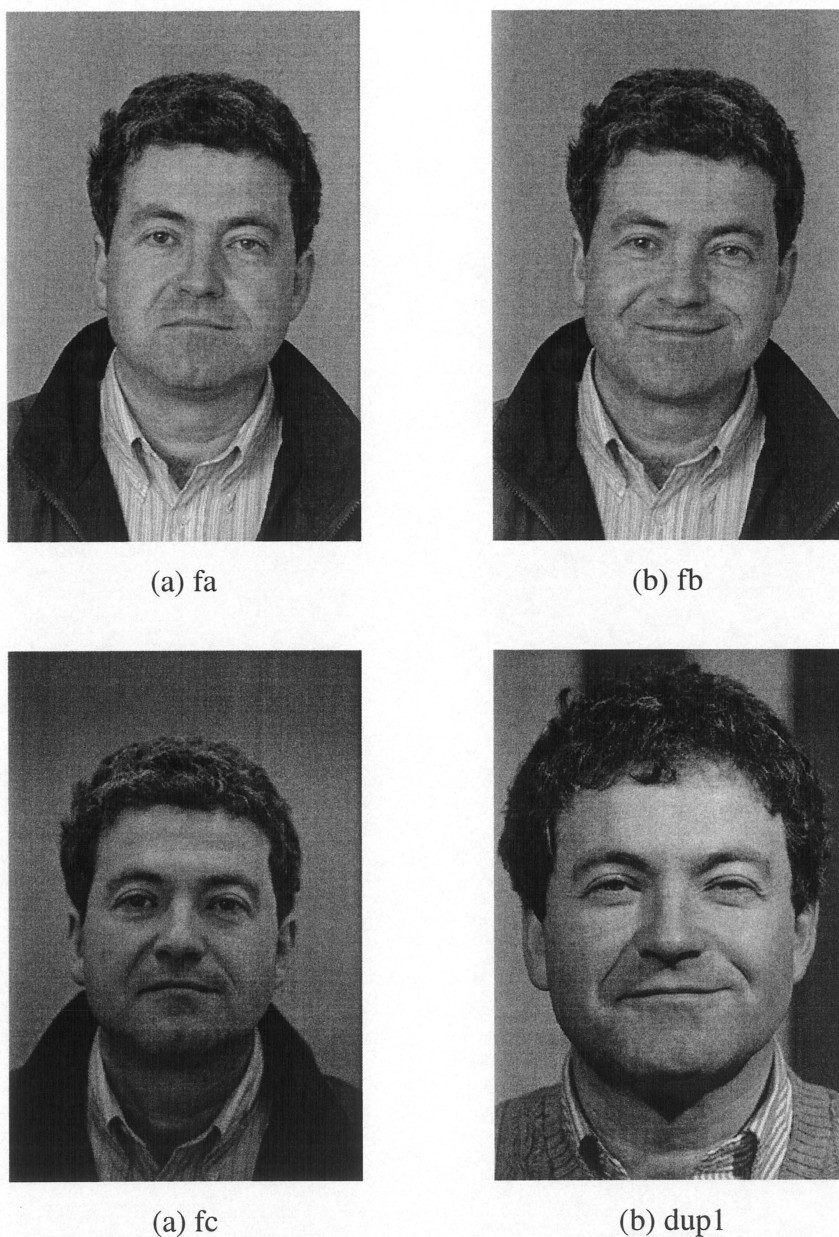


図 4.1: 本研究で用いる画像の例

している。学習データの内訳を表 4.1 に示す。

#### 4.1.2 評価用画像

評価用画像には、fb, fc, dup1 の 3 グループを用いる。性能評価のため、評価用画像の各グループを、人物を基準として 2 つのクラス (Class1, Class2) に分割する。実験では、それぞれのクラスに部内者、部外者を割り当て、一方を部内者として扱う場合、もう一方を部外者として扱い、相互検証によって評価を行う。評価用画像の内訳を表 4.2 に示す。



表 4.1: 学習データの内訳

| mean number<br>per person | fa  | fb  | fc | dup1 | number of<br>person | total |
|---------------------------|-----|-----|----|------|---------------------|-------|
| 1                         | 125 | 1   |    |      | 125                 | 126   |
| 2                         | 125 | 125 |    |      | 125                 | 250   |
| 3                         | 125 | 125 | 60 | 65   | 125                 | 375   |
| 4                         | 125 | 125 | 60 | 191  | 125                 | 501   |

表 4.2: 評価用画像の内訳

| Group name | Class1 | Class2 | Total |
|------------|--------|--------|-------|
| fb         | 535    | 535    | 1070  |
| fc         | 67     | 67     | 134   |
| dup1       | 236    | 243    | 479   |

