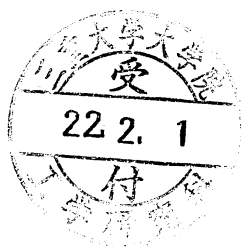


修 士 論 文

回転不変濃度こう配特徴と その顔検出への応用



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

横田 智彦

はじめに

画像中の物体の認識・検出は、検索質問を画像とする大規模画像検索 [1]、特定物体の認識・検出、生体機能である視覚のモデル化など広く応用がなされている。制約のない環境下で取得された画像中から特定の物体を認識・検出することは、計算機にとっていまだ難しい問題である。実環境で撮影された画像中から安定して物体を認識・検出する問題は、画像中の対象物体の位置、姿勢、個数などの課題を克服する必要がある。

本研究では、画像平面上の回転 (**Rotation In Plane:RIP**) に基づく物体の姿勢の変化があっても安定に対象物体を検出するために、画像中の物体から **RIP** に対して不変な特徴ベクトルを抽出する手法を提案する。また、提案手法を顔検出処理に応用し、回転を含む顔の自動検出を行う。回転を含む顔の検出は、自動視点顔検出 (**Multi-View Face Detection**) として顔画像認識の研究分野で重要な課題である。

提案手法は、画像中の物体認識、物体検出の高精度化に寄与する特徴量として有効性が示されている濃度こう配特徴 [20] を応用する。物体の回転に伴う濃度こう配の向きの変動と局所領域中の画素の位置変化に不変となるように特徴抽出処理を行い、回転不変な濃度こう配特徴を記述する。無回転の物体と回転を含む物体とのそれぞれから、提案手法を用いて抽出した特徴ベクトルは、相互相関が高かった。このことから、提案手法を用いることで回転に不変な濃度こう配特徴が得られることがわかった。

また、提案手法を顔検出処理に応用した実験の結果から、提案手法を用いた顔検出処理は、画像中の顔領域の回転に依存しないことが確認された。実験の結果から、検出成功率の最大値 91.1% が得られた。

濃度こう配特徴はこれまでに対象物体の拡大、縮小に対して不変な特徴抽出方法が提案されてきたが、画像平面上の回転に対して不変な抽出方法は十分に検討されていなかった。本研究は、濃度こう配特徴の応用と回転不変特徴の抽出手法を新たに提案した。提案手法の実現は、自動視点顔検出などの回転による姿勢変化を含む物体認識、検出の研究分野へのアプローチの単純化に寄与すると考える。

今後の課題として、(1) 回転を伴う対象として図面中の文字認識や一般物体検出などの分野への応用、(2) 他の回転不変特徴 [3] [4] [5] との比較が挙げられる。顔検出以外への

応用や他の回転不変特徴との比較を通じて，提案手法の有効性を示したい．

目次

| | |
|---------------------------|-----------|
| はじめに | i |
| 第 1 章 緒言 | 1 |
| 1.1 研究の背景 | 1 |
| 1.2 研究の目的 | 2 |
| 1.3 関連研究 | 2 |
| 1.4 論文の構成 | 3 |
| 第 2 章 濃度こう配特徴 | 4 |
| 2.1 濃度こう配特徴とは | 4 |
| 2.2 特徴抽出処理の流れ | 4 |
| 2.2.1 平滑化 | 6 |
| 2.2.2 濃度こう配の抽出 | 6 |
| 2.2.3 小領域分割と局所方向ヒストグラムの生成 | 6 |
| 2.2.4 特徴ベクトルの次元数の削減 | 6 |
| 第 3 章 回転不変濃度こう配特徴 | 8 |
| 3.1 提案手法の概要 | 8 |
| 3.2 濃度こう配の向きの変換 | 8 |
| 3.3 回転に対して不変な小領域分割 | 10 |
| 3.3.1 同心円形領域分割 | 10 |
| 3.3.2 セクター領域分割 | 10 |
| 3.3.3 同心円 + セクター領域分割 | 10 |
| 3.4 回転に対する不変性の評価 | 11 |
| 第 4 章 顔検出への応用 | 14 |
| 4.1 顔検出とは | 14 |
| 4.2 顔検出の代表的な手法 | 14 |

| | | |
|-------|---|----|
| 4.2.1 | 事前知識に基づく顔検出 | 15 |
| 4.2.2 | 特徴に基づく顔検出 | 15 |
| 4.2.3 | テンプレートマッチングに基づく顔検出 | 15 |
| 4.2.4 | 見え方 (<i>appearance</i>) に基づく顔検出 | 16 |
| 4.3 | 本研究における顔検出処理の流れ | 17 |
| 4.3.1 | サポートベクタマシン (SVM) | 17 |
| 第 5 章 | 評価実験 | 18 |
| 5.1 | FERET データベース | 18 |
| 5.2 | 評価方法 | 18 |
| 5.3 | 性能評価実験 | 20 |
| 5.3.1 | 分類実験 | 20 |
| 5.3.2 | 検出実験 | 20 |
| 5.4 | 考察 | 25 |
| 第 6 章 | 結言 | 27 |
| 6.1 | まとめ | 27 |
| 6.2 | 今後の課題 | 27 |
| 付録 A | 特徴抽出におけるパラメータの決定 | 28 |
| A.1 | 予備実験 | 28 |
| A.2 | 実験結果 | 28 |
| 付録 B | プログラムソースリスト | 31 |
| 付録 C | 画像ファイルソースリスト | 33 |
| 付録 D | 発表資料 | 36 |
| 謝辞 | | 37 |

