

修 士 論 文

円の Hough 変換と濃度こう配
ヒストグラムを用いる両目の検出

平成 23 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

伊藤 康孝

はじめに

個人および企業の情報セキュリティ強化は、現代の情報社会で重要であり、従来の個人認証よりも強力な手段が必要である。個人認証は、カードやパスワードを使用するのが一般的であり、容易である。しかし、この認証方法は、紛失や盗難、忘却などの問題を引き起こす。これと比較して、バイオメトリクス認証の一つである顔認識は、これらの問題は起こらない。また顔認識は非接触であるため、その応用例は幅広く、建物への入退室管理や Web カメラを用いたネットワークログオン、遊園地での迷子探しなど、多岐に応用できると考えられる。

高精度な顔認識を実現するために、目や鼻、口のような顔部品の検出は重要である。また、顔部品の検出は、顔検出の精度向上のためにも重要である。Qian らは、強度フィルタと K-means クラスタリングによる目の検出を提案した。しかし、精度向上のため局所特徴を用いることが今後の課題とされていた。本研究では、正面顔画像から目を検出する手法を提案する。

本手法では、濃度こう配の特徴量であるこう配強度とこう配方向とを用いる。この 2 つの特徴量と円の Hough 変換とを組み合わせ、人物画像から目の候補を検出する。そして、各目の候補の周囲の濃度こう配の局所方向ヒストグラム (HOG) を抽出し、Support Vector Machine (SVM) を用いて、候補の中から両目を決定する。この手法の利点は、円の Hough 変換を 2 次元のパラメータ空間で行っているため、3 次元の円の Hough 変換よりも効率的に半径未知の円を検出することができる。そして、画像の見かけを利用した目らしさの評価を濃度こう配ヒストグラムと SVM により行っている点である。

両目検出手法の有効性を、FERET データベースの正面顔画像 1,409 枚を用いた実験により評価した。実験の結果、提案手法は両目の候補数が 1 組の場合 97.94%、候補数が 3 組の場合は 98.94% の検出成功率で両目を検出することができた。この結果は、Qian らの検出率 91.95% を大きく上回り、提案手法の有効性と信頼性が、大規模な画像データベース上での実験により実証された。

今後の課題として、(1) その他の顔部品の検出、(2) 顔認識への応用が挙げられる。

本論文の構成は、第 1 章では研究の背景と目的、関連する先行研究について述べる。第

2 章では，顔画像処理に関する従来研究について紹介する．続いて，第 3 章では本研究における提案手法の概要とその処理の流れ，各処理の詳細についてを述べる．さらに，第 4 章では提案手法による評価実験と結果，考察について述べる．最後に，第 5 章にて本研究のまとめと今後の課題について述べる．

目次

はじめに	i
第 1 章 緒言	1
1.1 研究背景	1
1.2 研究目的	2
1.3 関連する先行研究	2
第 2 章 顔画像処理に関する従来研究	4
2.1 顔検出と顔認識, 顔部品検出	4
2.2 画像全体を用いる手法	5
2.3 局所的な特徴を用いる手法	7
2.4 画像全体と局所的な特徴の両方を用いる手法	8
第 3 章 両目検出手法	9
3.1 提案手法の概要	9
3.2 濃度こう配抽出	9
3.3 円の Hough 変換	11
3.4 目の候補の検出	12
3.5 半径の推定	13
3.6 特徴抽出	13
3.7 目の組の作成	15
3.8 対称性の評価	17
第 4 章 評価実験	18
4.1 FERET データベース	18
4.2 Support Vector Machine(SVM)	18
4.3 学習サンプル	19
4.4 評価方法	20

4.5	性能評価実験	21
4.6	結果と考察	21
4.7	比較実験	25
4.8	結果と考察	25
第 5 章	結言	28
5.1	まとめ	28
5.2	今後の課題	28
付録 A	プログラムソースリスト	29
A.1	コンパイル方法	30
A.2	実行方法	30
付録 B	研究データ	31
付録 C	発表資料	32
謝辞		33

第 1 章

緒言

1.1 研究背景

個人および企業の情報セキュリティ強化は、現代の情報社会で重要であり、従来の個人認証よりも強力な手段が必要である。現在、主流となっている個人認証は、カードやパスワードを使用するのが一般的であり、容易である。しかし、これらの認証方法は、紛失や盗難、忘却といった問題を引き起こす。これと比較して、バイOMETRICS認証の一つである顔認識は、固有の不変な物理的特徴を利用するため、これらの問題は起こらない。しかし、不変な物理的特徴にも関わらず、人間の顔の機械認識は、様々な要因の影響を非常に受けやすい。例えば、顔認識は表情や照明条件、隠れ、経年変化による影響を受ける。もし、これらの要因に対して頑健な顔認識技術が確立されれば、信頼性の高いバイOMETRICS認証が実現し、その応用先は幅広い。例えば、建物への入退室管理や Web カメラを用いたネットワークログオン、遊園地での迷子探し、銀行 ATM、不審者のチェックなど様々である。

正確な顔認識を実現するために、様々な研究が活発に行われている [1, 2, 3, 4]。顔の正中線抽出 [5, 6] は、完全な自動顔認識における計算コストを削減するための有望な技術の一つであり、高い抽出精度を達成することが示された。虹彩の検出 [7] や高速複合分離度フィルタによる円形オブジェクトの検出 [8] も研究されている。複合分離度フィルタは、垂直方向と水平方向、長方形の対角方向 2 つの分離度フィルタを組み合わせたものである。Active Appearance Model(AAM) 探索による目領域構造の自動抽出 [9] も行われている。これらは、2 次元画像データを対象としているが、顔認識の研究は、3 次元データに対しても行われている [3]。

1.2 研究目的

信頼性のある高精度な顔認識を実現するために、両目や鼻、口のような顔部品の検出は重要である。顔部品の検出は、顔認識の精度を向上させるためにも重要である。本研究では、正面顔画像から両目を検出することを目的とする。もし事前に両目の位置がわかれば、鼻や口の位置を決定することが容易になり、両目や鼻、口の位置関係を利用することで正確に顔を検出することが可能となる。このようにすることによって、画像中のどこに顔が存在するのか、正確な位置を把握することができる。さらに、両目の位置を知ることによって、顔の中心線である正中線を得ることができる。顔の正中線を知っていれば、顔の傾きの影響を受けにくい頑健な顔認識が可能となる。このように、両目は顔の大きさの正規化や顔位置の決定、顔画像平面内の回転補正などができる重要な要素であり、両目や鼻、口のような顔部品の検出は、顔認識の精度を向上させる重要なステップである。

1.3 関連する先行研究

Qian ら [10] は、強度フィルタと K-means クラスタリングによる目の検出を提案した。この手法は大きく分けて、目の位置を大まかに特定する第 1 段階と、目の中心の正確な位置を特定する第 2 段階の 2 つに分けられる。まず第 1 段階では、OpenCV に実装されている Viola-Jones の顔検出器を用いて顔を検出する。次に、検出した顔画像に対して肌の階調レベルの範囲を調べるためにヒストグラム解析をし、画像強調をグレースケール画像に対して行う。これにより顔画像のコントラストを向上させる。コントラスト強調した画像をガボールウェーブレットを用いて参照画像を作り、K-means クラスタリングアルゴリズムを用いて参照画像を解析することで、大まかに目の位置を特定する。第 2 段階では、眼球の周りの白い領域が最も強い強度を持っており、瞳孔が最も弱い強度を持っていることを利用する。第 1 段階で特定した目を含む大まかな領域内の全ての画素に対して、中心画素の値を 3×3 近傍の合計値で置き換えるフィルタを施す。続いて、中心画素を 5×5 近傍の合計値で置き換えるフィルタを施す。これらの処理によって得られた値から最小値を求める。この最小値が得られた位置を瞳孔の位置とし、目の中心位置としている。しかし、この手法は強度情報を主に使用するため、部分的な隠れがある場合に十分な性能を発揮できないとされている。したがって、精度向上のために局所特徴を用いることが今後の課題とされていた。

Song ら [11] は、組み合わせ 2 値エッジと強度情報を用いた目の検出を提案した。この手法は、スケールを変化させた高周波成分を用いてスケール毎に顔画像を再構成し、画素ごとの単純加算によって、単一スケールで得られた全ての BEIs をマージし、2 値化する。

このマルチ解像度ウェーブレット変換に基づいて、グレイスケール顔画像から 2 値エッジ画像 (BEIs) を抽出する。次に、目と鼻との間の垂直距離は、通常、目と眉との間の距離と、鼻と口との間の距離よりも大きいという点、顔には目と眉は 2 つ存在するが、鼻と口は 1 つという顔を構成するレイアウトに関する事前知識を利用し、目の領域を抽出している。そして、虹彩の強度はその周辺よりも低いということを利用し、目の位置を決定している。

Asteriadis ら [12] は、距離ベクトルマップを用いた顔特徴の検出を提案した。この手法は、まず Viola-Jones の顔検出器を用いて顔領域を検出する。次に検出された顔領域に対して Canny オペレータを施し、顔領域の距離ベクトルフィールドを算出する。候補領域の距離ベクトルフィールドと、水平方向と垂直方向とのベクトル成分マップであるテンプレートとを比較することで目の領域検出を行っている。

Active Appearance Model(AAM) 探索による目領域構造の自動抽出 [9] がある。これは、目領域の構造に関する情報をあらかじめ付与した学習サンプルから Active Appearance Model(AAM) を構築し、AAM 空間において未知入力に最も近い学習サンプルを見つけることで、その学習サンプルの構造情報から間接的に入力目領域の構造を自動取得するものである。しかし、この手法は、画像中の物体の輪郭を構成する点の座標がすべて与えられた学習サンプル群が必要であり、学習サンプルを集めるのが容易ではないという問題がある。