## 4.2 評価方法

本研究では、以下の 2 つの場合について性能を評価した。

1. 入力画像中の人物が、部内者か部外者かを決定する 2 クラス分類（照合）
2. 部内者についてどの人物であるかを特定する認識

照合においては、入力画像中の人物の分類結果には表 4.3 中の 4 種類がある。本来は部内者の人物を部外者に誤って分類する割合（False Rejection Rate : FRR）と本来は部外者の人物を部内者に誤って分類する割合（False Acceptance Rate : FAR）とが等しくなる点である等価エラー率（Equal Error Rate : EER）によって評価する。式 3.2 のしきい値  $T$  を変動させて FRR, FAR を求める。FRR, FAR はそれぞれ表 4.3 中の値を用いて以下の式で定義される。

$$\text{FRR} = \frac{b}{a+b} \quad (4.1)$$

$$\text{FAR} = \frac{c}{c+d} \quad (4.2)$$

認識においては、認識エラー率（Recognition Error Rate）により評価する。認識エラー率は、部内者の画像の総数に対する別の人物に誤認識した数の割合で定義する。以下のよう認識成功率を 1 から引くことで求められる。

表 4.3: 分類の種類 ( $a, b, c, d$  は画像数)

|    |     | 分類結果 |     |
|----|-----|------|-----|
|    |     | 部内者  | 部外者 |
| 正解 | 部内者 | $a$  | $b$ |
|    | 部外者 | $c$  | $d$ |

$$\text{RecognitionErrorRate} = 1 - \frac{\text{認識成功数}}{a + b} \quad (4.3)$$

## 4.3 実験

顔画像の照合、認識における濃度こう配特徴の有効性を実証するために評価実験を行った。濃度こう配特徴、濃淡特徴それぞれを用いた場合の照合、認識性能を比較した。それぞれの特徴に対し、次元削減手法は PCA のみの場合、PCA と LDA を組み合わせた場合の 2 種類を適用した。学習データにおける 1 人あたりの画像数を変化させ、照合、認識性能に与える影響を調べた。

## 4.4 結果と考察

学習データにおける 1 人あたりの平均画像数を変化させ、照合した結果を図 4.2 に示す。

特徴量に濃度こう配特徴を用い、特徴ベクトルの次元削減に PCA と LDA とを組み合わせ適用した場合のみ、学習データにおける 1 人あたりの画像数を増加させたときの EER の低下が確認できる。学習に 1 人あたり平均 4 枚の画像を用いた場合と、1 人あたり 1 枚の画像を用いた場合とを比較すると、EER が 28.64% から 15.93% まで低下した。しかし、特徴ベクトルの次元削減に PCA のみを適用した場合は EER の低下は見られなかった。この理由は、特徴ベクトルの次元削減に LDA を適用したことで個人内の顔の変動を正規化でき、個人間の分離度の算出に有効に働いたためであると考えられる。

1 人あたりの画像数が平均 4 枚の学習データを用いて、次元削減に LDA を適用し、濃淡特徴を用いた場合と濃度こう配特徴を用いた場合とを比較するために、式 (3.2) の  $T$  の値を変動させたときの各撮影条件における検出エラートレードオフ曲線 (Detection Error Trade-off curve : DET curve) を作成した。このとき、学習データには 1 人あたりの画像数が平均 4 枚のものを使用し、特徴ベクトルの次元削減には PCA と LDA を組み合わせて適用した。DET curve を図 4.3 に示す。どの撮影条件においても、濃淡特徴よりも濃度こう配特徴の方が照合性能が高いことが確認できる。Total では EER が 27.88% から 15.93% ま

で低下した。fb では 21.08% から 10.36% まで、fc では 29.44% から 18.75% まで、dup1 では 34.76% から 10.72% までそれぞれ EER が低下した。濃度こう配特徴が濃淡特徴よりも優れた照合性能を示した理由は、(1) 顔画像を小領域に分割し、各小領域内で特徴を抽出しているため、顔の位置ずれや表情変化などによる目、口などの顔部品の形状の変化を吸収できること、(2) 局所的な濃度値の変化量を特徴として抽出しているため、大域的な光源の変化に対する特徴量の変動が小さいことであると考えられる。

学習データにおける 1 人あたりの平均画像数を変化させたときの認識性能の結果を図 4.4 に示す。学習データにおける 1 人あたりの画像数を増加させたとき、全ての場合で認識エラー率の低下が確認できる。特に、濃度こう配特徴を用いて特徴ベクトルの次元削減に LDA を適用した場合は、認識エラー率の低下が大きい。学習に 1 人あたり平均 4 枚の画像を用いた場合、1 人あたり 1 枚の画像を用いた場合と比較すると、認識エラー率が 48.11% から 20.27% まで低下した。また、学習に 1 人あたり平均 4 枚の画像を用いた場合に、特徴ベクトルの次元削減に LDA を適用し、濃淡特徴を用いた場合と濃度こう配特徴を用いた場合とを比較した。濃度こう配特徴を用いた方が Total の認識エラー率が 34.37% から 20.27% まで低下した。fb では 21.31% から 11.4% まで、fc では 29.85% から 23.88% まで、dup1 では 64.97% から 39.21% までそれぞれ認識エラー率が低下した。