第 1 章

緒言

1.1 研究の背景

画像中の物体の認識・検出は，検索質問を画像とする大規模画像検索 [1]，特定物体の認識・検出，生体機能である視覚のモデル化など広く応用がなされている．例としては，車両のナンバープレート認識や空港における顔認識システムなどが挙げられる．ナンバープレート認識や顔認識システムは，主要幹線道路での不審車両の監視や税関エリアにおける犯罪者の特定といったセキュリティ対策として利用されている．また，車載カメラ画像から道路交通標識を認識する研究，自動走行する災害救助ロボットの開発など，画像中の物体を認識・検出する技術は盛んに研究がなされている．

制約のない環境下で取得された画像中から特定の物体を認識・検出することは，計算機にとっていまだ難しい問題である．実環境で撮影された画像中から安定して対象物体を認識・検出するために克服すべき課題を以下に示す．

1. **物体の位置**．対象物体が画像中のどの位置に，どの大きさで含まれているかは未知である．
2. **物体の姿勢**．被写体となる物体とカメラとの相対的な位置関係の変化により，物体の見え方は大きく変動する．
3. **物体の隠れ**．対象物体の一部，あるいは全部が他の物体で隠される場合がある．
4. **物体の個数**．対象物体が画像中にいくつ含まれているか，あるいは含まれていないかは未知である．
5. **見え方の変化**．撮影環境の光源位置やデバイスの性能，特性により画像中の物体の見え方は容易に変化する．

上に示した課題のうち，物体の姿勢の変化は主に物体の回転に基づく．画像中の物体の回転は，画像平面上の回転（Rotation In Plane:RIP）と画像平面外の回転（Rotation Off

Plane:ROP) とに分離できる. 本研究では **RIP** のみを対象とする. 以下, 本文中では **RIP** のみを指して回転と呼ぶ.

一般的に画像中の対象物体の見かけを記述する特徴量においては, 対象領域内部の見かけが物体の回転により変動した場合, 対象領域から抽出される特徴量も変動する. しかし, 特徴量が適用される問題によっては, 回転に伴う特徴量の変動が望ましくない場合がある. 例として, 対象物体の隠れを伴わない回転に対して, 特徴量の変動する場合を考える. 物体が回転している場合でも, 安定して認識, 検出する必要がある場合, 回転に伴う特徴量の変動は分類器で吸収される必要がある. 具体的には, 文献 [2] に示されるように複数の分類器を競合させたり, 回転を含む学習データセットを用いた分類器の学習が要求される. 回転に対して頑健な特徴量の実現は, このような問題に対する認識, 検出アプローチの単純化に寄与する.

1.2 研究の目的

本研究では, 画像中の物体から回転に対して不変な特徴ベクトルを抽出する手法を提案する. 具体的には, 画像中の物体認識, 物体検出の高精度化に寄与する特徴量として有効性が示されている濃度こう配特徴を, 回転不変な特徴量にする. さらに, 提案手法を用いて, 回転を含む物体の学習や複数の分類器を競合することなしに, 画像中の物体が回転を含んでいても高精度に認識, 検出するシステムの構築を目指す.

また, 提案手法を顔検出処理に応用し, 回転を含む顔の自動検出を行う. 回転を含む顔の検出は, 自動視点顔検出 (**Multi-View Face Detection**) として顔画像認識の研究分野で重要な課題である.

1.3 関連研究

画像中の物体から **RIP** に対して不変な特徴を抽出する手法としては, **Lowe** が提案した **Scale Invariant Feature Transform (SIFT)** [3] が挙げられる. **SIFT** は, まず入力画像とガウス関数とを畳み込んだ平滑化画像を作成し, スケールの異なる平滑化画像間の差分処理 (**Difference of Gaussian:DOG**) により, 特徴抽出に適した点 (キーポイント) の検出とスケールの決定を行う. 次に, 検出されたキーポイント周辺の輝度こう配ヒストグラムから, 基準となる方向 (オリエンテーション) を割り出す. キーポイント周辺の領域をオリエンテーション方向に回転して特徴抽出を行うことで, **SIFT** は回転不変性を得ている.

これ以外に松浦らが提案した二次元フーリエ級数展開を用いた特徴抽出 [4] がある. この手法では, まず画像の画素値が与えられた離散データを直交座標における関数値として持つ連続関数を求め, その極座標表現を得る. 極座標表現された画像を二次元フーリエ級

数に展開することで、RIP に対して不変な特徴量を抽出している。

他の手法グループとして、複素対数極座標変換 (Log Polar Mapping:LPM) がある。LPM は、人の視界のように特徴抽出領域中心の解像度が高く、領域の外縁に近いほど解像度が低くなる中心窩視覚センサの最も簡単なモデルである。LPM 自体は特徴を回転不変にするものではないが、LPM を施した画像上では、物体の回転は平行移動として捉えることができる。栗田らは LPM と自己相関関数を用いて RIP に不変な物体認識を提案している [5]。

SIFT は回転不変性を得るために、一度生成したこう配ヒストグラムから基準となる向きを割り出して、再度こう配ヒストグラムを生成する。提案手法は、特徴抽出領域の中心点を基準として、各画素と各こう配の向きとの相対角度を求めることで、こう配特徴の回転不変性を得る。また、松浦らの手法では濃淡特徴に対して二次元フーリエ級数展開を施して回転不変特徴を抽出しているが、提案手法では、こう配特徴に対して回転方向に一次元のフーリエ変換を施すことで回転不変特徴を抽出する。

1.4 論文の構成

本論文ではまず、第 2 章にて濃度こう配特徴について紹介し、その記述方法を述べる。続いて、回転に不変な濃度こう配特徴を抽出するための提案手法について第 3 章で述べ、抽出した特徴量の回転不変性を評価した実験結果について述べる。第 4 章では、顔画像処理に関する従来研究について紹介し、提案手法の顔検出処理への応用について述べる。さらに、第 5 章では提案手法の顔検出処理における性能を評価した実験結果と考察について述べ、第 6 章にてまとめと今後の課題について述べる。