また、顔画像から両目を検出する手法として、仲尾らが提案する正中線の自動検出とその両目検出への応用 [5] がある。この手法は、入力画像に対してエッジ抽出、2 値化、ノイズ除去を行った画像とその鏡像画像をテンプレートとして用いる。まずエッジ点の重心によって参照点を決定し、一般化 Hough 変換を行う。一般化 Hough 変換の目的は、参照点と左右対称関係にある抽出点を検出することである。この 2 点を用いることで、正中線を決定している。次に、正中線が垂直になるように画像を回転させることで、左右の目の高さが等しく、また両目は正中線に対し線対称な位置に存在すると仮定し、右目と左目を対で探索している。さらに目領域内の濃度こう配強度の平均がしきい値以下であれば、その領域は顔部品を含まないとみなし探索の対象から外している。そして、削減した探索範囲内に対して、任意の目領域のペアを生成し、テンプレートマッチングによる評価を行うことで両目を決定している。しかし、仲尾らの両目検出手法は正中線の検出精度に依存し、しきい値の決定にも改善が必要であるとされていた。

第 2 章

顔画像処理に関する従来研究

顔画像の自動認識は，建物への入退室管理，銀行 ATM などでの個人識別，ロボットと人間との間のインタフェース，ビデオ映像検索クエリなど，様々な応用分野があり，盛んに研究されている．この研究分野の最終的な目標は，実環境下における安定した顔画像の自動認識であるが，これは計算機にとって未だに難しい課題である．

ここでは，従来の顔画像処理に関する研究 [1, 2, 3] について本研究に関連する顔検出を取り扱うものを中心に紹介し，この分野のこれまでの研究動向について概説する．

2.1 顔検出と顔認識，顔部品検出

顔を含んだ画像，または画像列 (シーケンス) は，例えば視覚に基づいた知的マンマシンインタフェースなどにおいて基本となる情報である．顔認識や顔の追跡，表情の認識など多数の顔画像処理に関する研究が報告されているが，これらのほとんどは入力画像中から適切に顔領域が切り出されていることを仮定している．

ここではまず，顔画像処理の分野における顔検出，顔認識，顔部品検出のそれぞれについて概略と相互の関係について説明する．

入力画像に複数の顔が含まれていた場合，それらの顔を発見し適切な大きさで画像から切り出す処理を顔検出と呼ぶ．顔検出の最終的な目標は，画像中の顔のサイズや向き，形状の変動に関わらず，正しく顔を検出することである．

顔認識とは，入力画像中の顔が事前に登録されているデータベースに含まれるどの人物の顔かを判定する処理である．この処理では，入力画像とデータベース内のすべての顔が比較される．多くの顔認識手法では，額からあご，左右の頬までの顔面画像を認識対象としている．これは，頭髮や衣服の変化による画像変動が認識結果に与える悪影響を低減するためである．したがって，顔認識を目的とした顔検出では，余分な領域を含まない顔面領域のみを正しく切り出す必要がある．

顔検出手法の多くは、顔の検出に重点を置いたため、顔の境界線を正確かつ厳密に決定することが困難である。後述する顔特徴に基づく顔検出で利用されているように、顔部品（両目、鼻、口）などを検出し、それらの位置情報をもとに統計的モデルなどによって顔の境界線を決定するアプローチが提案されている。顔部品検出は、そのための基本的な処理であり、表情認識や動作認識にも応用されている。

顔認識の手法は大きく次の3つに分類することができる。

- 画像全体を用いる手法
- 局所的な特徴を用いる手法
- 画像全体と局所的な特徴の両方を用いる手法

以下では、上記3つのそれぞれに属する手法について簡単に述べる。

2.2 画像全体を用いる手法

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)[13, 14] である。その他にも、線形識別分析 (Linear Discriminant Analysis : LDA) を用いた手法も提案されており、高い性能が得られることが知られている [15, 16]。

2.2.1 固有顔 (eigenfaces)

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。原画像は高い冗長性を含んでいるため、PCA の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利点がある。

機械による顔画像認識の最初の成功例は、この固有顔 [17] 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ため、ある未知の顔画像に対して、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。これにより、入力顔画像はこの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 [13]。

2.2.2 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 [18] 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．

2.2.3 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ある．

2.2.4 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，顔検出モジュール，目位置特定モジュール，顔認識モジュールの 3 つのモジュールから構成される．

その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 [13]．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 14×10 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．

2.3 局所的な特徴を用いる手法

この手法は局所的な特徴に基づく手法と局所的な見え方に基づく手法との2つに大別される。局所的な特徴に基づく手法とは、まず局所的な特徴を抽出し、見つけられた特徴点に関する特徴を抽出する。局所的な見え方に基づく手法は、サブ領域に顔画像を分割し、領域に基づいて局所的な特徴を直接抽出する。このように、この手法では目や鼻、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方の統計量などが分類器に与えられ、認識が行われる。

このグループに属する代表的な手法には、隠れマルコフモデル (Hidden Markov Model : HMM) の手法 [22]、ガボールウェーブレット分解に基づいた顔の特徴検出 [23]、ガボール特徴に基づいた手法 [24]、Self-Organizing Maps(SOM)[25]、Elastic Bunch Graph Matching(EBGM)[26] などがある。特に EBGM は最も成功したアルゴリズムの1つである。

局所的な見え方に基づく手法は、局所領域分割、特徴抽出、特徴選択、分類の4ステップから成り、各ステップにおいて様々な手法が提案されている。局所領域分割の形について、長方形窓 [27]、楕円 [28]、帯状の形 [29] などが研究されている。また大きさについて、局所領域の特徴数とロバスト性に直接的な関係があるとされている。特徴抽出については、濃淡値特徴 [27] がテクスチャ情報の損失のない最も簡単な特徴とされ、その他にもガボールウェーブレット [24]、Harr ウェーブレット [30]、フラクタル特徴などがある。

2.3.1 Elastic Bunch Graph Matching(EBGM)

EBGM は、まず画像を様々な周波数と方向を持ったガボールフィルタで畳み込み、それらの応答の集合を Jet とし、Bunch Graph と呼ばれるモデルを生成する。この Bunch Graph は、複数人のデータから生成しておく。次に、Bunch Graph と入力画像の各特徴点の Jet との間で類似度を計算し、特徴点の座標を推定している。この手法は、外観の変化に対する高いロバスト性を示している。

2.3.2 Self-Organizing Maps(SOM)

SOM は、まず各顔画像を複数個の異なる局所サブブロックに分割する。SOM ネットワークは、クラスに関係なく全ての利用可能な学習画像から得られた全てのサブブロックを用いて学習する。SOM マップを学習した後、同じ顔画像のそれぞれのサブブロックは、最近傍法により、対応する best matching units(BMUs) に写像される。位置ベクトルは、その 2D SOM 位相空間の位置を意味し、顔画像を表現できる。

2.4 画像全体と局所的な特徴の両方を用いる手法

この手法は全体的な特徴と局所的な特徴との両方を用い，Hybrid LFA[31]，Local probabilistic method[28]，フレキシブルな見え方モデルに基づく手法 [32]，modular eigenfaces[33] などがある．modular eigenfaces は，全体の固有顔と局所的な固有顔の両方を用いる．Local probabilistic method は，画像全体を用いる手法で新たな学習サンプルを生成し，局所的な特徴を用いる手法を認識部において利用している．

2.4.1 フレキシブルな見え方に基づく手法

この手法は，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．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．最後に，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．

第 3 章

両目検出手法

3.1 提案手法の概要

提案手法による両目検出の流れを図 3.1 に示す．提案手法は，大きく 2 つの特徴がある．1 つ目の特徴は，半径未知の円を検出するための 2 次元空間による円の Hough 変換である．提案する円の Hough 変換は，まず Sobel オペレータ [34] により顔画像から濃度こう配を算出し，濃度こう配の強度と方向を得る．この濃度こう配の強度と方向を用いることにより，本手法は 2 次元空間上で円を検出することができる．瞳の形状から目は円であるため，本手法を用いることで顔画像から目の候補を検出する．詳しい検出方法については後述する．

2 つ目の特徴は，濃度こう配ヒストグラム (HOG) と SVM を用いた目らしさの評価を行うことである．前述した円の Hough 変換により顔画像から多数の目の候補が得られるため，画像の見かけを利用し，目らしさの評価を行うことで多数の目の候補の中から両目を決定する．これは，まず円の Hough 変換により検出した候補点を中心に長方形窓を生成し，濃度こう配ヒストグラム (HOG) を抽出する．抽出された特徴ベクトルは，SVM によってその領域の目らしさが評価される．そして，事前に定義された条件を満たす目の組を生成し，両目の目らしさの評価値の合計によって順位付けをすることで両目を決定する．それぞれの処理について，次節以降で詳しく述べる．

3.2 濃度こう配抽出

濃度こう配とは，各画素値の変化の度合を示すものであり，2 つの特徴量を持つ．この 2 つの特徴量は，こう配強度とこう配方向である．本手法では，カーネルサイズが 3×3 の Sobel オペレータを用いて，各画素における濃度こう配の強度と方向とを算出する．Sobel オペレータとは，こう配に基づいたエッジ検出法であり，近傍領域内の濃度の和を求める