また、学習データに含まれる人物は、評価用画像には含まれない。つまり別の人物の顔画像を用いて学習を行っている。別の人物の顔画像を用いて学習しても照合、認識性能が向上することがわかった。この理由は、撮影条件の変化による顔の変化のしかたは、異なる人物の間で同じであるためである。

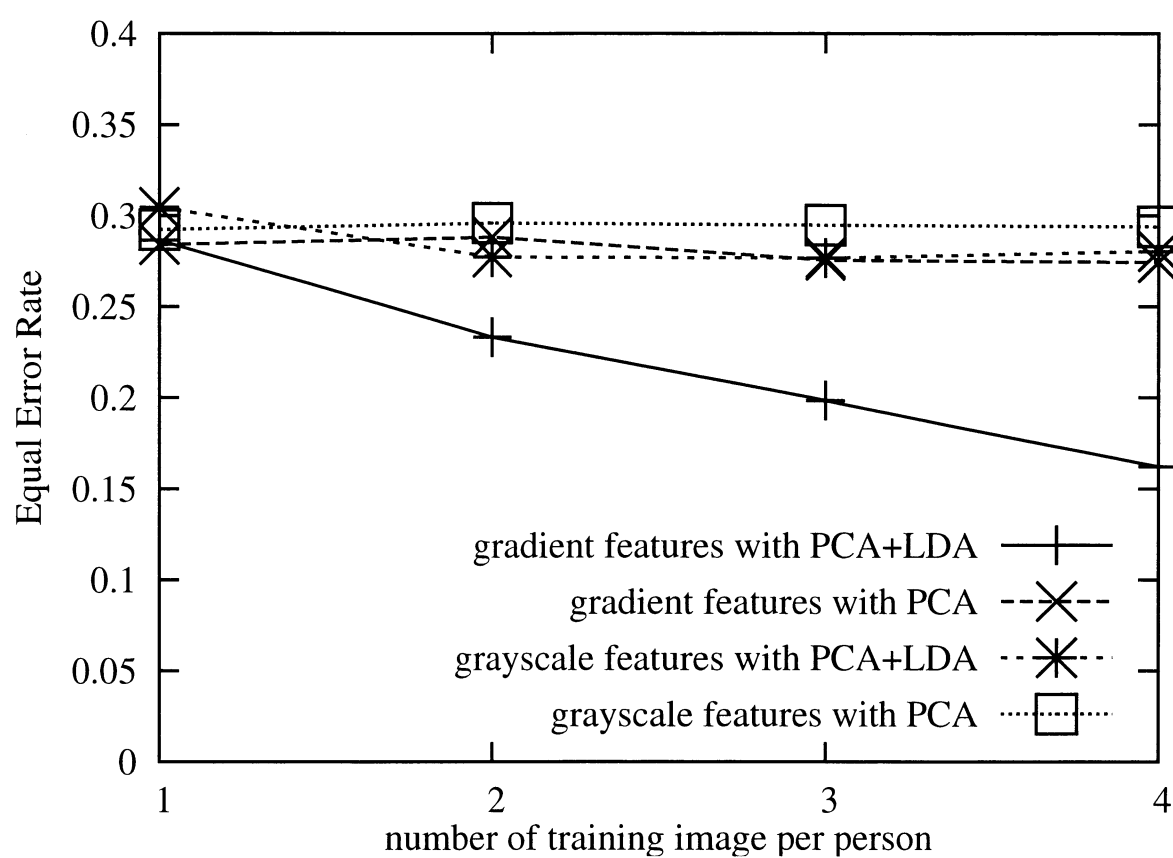


図 4.2: 学習データにおける 1 人あたりの画像数の変化に伴う照合性能の変化

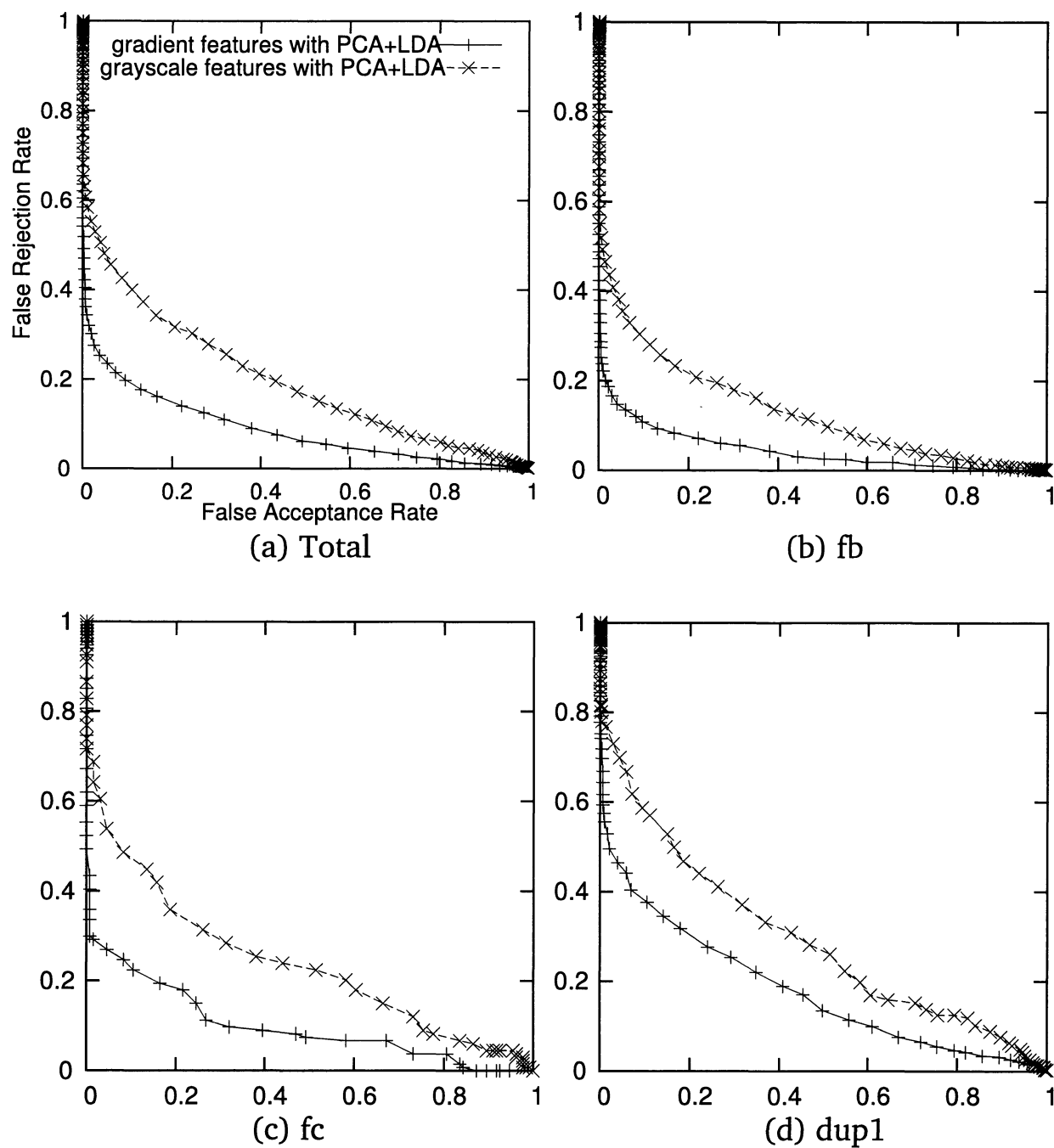


図 4.3: 各撮影条件における照合の DET curve

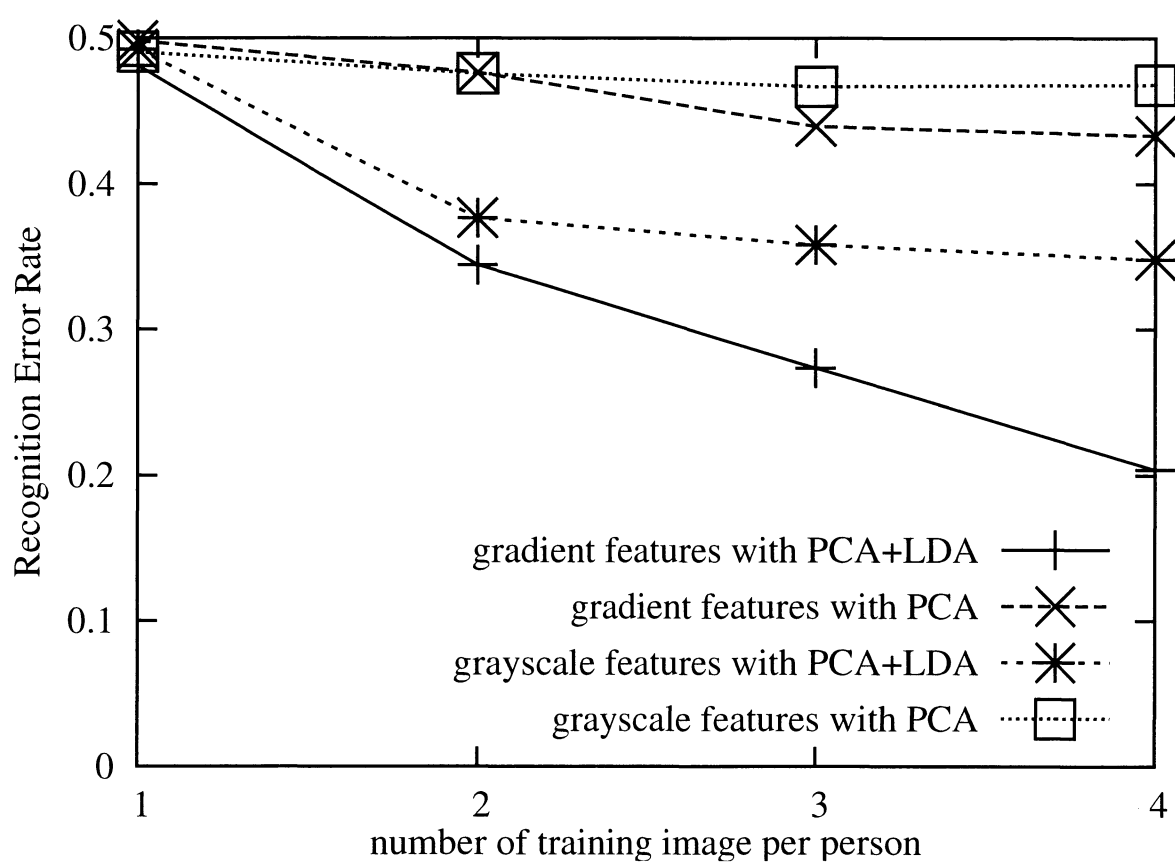


図 4.4: 学習データにおける 1 人あたりの画像数の変化に伴う認識性能の変化

## 第 5 章

# 結言

### 5.1 まとめ

本研究では，正面顔画像を用いた顔画像の照合，認識において，顔の見かけを濃度こう配特徴で表現する手法を提案し，その有効性を検証した．また，学習データに含まれる 1 人あたりの画像数と照合性能との関係を調べた．評価用画像の人物は学習データには含まれておらず，別の人物を用いて学習を行った．その結果，学習データにおける 1 人あたりの画像数を増加させたとき，特徴量に濃度こう配特徴を用い，次元削減に PCA と LDA を組み合わせて適用した場合に EER が低下していくことがわかった．