第 2 章

濃度こう配特徴

この章では、濃度こう配特徴とその抽出方法について述べる。濃度こう配特徴の抽出方法には、抽出処理の過程で用いられる手法にいくつかのバリエーションがある。Dalal らは文献 [19] にて抽出の手順に含まれる種々のパラメータについて詳細な検討を行ったが、抽出される特徴量の特性は基本的には同じである。

2.1 濃度こう配特徴とは

濃度こう配特徴は、濃淡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、画像の認識や画像からの対象物体検出において、認識や検出の高精度化に寄与する特徴量として有効性が示されている [22]。濃度こう配特徴は当初、若林ら [20] によって手書き数字認識の高精度化を目的として開発された。その後、本郷ら [21] による 4 方向面特徴の顔認識への応用や、澤ら [23] による手書き漢字認識への応用、Dalal ら [19] による HOG 特徴の人物検出への応用など、様々な分野への応用が行われている。

2.2 特徴抽出処理の流れ

本研究における濃度こう配特徴抽出処理の流れを図 2.1 に示す。こう配特徴の抽出処理では、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向きとを算出する。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。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。

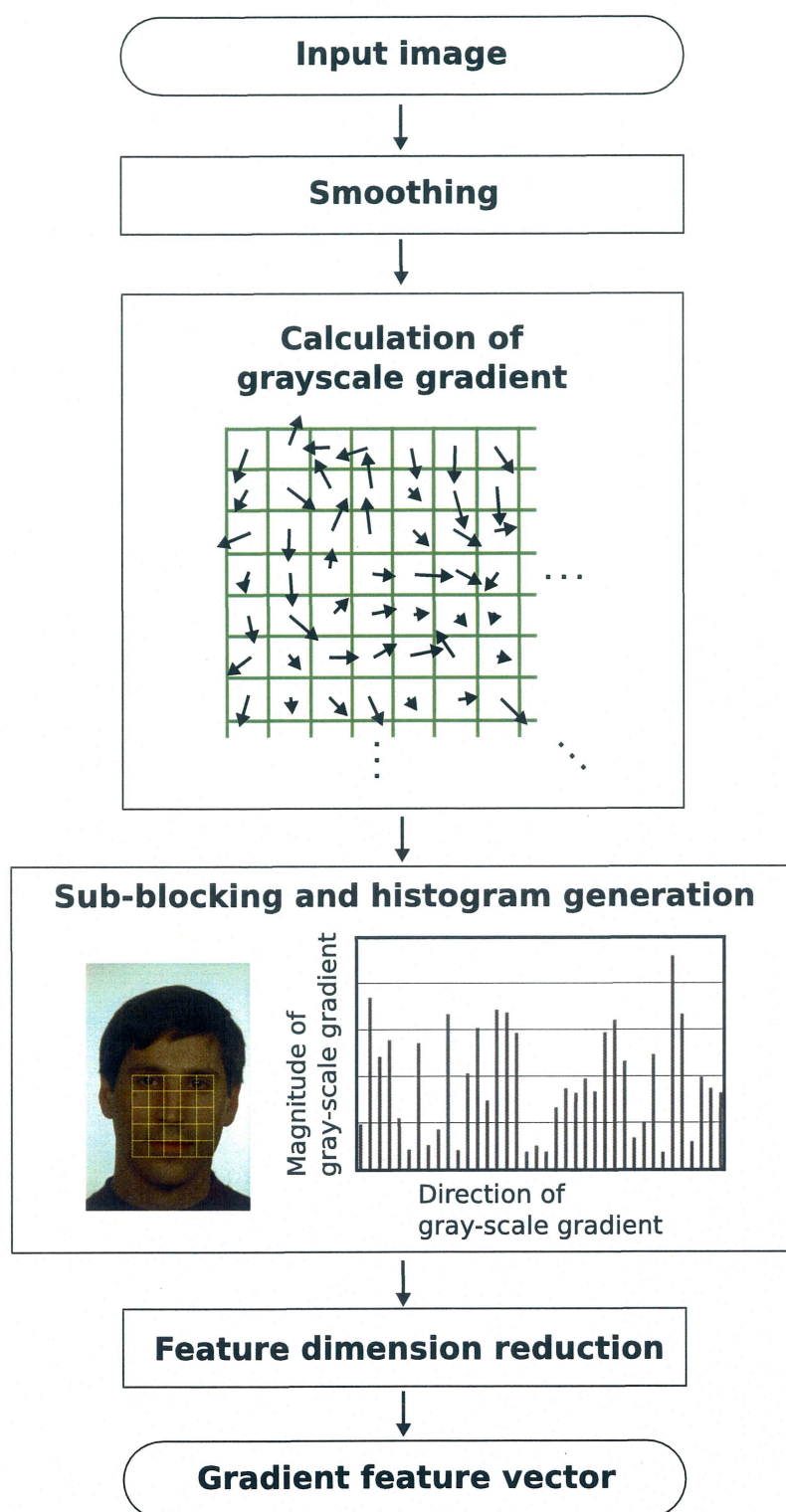


図 2.1: 特徴抽出処理の流れ

2.2.1 平滑化

対象画像 $I = I(p)$ の画素 $p = (p_x, p_y)$ に対してガウシアンフィルタを適用する．ガウシアンフィルタは以下の式で与えられる． N , σ はそれぞれカーネルサイズと重み係数である．なお, N は 5, $\sigma = 1.5$ とする．