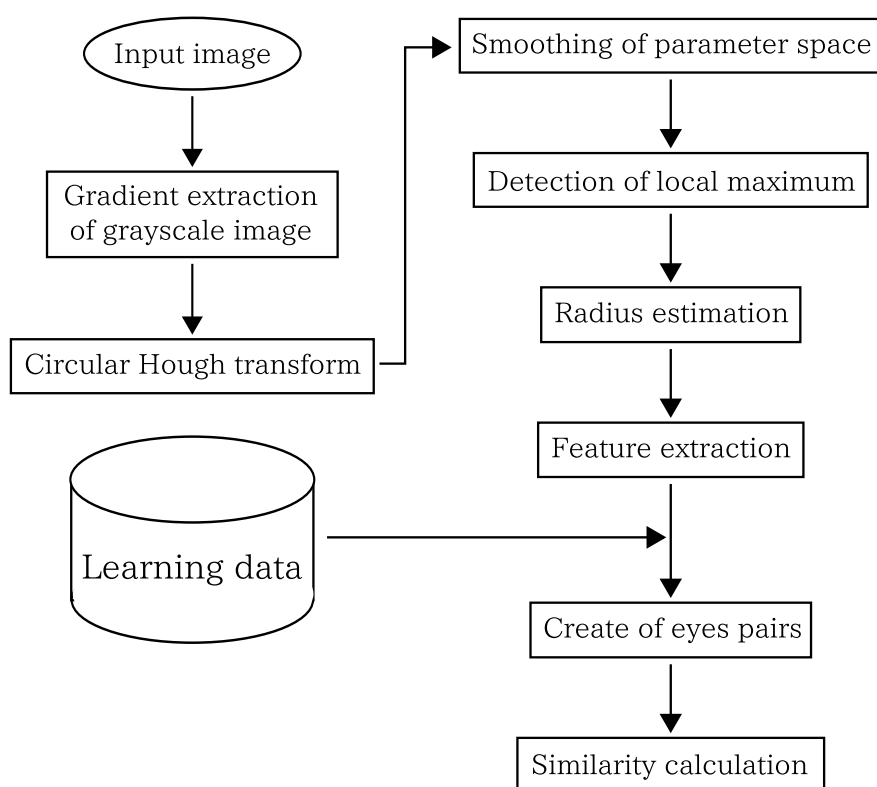


図 3.1: 両目検出処理の流れ

際に，上下左右の画素には対角線上のものより大きな重みを与えるという考えに基づいた加重マトリクスで，最もよく用いられる差分オペレータである．これは，水平方向，垂直方向それぞれの偏微分を計算することにより画像からエッジを検出することができる．

水平方向，垂直方向それぞれのカーネルを図 3.2 に示す．画素 $p(x, y)$ におけるこう配の水平方向，垂直方向の成分をそれぞれ $G_x(p)$ ， $G_y(p)$ とし，こう配強度を $|G(p)|$ ，こう配方向を $\theta(p)$ とする．ここで，こう配強度とこう配方向はそれぞれ式 3.1，3.2 で与えられる．図 3.2 のそれぞれのカーネルを適用することで，水平方向，垂直方向の成分である $G_x(p)$ ， $G_y(p)$ を取得し，式 3.1，3.2 により，こう配強度 $|G(p)|$ とこう配方向 $\theta(p)$ を求める．

以上により抽出した濃度こう配の画像例を図 3.3 に示す．これは，こう配強度を明度で表し，こう配方向を色相によって表している．図 3.3 を見ると，目に強く特徴的なこう配分布があることが見てとれる．

$$|G(p)| = \sqrt{G_x(p)^2 + G_y(p)^2} \quad (3.1)$$

$$\theta(p) = \tan^{-1} \frac{G_y(p)}{G_x(p)} \quad (3.2)$$

-1	0	1
-2	0	2
-1	0	1

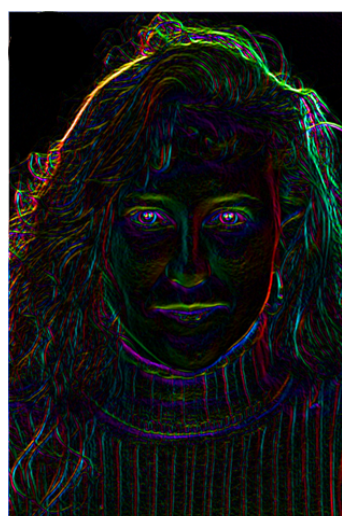
(a) 水平方向

-1	-2	-1
0	0	0
1	2	1

(b) 垂直方向

図 3.2: カーネルサイズ 3×3 の Sobel オペレータ

Grayscale image



Gradient image

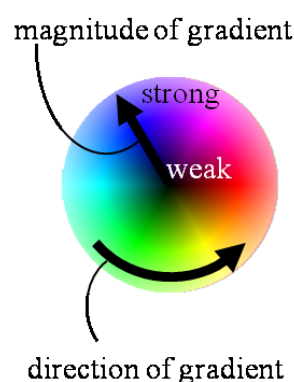


図 3.3: 濃度こう配の画像例

3.3 円の Hough 変換

瞳の形状から目は円形であるため、円形の図形検出を目的とする円の Hough 変換を用いることで、目の検出ができると考えられる。このことから、本手法では円の Hough 変換を用いる。円の Hough 変換は、画像中の特徴点が連続していなくても円を検出することができ、雑音にも強いことから、円を検出するのに頑健であると知られている [35]。

しかし、半径未知の円を検出するための円の Hough 変換は、 x, y 座標に円の半径 r を加えた 3 次元パラメータ空間 (x, y, r) を要するため、計算時間と記憶容量が大きくなるという問題がある。

一方、提案手法における円の Hough 変換は、まず 2 次元パラメータ空間 (x, y) 上から半径未知の円の中心のみを検出する。この円の Hough 変換は、前述した濃度こう配を利用する。まず Sobel オペレータにより顔画像から濃度こう配を算出し、画像の各エッジ点からこう配方向へ線分を描く。このとき、開始点のこう配強度を投票値とし、こう配強度

は，2次元パラメータ空間 (x, y) 上の線分に沿って，注目画素とその上下左右の4近傍に投票される．こう配方向に線分を描画することによって，2次元パラメータ空間 (x, y) に3次元パラメータ空間 (x, y, r) の投影を生成する．

このようにすることによって，たとえ半径未知の円に対しても円の中心位置に多くの投票がされ，パラメータ空間から極大点を検出することで円を検出することができる．さらに，白目領域に虹彩のような，より高いコントラストを持つ円は，開始点のこう配強度を投票することによって中心位置により多くの投票が集まる．このため，提案する円の Hough 変換を用いることで，目のようなコントラストのはっきりした円を検出するのに適している．図 3.4 に円の Hough 変換後のパラメータ空間を示す．

3.4 目の候補の検出

本研究では，円の Hough 変換により検出した円（極大点）を目の候補とする．円の Hough 変換後のパラメータ空間は，投票値の集まりにばらつきがある．極大点とは，8近傍に注目画素より大きい投票値がない点をいうので，平滑化処理を行わないと極大点が多数検出されることがある．これを防ぐためにパラメータ空間に対して，カーネルサイズが 3×3 の移動平均フィルタを用いて，平滑化を施す．図 3.5 は，円の Hough 変換後のパラメータ空間に対して平滑化を行ったものである．

平滑化後のパラメータ空間からしきい値を満たす極大点を検出し，目の候補とする．しきい値は，投票数の少ないノイズを検出しないためである．図 3.6 は，検出した目の候補をプロットした例である．



図 3.4: 円の Hough 変換後のパラメータ空間



図 3.5: 平滑化後のパラメータ空間

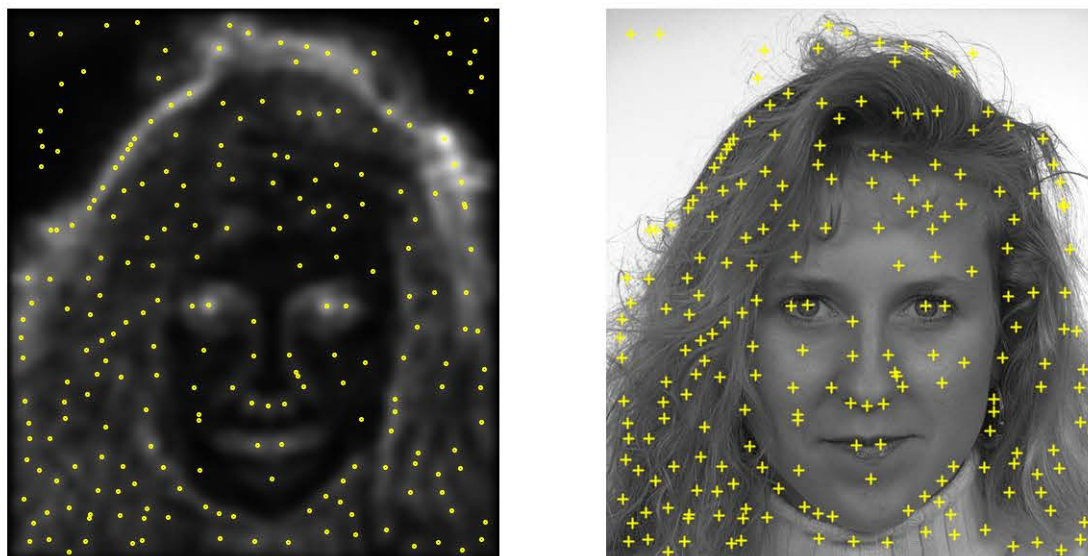


図 3.6: 検出した全ての目の候補

3.5 半径の推定

検出した目の候補の半径は，次のように推定している．まず，目の候補を中心とする固定半径の各円の円周に沿って，円周上のこう配強度を累積する．このとき，円周上の各画素におけるこう配強度は，中心に向かうベクトル n と実際のこう配ベクトル g との成す角 θ が $\cos \theta > 0.99$ を満たすもののみを累積し，累積する値はこう配強度と $\cos \theta$ とを掛け合わせた値とする．累積されたこう配強度は各半径で平均を取り，最大の平均強度を持つ半径を推定した半径とする．