学習に 1 人あたり平均 4 枚の画像を用いた場合，1 人あたり 1 枚の画像を用いた場合と比較すると，EER が 28.64%から 15.93%まで低下した．また，学習に 1 人あたり平均 4 枚の画像を用いた場合に，次元削減に LDA を適用し，濃淡特徴を用いた場合と濃度こう配特徴を用いた場合との照合性能を比較した．Total では EER が 27.88%から 15.93%まで低下した．表情変化では 21.08%から 10.36%まで，光源変化では 29.44%から 18.75%まで，経年変化では 34.76%から 26.02%まで EER がそれぞれ低下した．

次に，学習データに含まれる 1 人あたりの画像数と認識性能との関係を調べた．学習データにおける 1 人あたりの画像数を増加させたとき，全ての場合で認識エラー率の低下が確認できた．特に，濃度こう配特徴を用いて特徴ベクトルの次元削減に LDA を適用した場合は，認識エラー率の低下が大きいことが示された．学習に 1 人あたり平均 4 枚の画像を用いた場合，1 人あたり 1 枚の画像を用いた場合と比較すると，認識エラー率が 48.11%から 20.27%まで低下した．また，学習に 1 人あたり平均 4 枚の画像を用いた場合に，特徴ベクトルの次元削減に LDA を適用し，濃淡特徴を用いた場合と濃度こう配特徴を用いた場合との認識性能を比較した．濃度こう配特徴を用いた方が Total の認識エラー率が 34.37%から 20.27%まで低下した．fb では 21.31%から 11.4%まで，fc では 29.85%から 23.88%まで，dup1 では 64.97%から 39.21%まで認識エラー率がそれぞれ低

下した.

これらの実験結果より, 以下のことがわかった.

1. 顔画像の照合, 認識において, 特徴量に濃度こう配特徴を使用し, 次元削減に PCA と LDA を組み合わせて適用した場合, 学習データの 1 人あたりの画像数を増加させると照合, 認識の性能が向上する.
2. 評価用画像の人物とは別の人物の顔画像を学習に使用しても, 顔の見かけの変動を吸収でき, 照合, 認識の性能に良い影響を与える.

## 5.2 今後の課題

今後の課題としては, 実環境下での顔の照合, 認識を全自動で安定して行うために, 顔の照合, 認識の性能のさらなる改善が必要である. 学習データにおける 1 人あたりの画像数を更に増加させて実験を行うことや, 今回使用したユークリッド距離以外の距離関数を検討することなどが必要であると考えられる. また HOG 特徴, SIFT, SURF などの他の濃度こう配を用いた特徴量との性能を比較することも必要である.