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

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

2.2.2 濃度こう配の抽出

カーネルサイズが 3×3 画素の Sobel オペレータを用いて, 濃度こう配の強度と向きとを求める．画素 p における濃度こう配の強度 $|G(p)|$, 向き $\theta(p)$ はそれぞれ以下の式で与えられる．なお, $G_x(p)$, $G_y(p)$ はそれぞれ画素 p における濃度こう配の水平, 垂直方向の成分である．また, こう配の向きの値域は $[-\pi, \pi]$ である．

$$|G(p)| = \sqrt{G_x(p)^2 + G_y(p)^2} \quad (2.3)$$

$$\theta(p) = \tan^{-1} \frac{G_y(p)}{G_x(p)} \quad (2.4)$$

2.2.3 小領域分割と局所方向ヒストグラムの生成

それぞれの画素で算出された濃度こう配の強度と向きとを局所ヒストグラム化するために, 対象領域を小領域に分割する．一般的には, 対象領域を図 2.1 中のように正方格子形に分割してヒストグラム化する手法が用いられる．

2.2.4 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 5×5 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．さらに， L 段階に量子化された濃度こう配の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．これらの手順による加重平均フィルタを用いた特徴の数の削減処理は，特徴ベクトルの要素数を削減するだけでなく，小領域分割と方向量子化とが抽出された特徴に与える影響を低減する効果がある．

第 3 章

回転不変濃度こう配特徴

前節の手順で抽出された濃度こう配特徴は，画像の回転に対して不変ではない．その理由は以下の通りである．まず，物体の回転に伴う濃度こう配の向きの変動が挙げられる．さらに，対象領域を正方格子形に分割する手法では，画像中の物体が回転すると各小領域に含まれる画素が変動し，抽出対象となる局所領域に変化が生じる．この章では，本研究で提案する，回転に対して不変な濃度こう配特徴を抽出する手法について述べる．

3.1 提案手法の概要

提案手法の流れを図 3.1 に示す．提案手法は，従来の濃度こう配特徴の抽出方法（図 2.1）に二つの処理を追加して，回転不変な濃度こう配特徴を抽出する．まず，濃度こう配の向きが物体の回転に対して不変となるように，局所領域の中心点を基準として濃度こう配の向きを変換する．次に，回転に伴う画素の位置変化に対して不変となるように局所領域を小領域分割して，変換後のこう配の方向ごとの局所方向ヒストグラムを生成する．以下では，それぞれの手順について詳細を述べる．

3.2 濃度こう配の向きの変換

各画素におけるこう配の向き $\theta(\mathbf{p})$ を，図 3.2 に示すように特徴抽出領域の中心と各画素を結ぶ線分を基準にした向き $\phi(\mathbf{p})$ に変換する．画素 \mathbf{p} における $\phi(\mathbf{p})$ は以下の式で与えられる．なお， c_x, c_y はそれぞれ特徴抽出領域の中心の x, y 座標である．