3.6 特徴抽出

本研究では，目領域内の画像の見え方を濃度こう配ヒストグラム (HOG) を用いて表現する．濃度こう配ヒストグラムは，グレースケール画像中の対象物の見かけを濃度こう配の局所方向ヒストグラムで記述する特徴量であり，手書き文字認識において高い性能を有する特徴量として知られている [37, 36]．本手法では，この濃度こう配ヒストグラムをグレースケールの顔画像から抽出する．

特徴抽出の流れは，まず入力画像中の対象領域に対してノイズ除去のために平滑化を行う．続いて，Sobel オペレータにより各画素における濃度こう配の強度と方向とを算出する．そして対象領域を複数の小領域に分割し，濃度こう配の局所方向ヒストグラムを生成する．最後に，生成された局所方向ヒストグラムに対して平滑化処理などの後処理を行い，次元数の削減を行う．

以下では，それぞれの手順について詳細を述べる．

1. 平滑化

対象画像 $I = I(\mathbf{p})$ の画素 $\mathbf{p} = (p_x, p_y)$ に対して，以下のガウシアンフィルタを適用する．ここで N, σ は，それぞれカーネルサイズと重み係数であり， $N = 5, \sigma = 1.5$ とする．

$$I'(\mathbf{p}) = \frac{\sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} f(i, j) \cdot I(p_x + i, p_y + j)}{\sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} f(i, j)}, \quad (3.3)$$

$$f(i, j) = \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right). \quad (3.4)$$

2. 濃度こう配抽出

カーネルサイズが 3×3 の Sobel オペレータを用いて，濃度こう配の強度と方向とを算出する．画素 \mathbf{p} における濃度こう配の強度を $|G(\mathbf{p})|$ ，方向を $\theta(\mathbf{p})$ とすると，それぞれ以下の式で与えられる．なお， $G_x(\mathbf{p}), G_y(\mathbf{p})$ は，それぞれ画素 \mathbf{p} における濃度こう配の水平，垂直方向の成分である．

$$|G(\mathbf{p})| = \sqrt{G_x(\mathbf{p})^2 + G_y(\mathbf{p})^2} \quad (3.5)$$

$$\theta(\mathbf{p}) = \tan^{-1} \frac{G_y(\mathbf{p})}{G_x(\mathbf{p})} \quad (3.6)$$

3. 小領域分割と局所方向ヒストグラム生成

各画素ごとの濃度こう配の強度と方向とを，そのまま特徴ベクトルとして用いることは，その次元数が多すぎるため次元数の削減を行い，濃度こう配の特徴ベクトルを抽出している．まず，それぞれの画素で算出された濃度こう配の強度と方向とを局所ヒストグラム化するために，対象領域を $n_x \times n_y$ の小領域に分割する．さらに，量子化数 L により方向ごとの量子化を行う．続いて， $n_x \times n_y$ 個の小領域それぞれの内部で， L 次元の方向ごとに，濃度こう配の強度を積算していく．この処理によって， $n_x \times n_y \times L$ の原特徴ベクトルを得る．

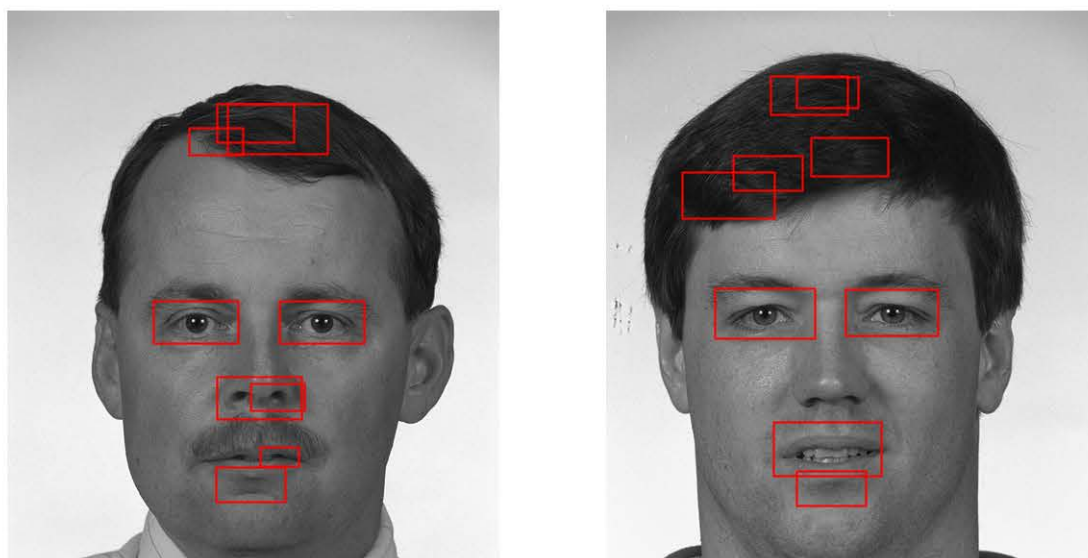


図 3.7: 特徴抽出窓の大きさの例

4. 特徴ベクトルの次元数の削減

抽出された特徴ベクトルの次元数を削減するために後処理を行う．ヒストグラム化された濃度こう配強度を分割された任意の小領域 $(2i, 2j)$ ($i = 0, 1, \dots, \frac{n}{2} - 1, j = 0, 1, \dots, \frac{m}{2} - 1$) を中心とする 5×5 の領域について，方向ごとにガウシアンフィルタを用いて平滑化する．これにより領域数を削減する．さらに， L 段階に量子化された濃度こう配の方向に対し，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，方向を $L/2$ に削減する．

なお，特徴抽出には様々なパラメータが存在するが，本研究では，小領域分割数 $n_x \times n_y$ は 5×5 ，方向量子化数 L は 32 で，次元削減は方向のみ行い，領域数の削減は行わない．したがって，特徴ベクトルには $5 \times 5 \times 16$ の 400 次元の特徴量を用いる．

これらの手順にしたがって，すべての目の候補から濃度こう配ヒストグラムを抽出する．まず，それぞれの目の候補点を中心とする長方形窓から濃度こう配ヒストグラムを抽出し，400 次元の特徴ベクトルを得る．このとき，特徴抽出窓の大きさは，推定した半径に基づき決定している．図 3.7 に特徴抽出窓の大きさの一例を示す．次に画像間のさまざまなコントラストに起因する変動を最小限に抑えるため，抽出した特徴ベクトルをノルムで正規化する．

3.7 目の組の作成

検出した目の候補を用いて，両目になり得る目の組を作成する．ここで全ての目の候補を用いて目の組を作成すると多数の目の組が作成されてしまい，処理時間がかかるだけで

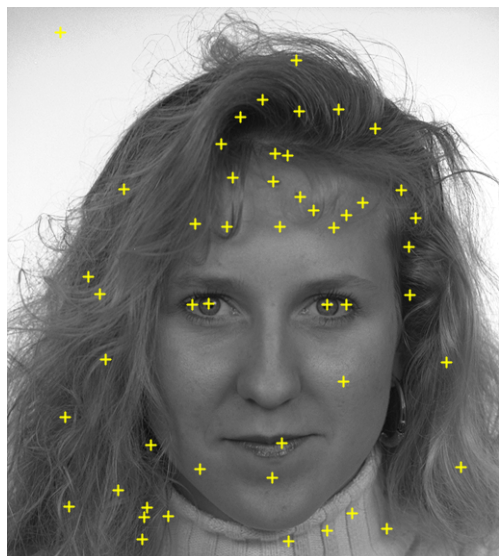


図 3.8: 上位 50 個の目の候補



図 3.9: 作成した目の組

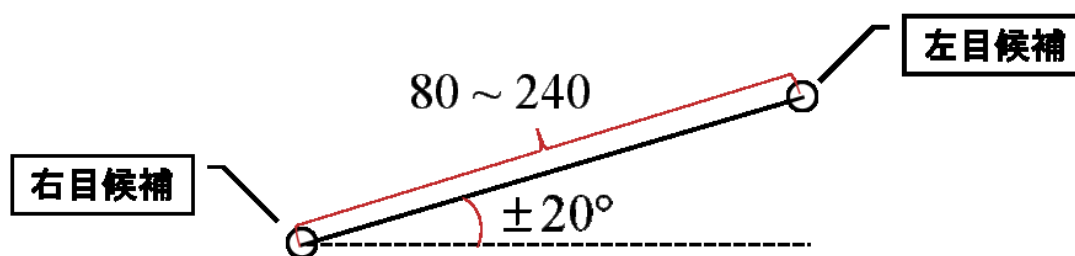


図 3.10: 目の角度と距離の関係

なく、本来両目であるはずの目の組を正しく求めることが困難になる。そのため、各目の候補を濃度こう配ヒストグラムで表現した目らしさに基づいて評価し、候補数の削減を行う。

まず全ての目の候補に対して SVM による評価を行い、画像の見かけによる目らしさの評価をする。目らしさの評価値は実数値で表され、この実数値により目の候補を降順に並び替える。この順序に基づき目の候補数を削減し、上位 50 個までを使用する目の候補とする。この使用する目の候補数については、4.6 節にて述べる。図 3.8 は、使用する 50 個の目の候補をプロットした例である。この上位 50 個までの目の候補を用いて、以下の条件を満たすように目の組を作成する。

1. 眼間線分の傾きが水平方向から ± 20 度以内である。
2. 眼間距離が $80 \sim 240$ pixel である。

この条件に基づき作成された各目の組は、SVM によって評価された左右の目の候補の目らしさの合計により順位付けを行う。図 3.9 は、目の組をプロットした例である。

3.8 対称性の評価

本来左目と右目は、顔の中心線である正中線を対称軸とする線対称な関係にある。したがって、左目候補の周辺領域と右目候補の周辺領域とを比較し、類似性を求めることで作成した目の組の対称性を評価することができる。まず各目の組の垂直 2 等分線によりそれぞれの目の組が水平になるように各候補ごとに画像の角度を補正する。次に候補間の距離を $2w$ とした時、候補点を中心とする 1 辺 $2w$ の正方形窓を生成し、Sobel オペレータによって正方形窓内の濃度こう配強度を求め、平滑化処理を加える。