本研究では, 入力画像の人物を識別するための ID 番号や名前などのラベルは付与されていない. そのため, 入力画像と, 事前に登録されている複数個人の全ての参照用テンプレートとを比較する 1 対 N の照合を行う. 入力画像にラベルを付与し, 同一のラベルが付与された参照用テンプレートとを比較する 1 対 1 の照合の評価を行いたい.

また, 顔の照合, 認識を全自動で行うには前処理として顔検出を行う必要がある. 本研究では, この顔検出の処理は正しく行われたと仮定して照合, 認識の評価をしているので, 顔検出部と統合した全自動の顔の照合, 認識の評価も行いたい.



## 付録 A

# 一般化固有値問題から標準固有値問題への変換

LDA を適用するには以下に示す一般化固有値問題を解き，固有ベクトル行列  $\Phi$  を求める必要がある．

$$S_b \Phi = S_w \Phi \Lambda. \quad (\text{A.1})$$

しかし，このアプローチは常に効果的で頑健な方法とはいえない．そのため，式 (A.1) の一般化固有値問題を以下の形式で表される標準固有値問題へと変換し，式 (A.1) の固有ベクトル行列  $\Phi$  を求める．一般化固有値問題を標準固有値問題へと変換する方法を説明する．

標準固有値問題

$$S_w \Phi_w = \Phi_w \Lambda_w \quad (\text{A.2})$$

を解く．ここで，

$$\Lambda_w = \begin{pmatrix} \lambda_{w1} & \cdots & \cdots & 0 \\ \vdots & \lambda_{w2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_{wn} \end{pmatrix} \quad (\text{A.3})$$

である．また，

$$\Lambda'_w = \begin{pmatrix} \lambda_{w1}^{\frac{1}{2}} & \cdots & \cdots & 0 \\ \vdots & \lambda_{w2}^{\frac{1}{2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_{wn}^{\frac{1}{2}} \end{pmatrix} \quad (\text{A.4})$$

とすると,  $S_w$  は次式で表すことができる.

$$S_w = \Phi_w \Lambda'_w \Lambda'_w \Phi_w^T. \quad (\text{A.5})$$

すると, 式 (A.1) は次式で表せる.

$$S_b \Phi = \Phi_w \Lambda'_w \Lambda'_w \Phi_w^T \Phi \Lambda. \quad (\text{A.6})$$

$\Lambda_w^i \Phi_w^T$  を両辺にかけると次式に変形できる.

$$\Lambda_w^i \Phi_w^T S_b \Phi = \Lambda'_w \Phi_w^T \Phi \Lambda. \quad (\text{A.7})$$

ここで,

$$\Lambda_w^i = \begin{pmatrix} \lambda_{w1}^{-\frac{1}{2}} & \cdots & \cdots & 0 \\ \vdots & \lambda_{w2}^{-\frac{1}{2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_{wn}^{-\frac{1}{2}} \end{pmatrix} \quad (\text{A.8})$$

である.

$V = \Lambda_w \Phi_w^T \Phi$  とおき, 次式に恒等変換する.

$$\Lambda_w^i \Phi_w^T S_b \Phi_w \Lambda_w^i \Lambda_w \Phi_w^T \Phi = V \Lambda. \quad (\text{A.9})$$

$G = \Lambda_w^i \Phi_w^T$ ,  $G^T = \Phi_w \Lambda_w^i$  とおき, 次式に変形する.

$$G S_b G^T V = V \Lambda. \quad (\text{A.10})$$

$N = G S_b G^T$  とおき, 以下の対称固有値問題を解く.