$$\phi(\mathbf{p}) = \theta_c - \theta(\mathbf{p}) \quad (3.1)$$

$$\theta_c = \tan^{-1} \left(\frac{p_y - c_y}{p_x - c_x} \right) \quad (3.2)$$

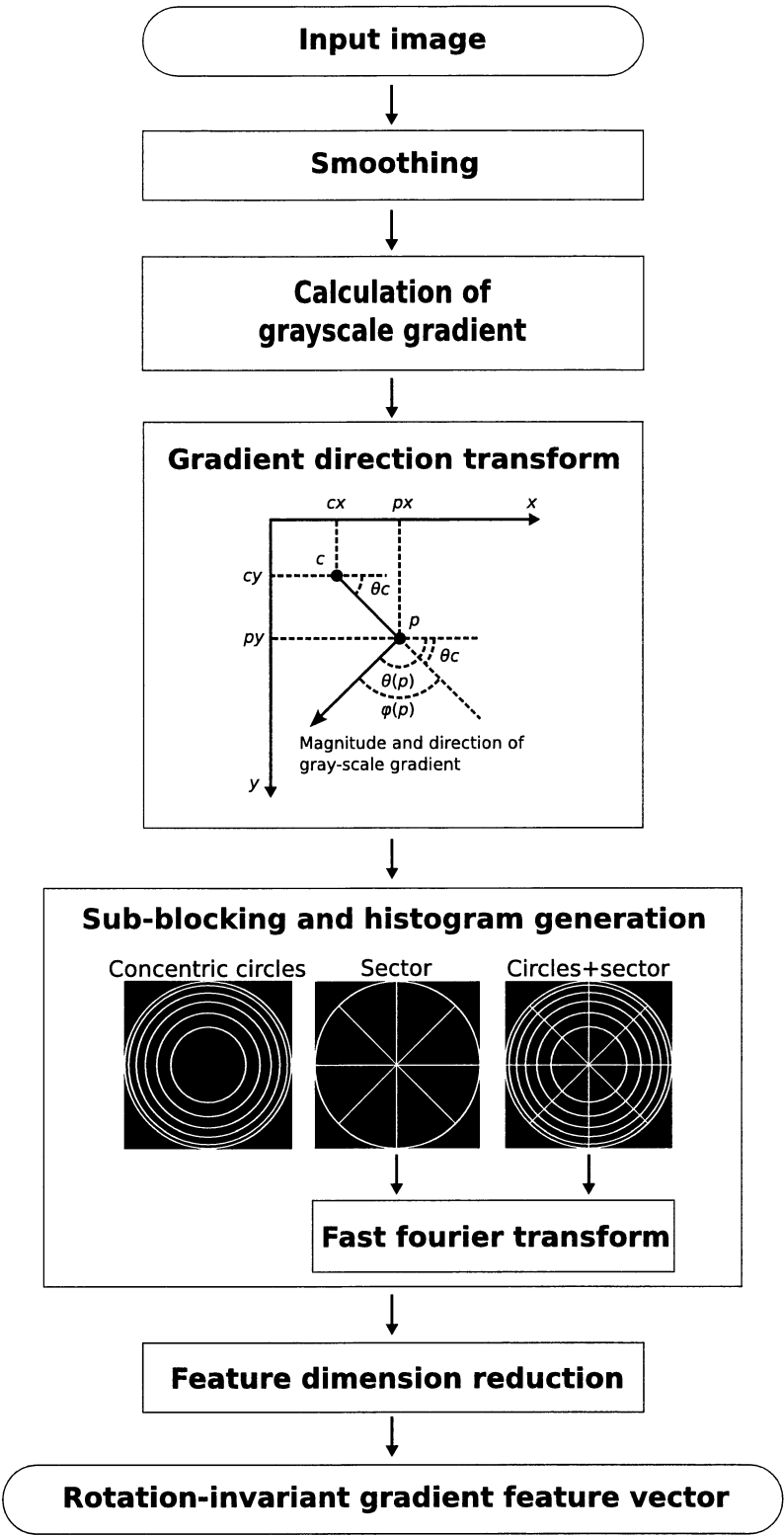


図 3.1: 特徴抽出処理の流れ

変換後のこう配の向き $\phi(\mathbf{p})$ ごとにこう配強度をヒストグラム化することにより画像平面での物体の回転に対し不変な特徴量が得られる。

3.3 回転に対して不変な小領域分割

局所領域を正方格子形に分割する手法では、画像中の物体の回転に伴い画素が移動するため、各小領域で作成される濃度こう配ヒストグラムが変動する。ここでは、物体の回転に伴う画素の位置の変動に対して頑健な小領域分割について述べる。

3.3.1 同心円形領域分割

ヒストグラム生成 (2.2.3) における小領域分割を、図 3.3(a) に示すように、画像を n 個の同心円形の小領域に分割する。このとき、同心円の半径を各小領域の面積が等しくなるように設定する。中心から数えて i 番目の円の半径 r_i は以下の式で与えられる。ただし、 r は最も外側にある円の半径である。

$$r_i = \sqrt{\frac{i}{n}} r \quad (3.3)$$

n 個の同心円形に分割した各小領域ごとに濃度こう配ヒストグラムを作成する。

3.3.2 セクター領域分割

図 3.3(b) に示すように、特徴抽出領域を n 個のセクターで小領域に分割する。ただし、 n は 2 の累乗である。 n 個のセクターに分割した各小領域で変換後のこう配の向き $\phi(\mathbf{p})$ ごとにこう配強度をヒストグラム化する。セクターによる領域分割では、格子形領域分割と同様に画像中の物体が回転すると各小領域に含まれる画素が変動する。そこで、抽出した局所方向ヒストグラムに対して、量子化されたこう配の向き ϕ ごとに円周方向にフーリエ変換を行い、得られた周波数スペクトルを特徴ベクトルとする。

3.3.3 同心円 + セクター領域分割

図 3.3(c) に示すように、特徴抽出領域を m 個の同心円と n のセクターで分割する。各小領域で変換したこう配方向 ϕ ごとにこう配強度をヒストグラム化する。さらに、抽出した特徴ベクトルに対して、量子化されたこう配方向 ϕ ごとに円周方向にフーリエ変換を施すことで画像中の物体の回転に対し不変な濃度こう配特徴を得る。

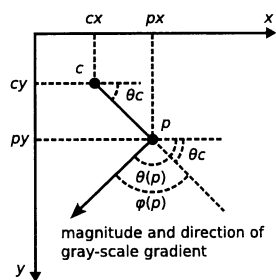


図 3.2: 濃度こう配の向きの変換

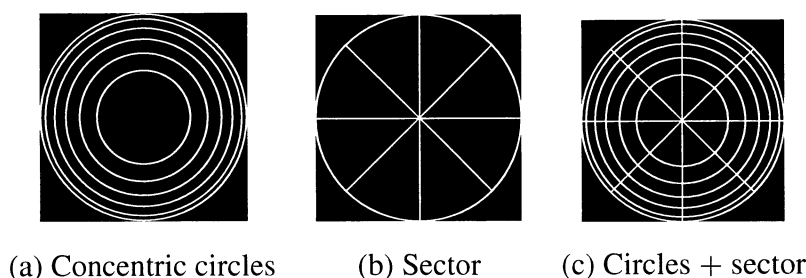


図 3.3: 小領域分割

3.4 回転に対する不変性の評価

回転を伴う画像から、提案した手法で抽出された濃度こう配特徴の実例を図 3.4 に示す。図中 (a) から (c) はそれぞれ図 3.3 の (a) から (c) に対応する。また、図中 (d) は従来の抽出方法で得られた濃度こう配特徴を示す。従来手法である格子状小領域分割 (d) では、抽出された特徴が画像の回転により大きく変動しているが、本研究にて提案した回転に対して不変な抽出手法で抽出された特徴量はそれほど変動がないことが確認できる。

画像の回転に対する不変性を定量的に評価するために、FERET データベース [24] に含まれる正面顔の画像に対して $0^\circ, 5^\circ, 15^\circ, 30^\circ, 45^\circ$ の回転を加え、それぞれからこう配特徴を抽出した。それぞれの回転角度の画像から抽出された特徴ベクトルと、無回転 (0°) の画像から抽出された特徴ベクトルとの類似度を相互相関で評価した。評価には正面顔の画像 400 例を用いた。実験に用いた顔領域の一例を図 3.5 に示す。