図 3.11 に抽出窓の大きさの例を示す。平滑化には移動平均フィルタを用いた。また対称性を求めるため、左右の正方形窓から抽出したエッジ画像のうち、一方は鏡像を用いる。図 3.12 は、平滑化後のエッジ画像を表す例である。

抽出されたエッジ画像と、もう一方の鏡像との類似性を評価する。類似性の尺度として用いる関数は、コサイン係数 (cosine coefficient) とする。

$$\text{cosine coefficient} = \frac{X \cdot Y}{\|X\| \|Y\|} \quad (3.7)$$

ここで、 X, Y はそれぞれ右側領域と左側領域のエッジ画像であり、一方は鏡像である。

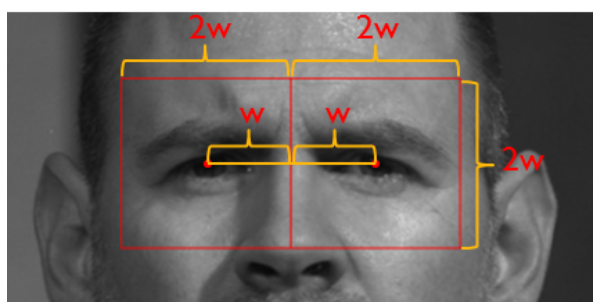


図 3.11: 抽出窓の大きさの例

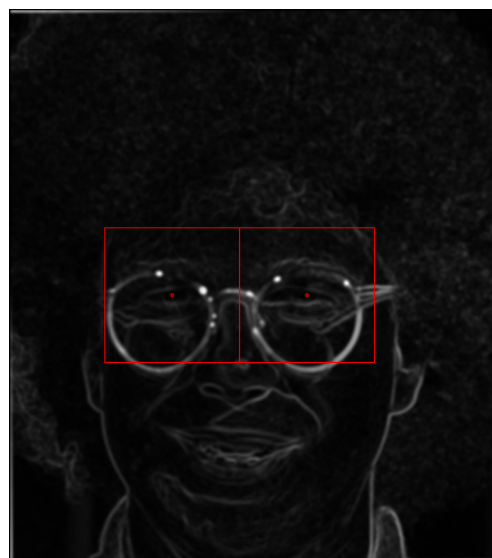


図 3.12: 抽出例

第 4 章

評価実験

4.1 FERET データベース

The Facial Recognition Technology (FERET) データベース [38] は、顔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、顔認識手法の評価では標準的に用いられている。データベース内の全ての画像は、無地または簡素な背景と 1 名の人物のみを含んでいる。撮影されている人物頭部の大きさにはばらつきがあり、表情の変化や眼鏡の着脱、首を傾けている人物のように様々な画像がある。画像サイズは 512×768 pixel である。

本研究では、FERET データベース内の正面顔を含む 2,409 枚を用いる。このうち、SVM の学習用に 1,000 枚、提案手法の性能評価用に 1,409 枚を使用する。なお、各画像には事前に目視により抽出された目の座標が校正用 (正解) 情報として与えられている。

図 4.1 に実験に使用する画像の一例を示す。

4.2 Support Vector Machine(SVM)

本研究では分類器に SVM を用いる。SVM は、現在知られている多くの手法の中で最も認識性能が優れた学習モデルの一つであり、主に 2 クラスの分類を行う分類機械である。SVM は学習サンプル中のクラス境界に最も近いサンプル (サポートベクトル) と、識別面との距離 (マージン) が最大となるように分離超平面を求めてクラスの分類を行う。本研究では SVM を、注目領域が目か、目以外かを分類するために利用する。



図 4.1: FERET データベースの画像例

4.3 学習サンプル

SVM に目か、目以外かを分類させるために学習を行う。学習には、FERET データベース内の画像 1,000 枚を用いて、1 画像あたり目の領域として 2 サンプル、目以外の領域として 5 サンプルの計 7 サンプルを使用する。したがって、SVM に学習させる合計サンプル数は 7,000 サンプルとなる。

目以外の学習サンプルは、円の Hough 変換時の投票数に応じて決定する。目以外の領域の 5 サンプルの選び方は、まず投票数上位 2 つ、次に投票数の中間値から 2 つ、最後に最も投票数の少ない 1 つとしている。このように選んだ理由は、円の Hough 変換の投票数の結果を用いることで、ランダム抽出より効率的に学習サンプルを決めることができると考えたからである。投票数が多ければ、目ではない円らしい学習サンプルが得られ、中間値からは髪や眉、鼻、口、頬など様々な領域の学習サンプルが得られる。そして、最も投票数の少ないところからは主に背景領域を学習できる。

図 4.2 に学習サンプルの一例を示す。白い点は全ての目の候補を示しており、黄色い枠は正解の目の領域、目以外の領域として、赤い枠は上位 2 つ、青い枠は中間値 2 つ、緑の枠は最も投票数の少ない 1 つを表している。

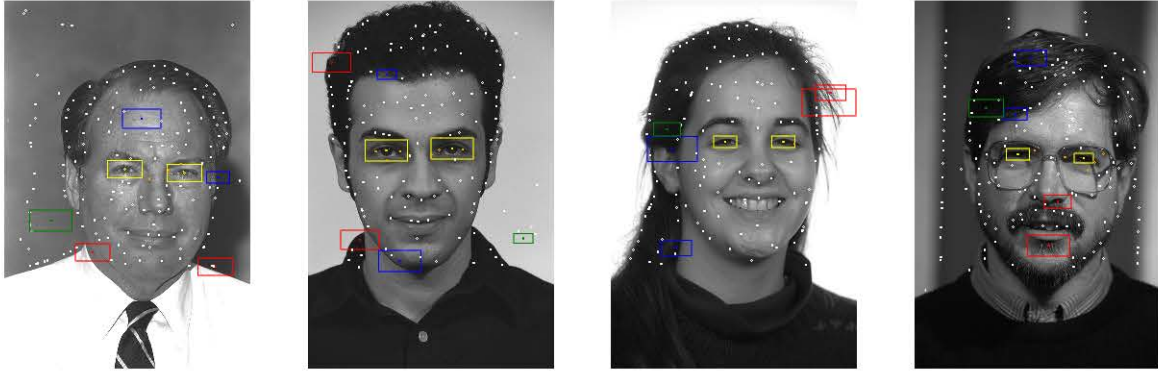


図 4.2: 学習サンプルの例

4.4 評価方法

FERET データベースに与えられている校正用情報を用いて、目の検出精度を定量的に評価する．この精度評価方法は、校正用情報として与えられている真の目の座標 $C_{r(x,y)}, C_{l(x,y)}$ と、検出した目の座標 $p_{r(x,y)}, p_{l(x,y)}$ とを比較する．なお、 $C_{r(x,y)}, p_{r(x,y)}$ は右目の座標を、 $C_{l(x,y)}, p_{l(x,y)}$ は左目の座標を表している．

評価基準にピクセル単位の絶対距離が用いられている場合、この評価基準は画像の解像度に依存する．そのため、本研究では評価方法として、Jesorsky ら [39] が用いている評価指標 d_{eye} を用いる．この評価基準は、解像度に依存しない相対誤差の尺度である．この評価指標の式を以下に示す．

$$d_l = \|C_{l(x,y)} - p_{l(x,y)}\|, d_r = \|C_{r(x,y)} - p_{r(x,y)}\| \quad (4.1)$$

$$d_{eye} = \frac{\max(d_l, d_r)}{\|C_{l(x,y)} - C_{r(x,y)}\|} \quad (4.2)$$

ここで、式 4.1 に示す d_l と d_r は、検出した目の座標と真の目の座標との誤差を表すユークリッド距離であり、 d_l は左目の誤差、 d_r は右目の誤差を表している．式 4.2 は、検出した目の座標と真の目の座標との距離の、左右における最大値を検出誤差とし、その誤差を眼間距離で正規化したものになる．例えば、 $d_{eye} < 0.1$ は眼間距離の 10% 以上の最大誤差を許容しないことを意味する．本研究では、 d_{eye} の値を 0.05 から 0.25 までの間で変化させ、目の検出率を求めた．

なお、検出率は以下に示す式 4.3 により求める．

$$\text{検出成功率 (\%)} = \frac{\text{検出成功枚数}}{\text{評価用画像数}} \times 100 \quad (4.3)$$

4.5 性能評価実験

提案手法の有効性を評価するために、FERET データベースの画像を用いて評価実験を行った。入力画像には、4.1 節で述べた 1,409 枚の正面顔画像を用いる。SVM は libSVM(ver.2.89)[40] を使用し、カーネルは RBF、 γ の値を 0.9 とした。実験は、目の組作成時に使用する目の候補数を 20 個、50 個、100 個と変化させて行った。

4.6 結果と考察

図 4.3 に、式 4.2 により表される目の検出の正しさの指標である d_{eye} において、 $d_{eye} < 0.1$ のときを検出成功とする累積ヒストグラムを示す。この図は、縦軸に両目の検出率、横軸に両目の候補数である RANK を表している。

図より、RANK1 では 20、50、100 個の場合でそれぞれ、97.94%、97.94%、97.94%を得た。RANK3 では 98.72%、98.94%、98.94%を得た。結果から、目の候補数を 50、100 個とした場合のそれぞれの両目の検出率に違いは表れなかった。さらに目の候補数が 50、100 個のときの検出率は、20 個の場合を上回る結果が得られた。このことから、目の組を作る際に使用する目の候補数は、50 個が適切であるといえる。提案手法による両目の検出率を見ても、高精度に両目を検出できていることがわかり、提案手法の有効性が確認できた。