$$NV = V\Lambda. \quad (\text{A.11})$$

次式により，式 (A.1) の固有ベクトル行列  $\Phi$  を求める．

$$\Phi = \Phi_w \Lambda_w^i V \quad (\text{A.12})$$

## 付録 B

# 研究データ

研究に使用したデータは deneb (IP アドレス : 133.67.33.11) のローカル

home/komaki/Thesis/csuFaceIdEval.5.0.arg/

に存在する。ただし、濃度こう配特徴を抽出するプログラムソースファイルは以下のディレクトリに存在する。

home/komaki/Thesis/MieUFaceRecog/

csuFaceIdEval.5.0.arg/ 以下のディレクトリを次に示す。またそれぞれのディレクトリについて概説する。詳細は、各ディレクトリにある README を参照されたい。

```
csuFaceIdEval.5.0.arg/    #研究に関するデータが存在するディレクトリ。
|                          #ソースファイルのコンパイルやプログラムの実行
|                          #などはここで行う。
|
|--bin/
|   '--x86_64/            #実行ファイルを保存するディレクトリ。
|
|--lib/
|   '--x86_64/            #オブジェクトファイルを保存するディレクトリ。
|
|--distances/            #distance ファイルが存在するディレクトリ。
|
|--data/
|   '--FERET/
|       |--gradient_sfi/  #濃度こう配特徴を抽出して生成した
```

```

|           | #原特徴ベクトルが存在するディレクトリ.
|           |
|           | --normSep2002sfi/ #濃淡特徴を抽出して生成した
|           | #原特徴ベクトルが存在するディレクトリ.
|           |
|           | --normSep2002pgm/ #前処理後の画像が存在するディレクトリ.
|           |
|           | --source/ #原画像が存在するディレクトリ.
|           |
|--src/ #C のソースファイルが存在するディレクトリ.
|
|--scripts/ #シェルスクリプトが存在するディレクトリ.
|
|--galleryVectors/ #次元削減後の特徴ベクトルが存在する
| #ディレクトリ.
|
|--imagelists/ #参照用データ, 学習データ, テストデータの
| #リストファイルが存在するディレクトリ.
|
|--Results/ #実験結果が保存されたディレクトリ.
|
|--trainingParam/ #学習データの特徴ベクトルを用いて算出した,
| #次元削減に必要なパラメータを保存した
| #ディレクトリ.
|
|--GradientImage/ #カラーの濃度こう配画像を生成するプログラムが
| #存在するディレクトリ.

```

プログラムソースファイルのコンパイルはディレクトリ `csuFaceIdEval.5.0.arg/` で `make` コマンドでできる.

プログラムの実行はディレクトリ `csuFaceIdEval.5.0.arg/` で以下のようにしてスクリプトを実行する.

```
./scripts/identification_gradient_PCA_LDA_ED.sh
```

# 謝辞

本研究に関し、数多くのアドバイスを下さり、指導していただいた木村文隆教授、若林哲史准教授、大山航助教、三宅康二名誉教授に深く感謝いたします。また日頃いろいろとお世話になった田中みゆき事務官に感謝します。

また、楽しい雰囲気を作って下さったヒューマンインターフェース研究室の皆さん、4年間の大学生活を共に過ごしてきた皆さんに感謝します。

最後になりましたが、私の大学生活を支えてくれた両親に今一度の感謝を表して、本論文の結びといたします。

## Acknowledgements

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office.

## 参考文献

- [1] W.Zhao, R.Chellappa, P.J.Phillip and A.Rosenfeld : Face Recognition : A Literature Survey, ACM Computin Surveys, Vol.35, No.4, pp.399-458, 2003.
- [2] Ross Beveridge, David Bolme, Marcio Teixeira and Bruce Draper 2003 : The CSU Face Identification Evaluation System Version 5.0 <http://www.cs.colostate.edu/evalfacerec/algorithms5.html>
- [3] P.J.Phillips, H.Wechsler, J.Huang and P.Rauss : The FERET database and evaluation procedure for face recognition algorithms, Image and Vision Computing J, Vol.16, No.5, pp.-295-306, 1998.
- [4] Kirby M. and Sirovich L. : Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Trans. Patt. Anal. Mach. Intell.12. 1990.
- [5] Sirovich L. and Kirby M. : Low-dimensional procedure for the characterization of human face. J. Opt. Soc. Am. 4, 519-524. 1987.
- [6] Ruderman D.L. : The statistics of natural images. Netw. Comput. Neural Syst. 5, 598-605. 1994.
- [7] Penev P. and Atick J. : Local feature analysis:A general statistical theory for object representation. Netw. Comput. Neural Syst. 7, 477-500, 1996.
- [8] Zhao W. Robust Image Based 3D Face Recognition, Ph.D. dissertation. University of Maryland, College Park, MD. 1999.
- [9] Turk M. and Pentland A. : Eigenfaces for recognition. J. Cogn. Neurosci. 3, 72-86. 1991.
- [10] Moghaddam B. Nastar C. and Pentland A. : A Bayesian similarity measure for direct image matching. In Proceedings, International Conference on Pattern Recognition, 1997.
- [11] Fisher R.A. : The statistical utilization of multiple measurements. Ann. Eugen. 8, 376-386, 1938.
- [12] Belhumeur P.N., Hespanha J.P., and Kriegman D.J. : Eigenfaces vs. Fisherfaces: Recog-