算出された相互相関の値を図 3.6 に示す。回転に不変な特徴量の相互相関は、全ての回転角度において 0.85 以上の値を示した。

評価に用いた画像は画像処理により人工的に生成しているため、回転後の補間処理により無回転の画像と完全には一致しない。回転不変特徴の相互相関が、回転角度の増加に伴い低下した原因として、補間処理による画像変化の影響が挙げられる。

一方、格子形領域分割を用いた特徴量は、 15° の時点で 0.5 を下回り、 45° では無相関

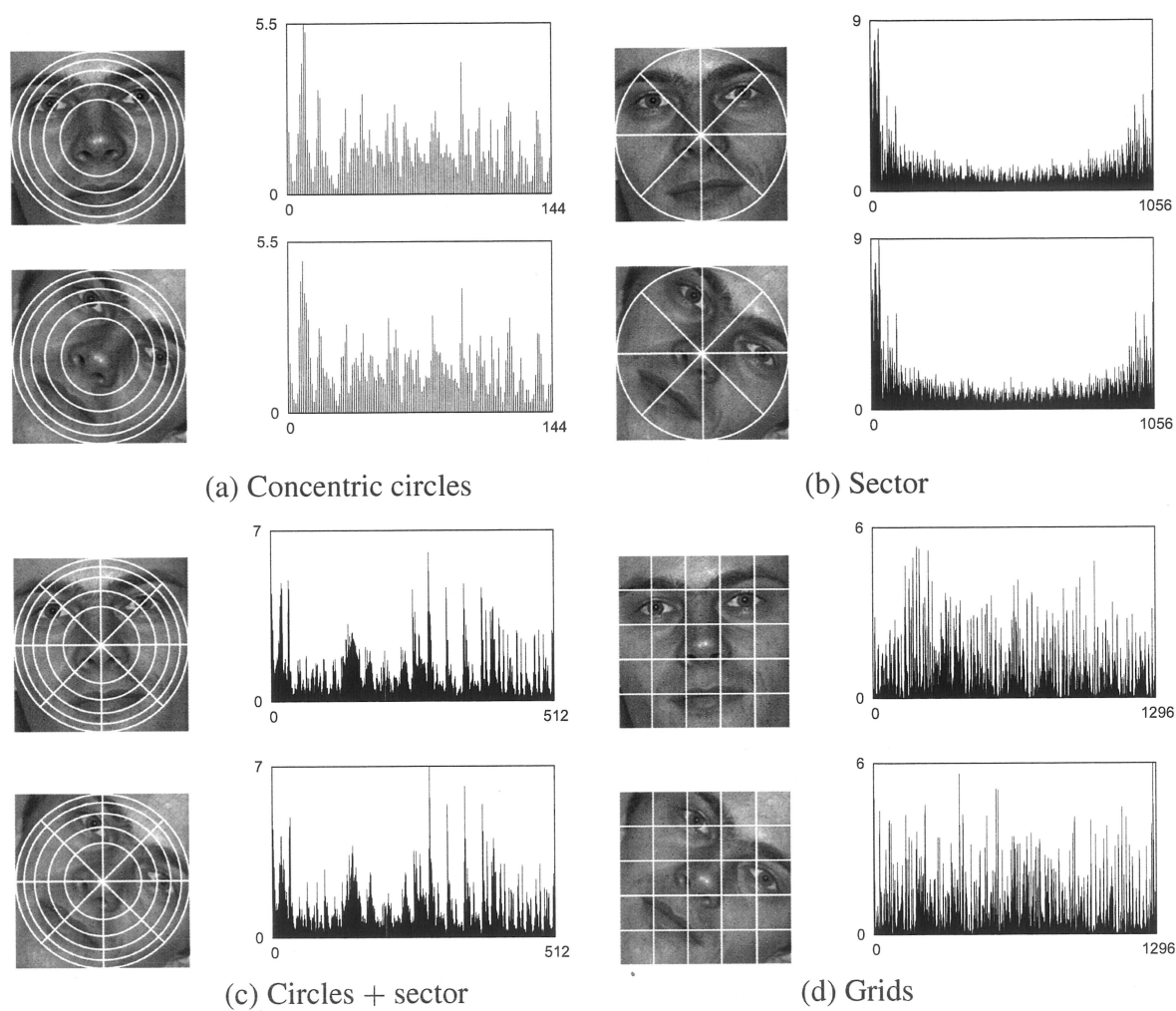


図 3.4: 各手法により抽出された濃度こう配特徴

となることが分かった。

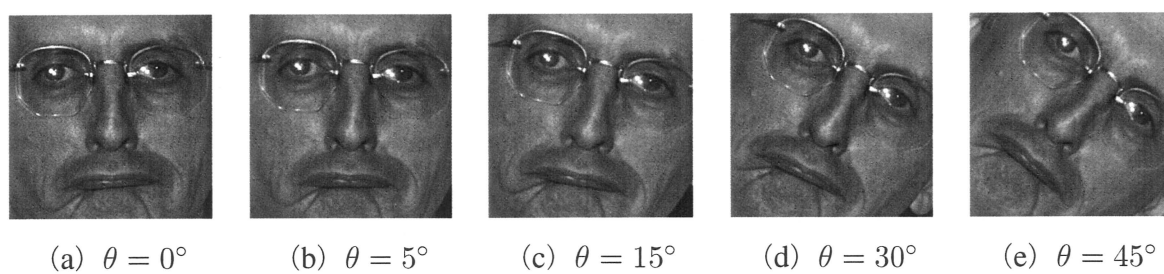


図 3.5: 回転を含む顔画像の一例

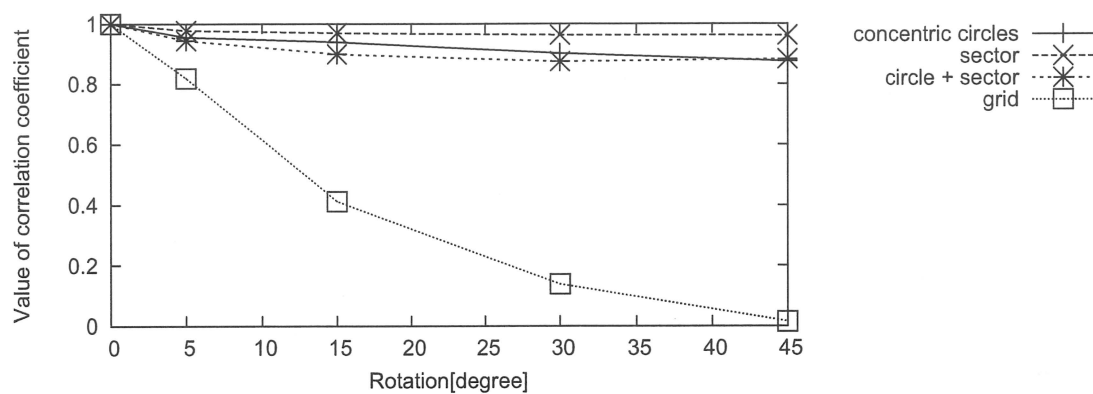


図 3.6: 相互相関係数の変化

第 4 章

顔検出への応用

顔画像の自動認識は，建物への入退出管理，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，ビデオ映像検索質問としての利用など，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いまだ難しい課題と言わざるを得ない．

ここでは，従来の顔検出処理に関する研究について紹介し，この分野のこれまでの研究動向について概説する．本章の内容よりも詳しいサーベイ論文 [6] [7] が発表されているので，それらも参照されたい．

4.1 顔検出とは

顔を含んだ画像，または画像系列は，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．

入力画像に顔が含まれている場合，画像中のすべて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．

4.2 顔検出の代表的な手法

顔が含まれた画像全体から顔のみを含む領域を検出し切り出す処理を，ここでは顔検出と呼ぶ．顔検出は，顔認識における重要な前処理である．ここでは顔検出技術に関するサーベイ論文である文献 [6] で紹介されている研究の一部を紹介する．

4.2.1 事前知識に基づく顔検出