対称性の評価をしない場合、RANK1 では 97.30%、RANK3 では 98.51%という結果となり、対称性の評価をすることで検出率が向上したことがわかる。

図 4.4 に検出に成功した両目の例を示す。各画像中の黄色の十字線は検出された目を表す。これらの結果画像を確認すると、提案手法による両目の検出は被写体の大きさの違いや、表情が変化していても正しく両目を検出できることがわかる。その他にも、照明の変化や眼鏡の着脱、顔の傾き、目を閉じている、などの様々な要因を含む画像に対しても両目の検出に成功していることがわかる。特に被写体である人物が目を閉じている例は、本来、円形である眼球が隠れているにも関わらず検出に成功している。本手法の特徴の 1 つである円の Hough 変換は円を検出するものであるが、今回の例のように目を閉じていたとしても、閉じた瞼付近に円の Hough 変換の投票が集まるため、目が検出できる場合がある。

図 4.5 に両目の検出に失敗した例を示す。各画像中の黄色の十字線は 3.4 節にて検出された全ての目の候補点を示している。両目の検出に失敗する主な原因は、目の候補点の中に正解の目が表れない場合であり、これらの画像を確認すると、本来の両目の位置に候補点が検出されていないことがわかる。図 4.5(d)、(e)、(f) は目元付近が極端に暗いため濃

度こう配の変化があまり見られず，3.3 節の円の Hough 変換を行う際に十分な投票が行われなかったのが原因であると考えられる．図 4.5(c) は眼鏡の映り込みの影響により，目の中心位置を特定できなかったことが原因であると考えられる．図 4.5(a)，(b)，(g)，(h) は，被写体である人物が目を閉じているため，本来，円形である黒目 (虹彩) が完全に隠れてしまい，円が検出できなかったのが原因である．なお，この失敗例は目の中心位置付近に検出されているが，式 4.2 の評価尺度である d_{eye} が $d_{eye} < 1.0$ を満たさず，瞼の検出に失敗した例である．

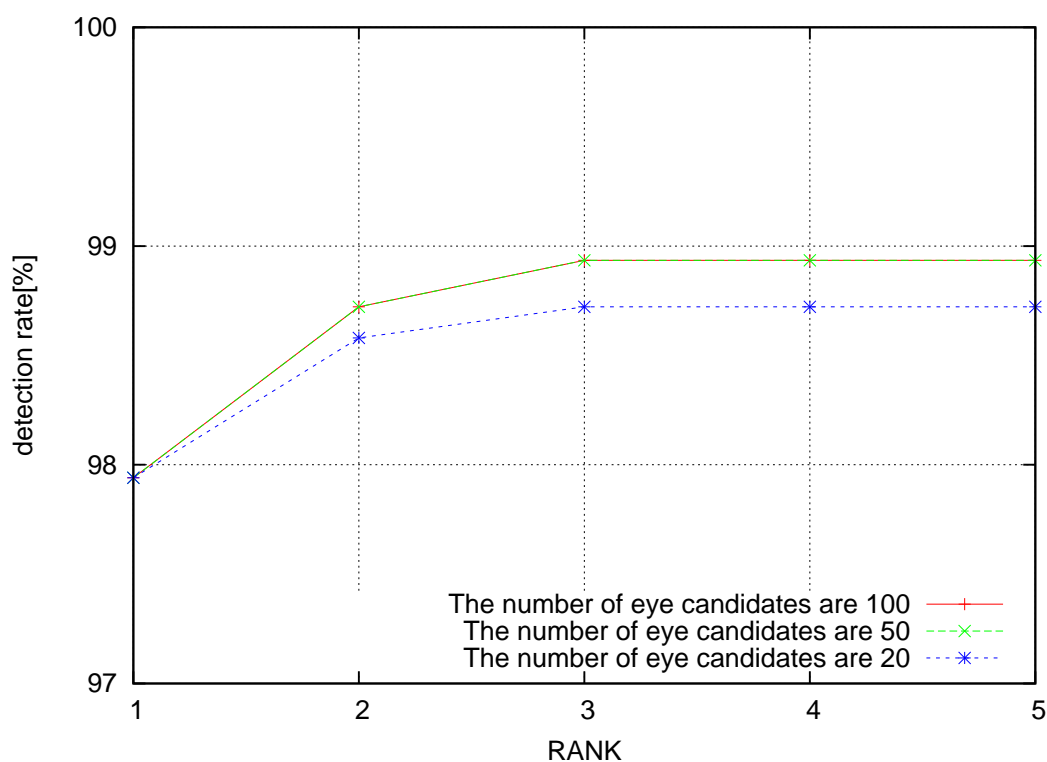


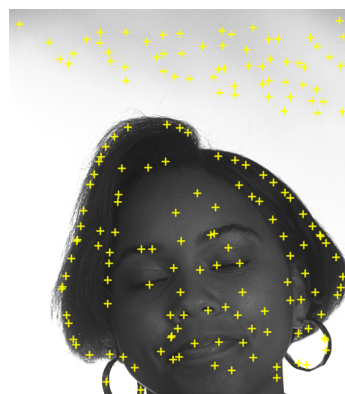
図 4.3: 提案手法による両目抽出精度 ($d_{eye} < 0.1$)



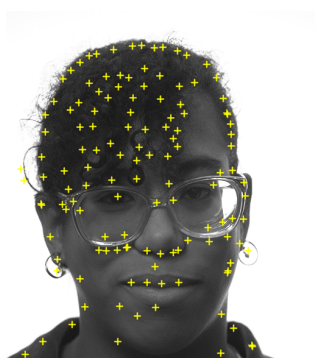
図 4.4: 提案手法による両目検出の例 (成功例)



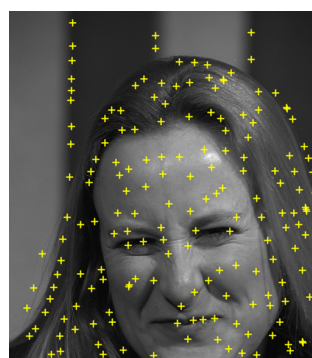
(a)



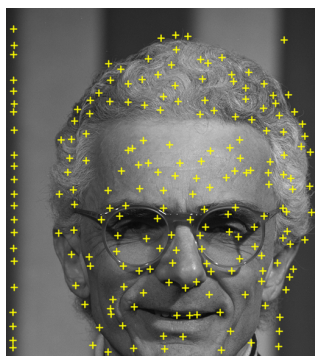
(b)



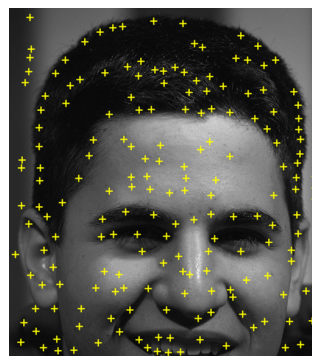
(c)



(d)



(e)



(f)



(g)



(h)

図 4.5: 提案手法による両目検出の例 (失敗例)

4.7 比較実験

提案手法の有効性を評価するために，他手法による両目検出の結果と比較を行う．比較に用いる手法は，Qian ら [10] と Song ら [11]，Asteriadis ら [12] の 3 手法である．各手法とも性能評価に FERET データベースを使用し，目の検出精度の指標には Jesorsky ら [39] が用いている式 4.2 に示す評価指標を用いる．各手法については，1.3 節を参照されたい．

4.8 結果と考察

図 4.6 に，提案手法と 3 つの他手法との比較結果を示す．この図は，縦軸に両目の検出率，横軸に式 4.2 から得られる許容誤差を表したものである．

結果のグラフから，提案手法は許容誤差が 0.06(6%) 以上の場合，他の 3 つの手法よりも高い検出率が得られることが確認できた．したがって，提案手法の有効性は他手法との比較実験においても実証できた．許容誤差が 6% 以下の場合に提案手法の検出率が低下するのは，正解の定義が異なるためと考えられる．FERET データベースが定める正解の目の位置は，目の領域全体の中心位置であり，黒目の中心ではない．一方，本研究では黒目の中心を正解の目の位置と定義するため，両者の目の位置の正解定義に違いがある．

図 4.7 に，FERET データベースにより与えられた正解の目の位置を図示した例を示す．図から，FERET データベースは黒目の中心を正解としていないことがわかる．図 4.8 は，図 4.7 に対して，正解の目の位置を本研究で定める目の位置に定義し直した場合の目の中心位置を示している．このように本研究では，黒目の中心位置を正解と定義している．また FERET データベースには，図 4.9 に図示するように正しく目の中心位置を捉えていない例も存在した．これらのことから，FERET データベースと本研究との間で正解の定義が異なるため，許容誤差が 6% 以下の場合に提案手法の検出率が低下すると考えられる．