- inition using class specific linear projection. *IEEE Trans. Patt. Anal. Mach. Intell.* 19, 711-720, 1997.
- [13] Etemad K. and Chellappa R. : Discriminant analysis for recognition of human face images. *J. Opt. Soc. Am. A* 14, 1724-1733, 1997.
- [14] Swets D.L. and Weng J. : Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Patt. Anal. Mach. Intell.* 20, 39-51, 1996.
- [15] Zhao W., Chellappa R. and Krishnaswamy A. : Discriminant analysis of principal components for face recognition . In *Proceedings, International Conference on Automatic Face and Gesture Recognition*. 336-341, 1998.
- [16] Fukunaga K. : *Statistical Pattern Recognition*, Academic Press, New York, NY, 1989.
- [17] Liu C. and Wechsler H. : Evolutionary pursuit and its application to face recognition. *IEEE Trans. Patt. Anal. Mach. Intell.* 22, 570-582, 2000.
- [18] Vapnik V.N. : *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, NY, 1995.
- [19] Bartlett M.S., Lades H.M. and Sejnowski T. : Independent component representation for face recognition. In *Proceedings, SPIE Symposium on Electronic Imaging : Science and Technology*. 528-539, 1998.
- [20] Lin S.H., Kung S.Y. and Lin L.J. : Face recognition/detection by probabilistic decision-based neural network. *IEEE Trans. Neural Netw.* 8, 114-132, 1997
- [21] Kung S.Y. and Taur J.S. : Decision-based neural networks with signal/image classification applications. *IEEE Trans. Neural Netw.* 6, 170-181, 1995.
- [22] Li S.Z. and Lu J. : Face recognition using the nearest feature line method. *IEEE Trans. Neural Netw.* 10, 439-443, 1999.
- [23] Phillips P.J. : Support vector machines applied to face recognition. *Adv. Neural Inform. Process. Syst.* 11, 803-809, 1998.
- [24] Kanade T. : *Computer recognition of human faces*. Birkhauser, Basel, Switzerland, and Stuttgart, Germany, 1973.
- [25] Kelly M.D. : *Visual identification of people by computer*. Tech. rep. AI-130, Stanford AI Project, Stanford, CA, 1970.
- [26] Samaria F. and Young S. : HMM based architecture for face identification. *Image Vis. Comput.* 12, 537-583, 1994
- [27] Samaria F. : *Face recognition using hidden markov models*. Ph.D. dissertation. University of Cambridge, Cambridge, U.K., 1994.
- [28] Okada K., Steffans J., Maurer T., Hong H., Elagin E. Neven H. and Malsburg C.V.D. : The Bochum/USC Face Recognition System and how it fared in the FERET Phase I



- II Test. In Face Recognition: From Theory to Applications, H. Wechsler, P.J. Phillips, V.Bruce, F.F. Soulie and T.S. Huang, Eds. Springer-Verlag, Berlin, Germany, 186-205, 1998
- [29] Wiskott L., Fellous J.M. and VON DER Malsburg C. : Face recognition by elastic bunch graph matching. IEEE Trans. Patt. Anal. Mach. Intell. 19, 775-779, 1997
- [30] Buhmann J., Lades M. and Malsburg C.V.D : Size and distortion invariant object recognition by hierarchical graph matching. In Proceedings, International Joint Conference on Neural Networks. 411-416, 1990.
- [31] Lades M., Vorbruggen J., Buhmann J., Lange J., Malsburg C.V.D., Wurtz R. and Konen W. : Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput. 42. 300-311, 1993
- [32] Kruger N., Potzsch M. and Malsburg C.V.D. : Determination of face position and pose with a learned representation based on labelled graphs. Image Vis. Comput. 15, 665-673, 1997.
- [33] Maurer T. and Malsburg C.V.D. : Singleview based recognition of faces rotated in depth. In Proceedings, International Workshop on Automatic Face and Gesture Recognition. 176-181, 1996.
- [34] Pentland A., Moghaddam B. and Starner T. : View-based and modular eigenspaces for face recognition. In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition.
- [35] Lanitis A., Taylor C.J. and Cootes T.F. : Automatic face identification system using flexible appearance models. Image Vis. Comput. 13, 393-401, 1995.
- [36] Heisele B., Serre T. Pontil M. and Poggio T. : Component-based face detection. In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, 2001.
- [37] Huang J., Heisele B. and Blanz V. : Component-based face recognition with 3D morphable models. In Proceedings, International Conference on Audio- and Video-Based Person Authentication, 2003.
- [38] Blanz V. and Vetter T. : A Morphable model for the synthesis of 3D faces. In Proceedings, SIGGRAPH'99, 187-194, 1999.
- [39] Meng Shi, Yoshiharu Fujisawa, Tetsushi Wakabayashi and Fumitaka Kimura 2002 : Handwritten numeral recognition using gradient and curvature of gray scale image, Pattern Recognition, Vol.35, Issue 10, pp.2051-2059
- [40] Navneet Dalal, Bill Triggs : "Histograms of Oriented Gradients for Human Detection", International Conference on Computer Vision & Pattern Recognition, Vol.2, pp.886-893, 2005.

- 
- [41] David G. Lowe : “Distinctive Image Features from Scale-Invariant Keypoints”, International journal of computer vision, Vol.60, No.2, pp.91-110, 2004.
- [42] Bay H., Tuytelaars T. and Van Gool L. : “ SURF : Speeded Up Robust Features”, Lecture notes in computer science, Vol.3951, pp.404, 2006.