このアプローチでは、顔に関する事前知識に基づく規則によって顔検出手法が定義される。例えば、画像中にそれぞれ対称な位置関係で両目があり、その中心線上に鼻、口が存在する、のように顔部品とそれらの位置関係を説明することの容易性が、この方法が採用された理由である。主な手順として、何らかの処理により顔の候補を複数抽出し、各候補を規則に従って真に顔であるか否かを判定する。これにより擬陽性 (False Positive) 検出を減少できる。

これらの手法の問題は、多くの場合に対して、一般的に有効な規則の作成が極めて困難なことである。具体的には、特定の撮影環境で撮影された画像に対して有効な規則が開発されたとしても、その規則が想定しない範囲の変動が画像中の顔に存在した場合、顔の検出が失敗しやすくなる。

このアプローチの代表的な手法として、Yang らによる Multiresolution rule-based method [8] がある。

4.2.2 特徴に基づく顔検出

事前知識に基づく手法に対して、顔検出に利用できるような顔の不変な特徴を見つけ出す研究が行われてきた。例えば、画像から顔部品や顔に特有な肌色領域を取り出すことで顔を検出する研究である。これらの特徴が得られれば、統計的モデルによりそれらの位置関係などを検証することで、対象領域の顔らしさを評価することができる。

これらの手法の問題点は、このような顔特徴の画像での見え方が光源環境に悪影響を受けやすいこと、顔特徴の境界が顔そのもののそれよりも弱く表現されることなどである。また、顔特徴の検出が顔そのものを検出する問題と同じ困難さを持っていることに注意する必要がある。

このアプローチの代表的な手法として、顔部品を用いる手法 [9] [10] [11]、顔テクスチャを解析する手法 [12]、肌の色を利用する手法 [13] [14] などが存在する。

4.2.3 テンプレートマッチングに基づく顔検出

事前に手動で定義またはパラメータ表現された標準的な顔テンプレートと入力画像とのマッチングにより顔を検出する手法である。入力画像が与えられたときには、例えば、顔輪郭、両目、鼻、口などでそれぞれ独立に標準テンプレートとの相関を計算し、相関の値に基づいて顔の有無を決定する。この手法の利点は実装が容易なことである。しかし、この手法は顔の大きさ、姿勢、形状の変動に対して十分に対応できない場合がある。これら