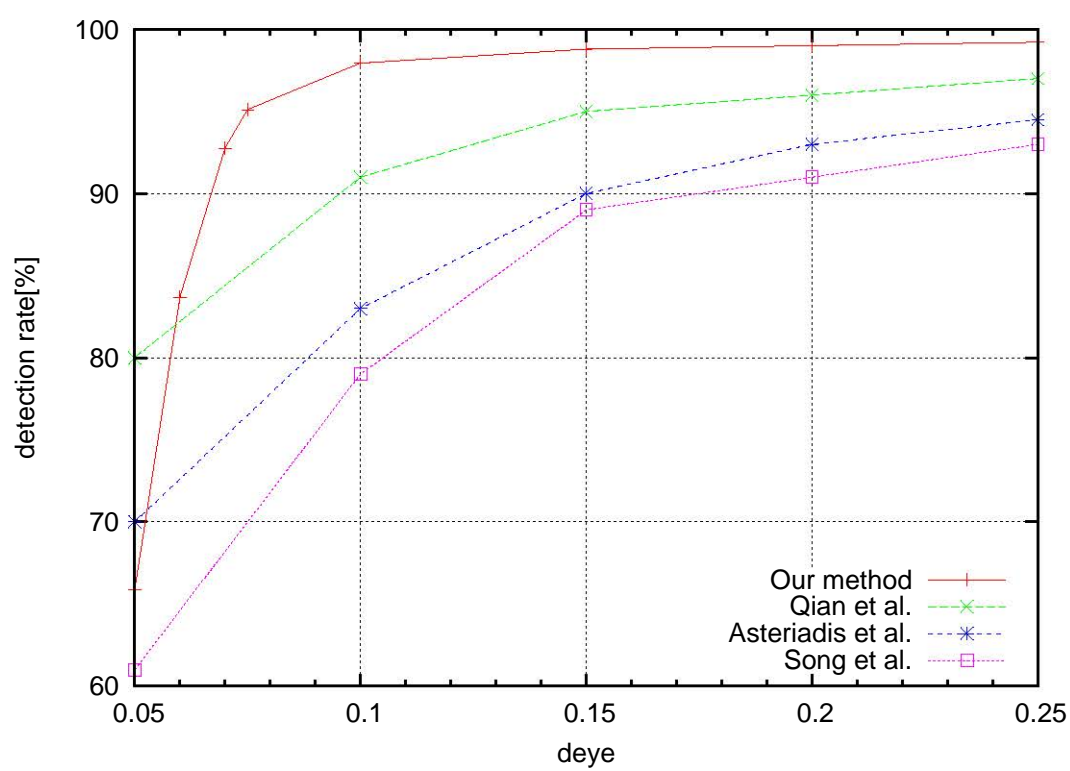


図 4.6: 提案手法と 3 つの他手法との比較結果

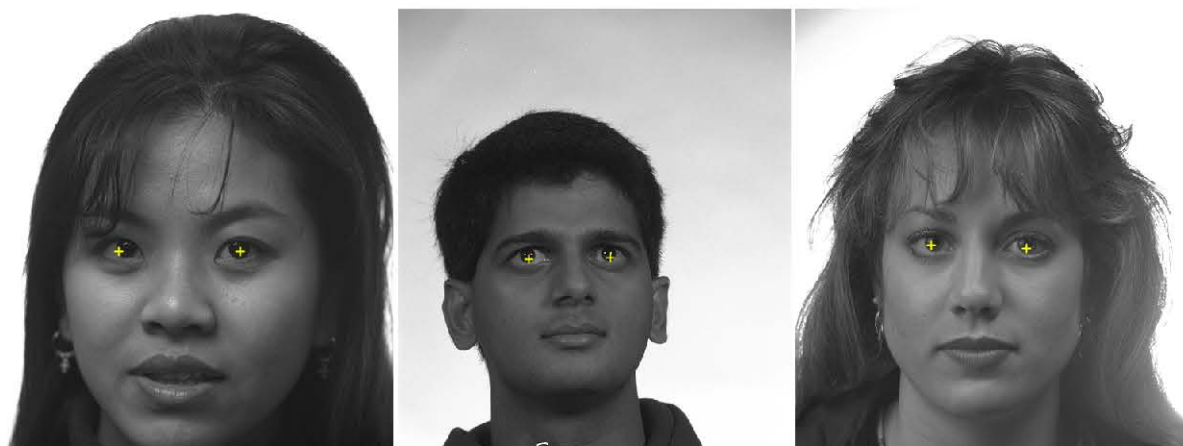


図 4.7: FERET データベースの正解の目の位置の例

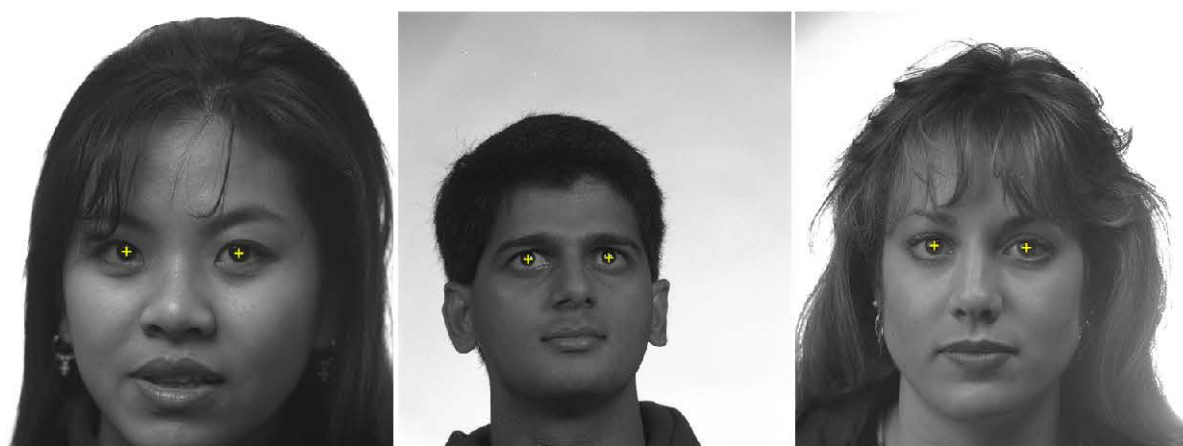


図 4.8: 本研究における正解の目の位置の例

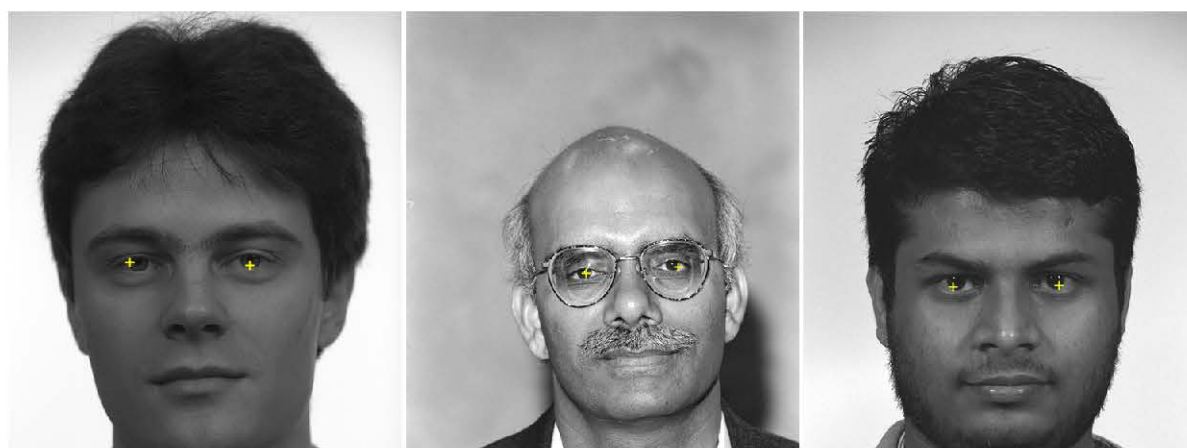


図 4.9: FERET データベースの正しく目の中心位置を捉えていない例

第 5 章

結言

5.1 まとめ

本研究では、濃度こう配を利用する円の Hough 変換と濃度こう配ヒストグラム、SVM による両目の検出手法を提案し、その有効性を検証した。提案手法の特徴は、円の Hough 変換を 2 次元のパラメータ空間で行っているため、3 次元の円の Hough 変換よりも効率的に半径未知の円を検出することができる。そして、画像の見かけを利用した目らしさの評価を濃度こう配ヒストグラムと SVM を用いて行っている点である。

FERET データベースの正面顔画像 1,409 枚を用いた評価実験の結果、検出した両目の組を 1 組とした場合、97.94%の検出率が得られ、検出した両目の組を 3 組とした場合、98.94%の検出率が得られた。この結果は、Qian らの検出率 91.95%を大きく上回り、提案手法の有効性と信頼性が大規模な画像データベース上での実験により実証された。これらのことから、提案手法を用いることで正面顔画像中から高精度に両目を検出することが可能であるとわかった。

5.2 今後の課題

本研究では、無地または単調な背景の画像を多く含む FERET データベースの画像を用いて、提案手法の有効性を評価したが、複雑な背景を持つ画像に対しては、その有効性を確認していない。そのため、今後は複雑な背景を持つ画像に対して、提案手法の有効性を確認したい。また、鼻や口のような両目以外の顔部品の検出、顔認識への応用が今後の課題である。

付録 A

プログラムソースリスト

本研究に関するプログラムを以下のディレクトリ

yasutaka/Research/M2/PROGRAM

に置く。ただし、濃度こう配ヒストグラム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置く。

yasutaka/Research/M2/lib

yasutaka/Research/M2/include

ディレクトリ構造と各プログラムの概略を以下に示す。なお、詳細については各ディレクトリの README に記す。

PROGRAM

```
|--before/
|   # SVM に評価させる特徴ベクトルを抽出する一連のプログラム。
|   # 学習サンプルを作成する際にも用いる。
|
|--libsvm-2.89/
|   |--svm-train
|       # 実験で用いる分類器を設計する。
|
|   |--svm-predict
|       # 分類器の性能評価を行う。
|       # 本研究ではこのプログラムを用いて、目らしさの評価を行った。
|
|
```

```
|--after/  
|  # SVM が評価した目らしさを記述したファイルを読み込み ,  
|  # 両目を検出する .  
|  
|--list/  
|  # 実験に使用するデータファイルの絶対アドレスを記述したファイル群 .  
|  
|--do_*.sh  
    # 各種実験を行うためのシェルスクリプト .
```

A.1 コンパイル方法

各プログラムは以下のコマンドでコンパイルすることができる .

```
gmake
```

A.2 実行方法

実行ファイルの実行方法を以下に示す .

```
./sample [Input_image] [Output_image] [Image_type] [Delete 1/4] [Dimension]
```

Input_image:入力画像名

Output_image:出力画像名

Image_type:出力画像の種類 (0:Sobel, 1:Sobel_half_binary, 2:Hough, 3:Original_binary, 4:Sobel_binary, 5:Hough 空間)

Delete 1/4:画像下部 1/4 を削除 (0:する, 1:しない)

Dimension:Hough 変換の次元数 (2:2 次元, 3:3 次元 (2 を指定する))

実行例を以下に示す .

```
./afterSVM_process /home/FERET/fafb/00001_930831_fa.a.pgm test.png 5 0 2
```

シェルスクリプトの実行例を以下に示す .

```
sh do_usingOutputFile_aveCoef2svm.sh < list/eval_data_1409.list
```

付録 B

研究データ

研究に使用したデータは venus のローカル

/home/FERET/

に置く．ディレクトリ構造と各ファイルの概略を以下に示す．

FERET

```
|--fafb/
|   # 実験に使用する人物画像が存在するディレクトリ
|
|--ground_truth/
|   # 画像ファイルの正解データが記述された.txt が存在するディレクトリ（黒目の中心に座標を修正済み）．
|
|--ground_truth_org/
|   # FERET オリジナルの正解データが記述された.txt が存在するディレクトリ．
|
|--output_file_7GAKUSYUU/
|   # 実験に使用した人数分の ,SVM が出力した output.txt ファイルが存在するディレクトリ．
|   # 使用した SVM は 1 枚あたり , 目:2 サンプル , 目以外:5 サンプルの合計 7 サンプルを学習．
```

付録 C

発表資料

修論発表で用いたプレゼンテーション資料を本論文の末尾に掲載する。

謝辞

本研究に関し，研究に対する基本姿勢や理論的裏付け，多くのアイデアを御教授下さった木村文隆教授，研究における様々な技術を御教授下さった若林哲史准教授，研究に対する指摘や資料作成，発表のコツ，プレゼンテーションの作り方などをご指導いただきました大山航助教，深く幅広い知識やアイデア，専門的観点からの指摘など多くを御教授下さった三重大大学三宅康二名誉教授に深く感謝します．また，日頃お世話になった田中みゆき事務，多くの有益なアドバイスを下さった研究室の先輩方，共に切磋琢磨した同期の皆様，楽しい学生生活を与えて下さった研究室の仲間に感謝します．

最後になりましたが，長きにわたる私の学生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.Phillips, and A.Rosenfeld:“Face recognition: A literature survey”, Acm Computing Surveys (CSUR), Vol.35, No.4, pp.399-458, 2003.
- [2] X.Tan, S.Chen, Z.H.Zhou, and F.Zhang:“Face recognition from a single image per person: A survey”, Pattern Recognition, Vol.39, No.9, pp.1725-1745, 2006.
- [3] A.F.Abate, M.Nappi, D.Riccio, and G.Sabatino:“2D and 3D face recognition: A survey”, Pattern Recognition Letters, Vol.28, No.14, pp.1885-1906, 2007.
- [4] B.Fasel and J.Luettin:“Automatic facial expression analysis: a survey”, Pattern Recognition, Vol.36, No.1, pp.259-275, 2003.
- [5] W.Ohyama, N.Nakao, T.Wakabayashi, and F.Kimura: “Automatic detection of facial midline and its contributions to facial feature extraction”, ELCVIA, Vol.6, No.3, pp.55-65, 2007.
- [6] 岡本 大地, 大山 航, 若林 哲史, 木村 文隆:“チェインコード化 Merlin-Farber 法を用いた正面顔正中線の自動抽出”, 電学論 C, Vol. 130, No. 12, pp.2159-2166 (Dec. 2010)
- [7] P.Gupta, H.Mehrotra, A.Rattani, A.Chatterjee and A.K.Kaushik:“Iris recognition using Corner Detection”, Proc. 23rd Int ' l Biometric Conference, 2006.
- [8] Y.Ohkawa, C.H.Suryanto, and K.Fukui: “Fast Combined Separability Filter for Detecting Circular Objects”, MVA2011, IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN(4-16)
- [9] 森山 剛, 金出 武雄, Jeffrey F.Cohn, 小沢 慎治:“Active Appearance Model 探索による目領域構造の自動抽出”, 画像の認識・理解シンポジウム (MIRU2006), 2006.
- [10] Z.Qian, and D.Xu: ” Automatic eye detection using intensity filtering and K-means clustering”, Pattern Recognition Letters, Vol.31, No.12, pp.1633-1640, 2010.
- [11] J.Song, Z.Chi, and J.Liu : “A robust eye detection method using combined binary edge and intensity information”, Pattern Recognition, Vol.39, No.6, pp.1110-1125, 2006.
- [12] S.Asteriadis, N.Nikolaidis, and I.Pitas : “Facial feature detection using distance vector fields”, Pattern Recognition, Vol.42, No.7, pp.1388-1398, 2009.

- [13] M.Kirby and L.Sirovich : “Application of the Karhunen-Loeve procedure for the characterization of human faces”, IEEE Trans. Patt. Anal. Mach. Intell.12. 1990.
- [14] L.Sirovich and M.Kirby: “Low-dimensional procedure for the characterization of human face”, J. Opt. Soc. Am. 4, 519-524. 1987.
- [15] A.M.Martinez, and A.C.Kak:“Pca versus lda”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.23, No.2, pp.228-233, 2001.
- [16] J.Lu, K.N.Plataniotis, and A.N.Venetsanopoulos:“Face recognition using LDA-based algorithms”, Neural Networks, IEEE Transactions on, Vol.14, No.1, pp.195-200, 2003.
- [17] M.Turk, and A.Pentland:“Eigenfaces for recognition”, Journal of cognitive neuroscience, Vol.3, No.1, pp.71-86, 1991.
- [18] B.Moghaddam, C.Nastar, and A.Pentland:“A Bayesian similarity measure for direct image matching”, Pattern Recognition, Proceedings of the 13th International Conference on, Vol.2, pp.350-358, 1996.
- [19] M.S.Bartlett, J.R.Movellan, and T.J.Sejnowski:“Face recognition by independent component analysis”, Neural Networks, IEEE Transactions on, Vol.13, No.6, pp.1450-1464, 2002.
- [20] S.H.Lin, S.Y.Kung, and L.J.Lin:“Face recognition/detection by probabilistic decision-based neural network”, Neural Networks, IEEE Transactions on, Vol.8, No.1, pp.114-132, 1997.
- [21] S.Y.Kung and J.S.Taur : “Decision-based neural networks with signal/image classification applications”, IEEE Trans. Neural Netw. 6, 170-181, 1995.
- [22] F.Samaria, and S.Young:“HMM-based architecture for face identification”, Image and vision computing, Vol.12, No.8, pp.537-543, 1994.
- [23] BS.Manjunath, R.Chellappa, and von der Malsburg, C.: “A feature based approach to face recognition”, Computer Vision and Pattern Recognition, 1992. Proceedings CVPR’92., 1992 IEEE Computer Society Conference on, pp.373-378, 1992.
- [24] B.Kepenekci, Boray Tek, F. and Bozdagi Akar, G: “Occluded face recognition based on Gabor wavelets”, Image Processing. 2002. Proceedings. 2002 International Conference on, Vol.1, pp.1-293, 2002.
- [25] X.Tan, S.Chen, Z.H.Zhou, and F.Zhang: “Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble”, Neural Networks, IEEE Transactions on, Vol.16, No.4, pp.875-886, 2005.
- [26] L.Wiskott, J.M.Fellous, N.Kuiger, and von der Malsburg, C.: “Face recognition by elastic bunch graph matching”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.19, No.7, pp.775-779, 1997.

- [27] S.Chen, J.Liu, and Z.H.Zhou: “Making FLDA applicable to face recognition with one sample per person”, Pattern recognition, Vol.37, No.7, pp.1553-1555, 2004.
- [28] A.M.Martinez: “Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.24, No.6, pp.748-763, 2002.
- [29] F.Samaria, and Olivetti Research Ltd: “Face segmentation for identification using hidden Markov models”, 1993.
- [30] H.S.Le, and H.Li: “Recognizing frontal face images using hidden Markov models with one training image per person”, Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, Vol.1, pp.318-321, 2004.
- [31] P.S.Penev, and J.J.Atick: “Local feature analysis: A general statistical theory for object representation”, Network: computation in neural systems, Vol.7, No.3, pp.477-500, 1996.
- [32] A.Lanitis, C.J.Taylor, and T.F.Cootes: “Automatic face identification system using flexible appearance models”, Image and vision computing, Vol.13, No.5, pp.393-401, 1995.
- [33] A.Pentland, B.Moghaddam and T.Starner: “View-based and modular eigenspaces for face recognition”, In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition.
- [34] 田村秀行 (編): コンピュータ画像処理, pp.182-199. オーム社, 2002.
- [35] 田村秀行 (編): コンピュータ画像処理, pp.200-206. オーム社, 2002.
- [36] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二, “特徴量の次元数増加による手書き数字認識の高精度化”, 電子情報通信学会論文誌 D-II, Vol.J77-D-II, No.10, pp.2046-2053, 1994.
- [37] Tetsushi Wakabayashi, Shinji Tsuruoka, Fumitaka Kimura and Yasuji Miyake : “Increasing the feature size in handwritten numeral recognition to improve accuracy”: Systems and Computers in Japan, Volume 26 Issue 8, pp.35-44, 1995.
- [38] 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.
- [39] O.Jesorsky, K.J.Kirchberg, and R.W.Frischholz : “Robust face detection using the hausdorff distance”, In Proc Int ' l Conf. Audio- and Video-Based Biometric Person Authentication, pp.90-95, 2001.
- [40] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>