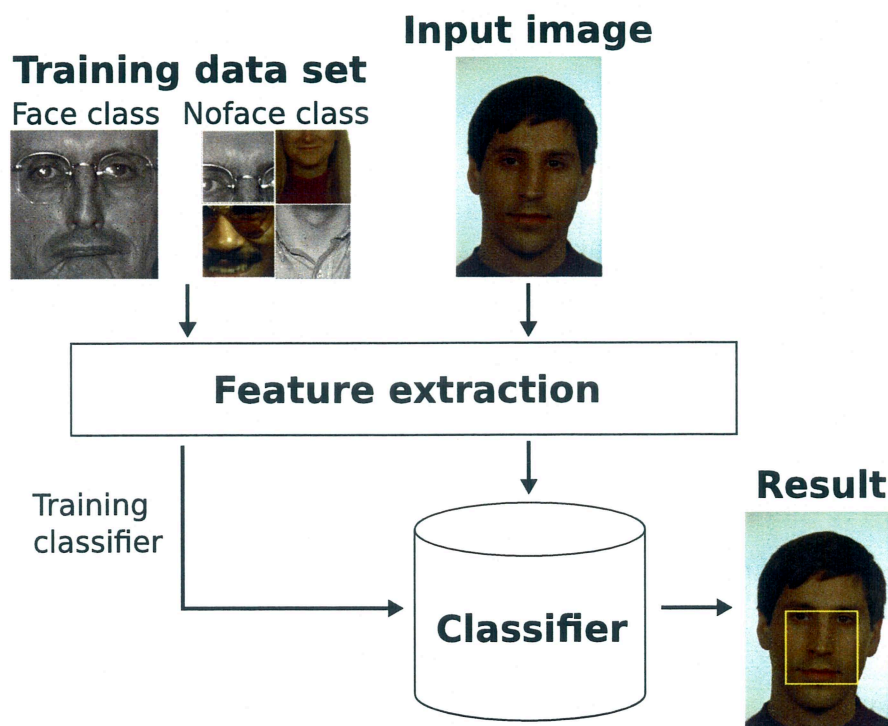


図 4.1: 顔検出処理の流れ

の変動に対応するために多重分解能，多重スケールテンプレートの導入や，可変形状テンプレートなどが提案されている。

このアプローチの代表的な手法としては，形状テンプレート [15]，Active Shape Model [16] などがある。

4.2.4 見え方 (*appearance*) に基づく顔検出

見え方に基づく顔検出手法と，テンプレートマッチングとの大きな違いは，テンプレートが事前に定義されるのではなく，画像中から収集した実例の学習によって得られる点である。基本的に，このアプローチは顔画像と非顔画像との特徴を的確に記述するために，統計的解析と機械学習の手法を用いる。学習された特徴は分布モデルや判別関数の形式で記述され，顔検出に利用される。計算効率と検出性能の向上のために次元削減が行われる。

このアプローチの代表的な例は，主成分分析を利用して顔の見かけ特徴を抽出する Eigenface [17]，ニューラルネットワークを用いる手法 [18] などである。

4.3 本研究における顔検出処理の流れ

本研究における顔検出処理の流れを図 4.1 に示す。以下では本研究における顔検出処理の手順について述べる。

まず、検出に用いる分類器の学習のために学習用画像から特徴を抽出する。学習用画像は、無回転の顔画像 400 枚と無回転の非顔画像 6000 枚からなる。抽出した特徴を用いて分類器を学習させる。ここでの学習とは、分類器に顔クラスと非顔クラスとの境界を決定させることを示す。

次に、入力画像を縮小しながら画像中で特徴抽出領域を走査する。特徴抽出領域から抽出した特徴ベクトルを、分類器によって顔クラスか非顔クラスかに分類する。入力画像全体を走査し終わるまでこの処理を繰り返し、顔クラスに分類された領域の中で最も顔らしい領域を、入力画像中の顔領域として検出する。

4.3.1 サポートベクタマシン (SVM)

本研究では、分類器にサポートベクタマシン (SVM) を用いる。SVM とは、基本的には 2 クラスの分類を行う分類機械である。本研究では、SVM を注目領域が顔か、非顔かを分類するために利用する。SVM は、学習サンプル中でクラス境界近傍に位置する学習サンプル (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平面を求めてクラスの分類を行う。

第 5 章

評価実験

この章では、提案手法を顔検出に応用した場合の性能について述べる。第 3 章で述べた手法を用いて、人物画像から顔領域を検出する実験を行う。

ここではまず、評価に用いる人物画像と評価の方法について述べる。続いて、実験の内容と結果について述べ、章末にて考察する。

5.1 FERET データベース

The Facial Recognition Technology (FERET) データベース [24] 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、背景領域と 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、顔が正面を向いていない人物、首を傾けている人物も含まれる。

本実験では、FERET データベース内の正面顔を含む画像 400 枚に対して角度 θ ($\theta = 0^\circ, 5^\circ, 15^\circ, 30^\circ, 45^\circ$) の回転を加えた計 2000 枚の画像を用意した。画像の一例を図 5.1 に示す。分類器に用いる SVM の設計には、無回転 ($\theta = 0^\circ$) の画像 400 枚を用いた。性能評価には各回転角度ごとに 400 画像、合計で 2000 枚の人物画像を用いた。

5.2 評価方法

顔の検出成功を以下のように定める。

- 顔部品 (左目, 右目, 鼻, 口) の座標すべてが、検出された顔領域に含まれる。
- 検出された顔領域の面積が、真の顔領域の面積の 150% 以下である。

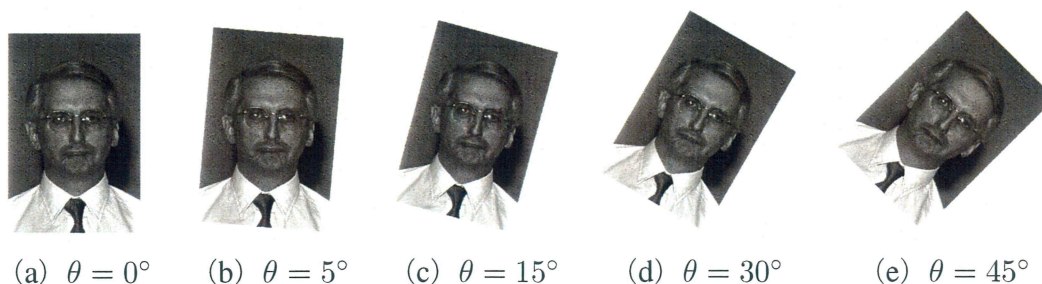


図 5.1: 回転を加えた画像の例



(a) 成功例

(b) 失敗例:顔部品の座標が
検出された領域に含まれていない(c) 失敗例:検出された領域の面積が
真の面積の 150%より大きい

(d) 失敗例:検出された領域がない

図 5.2: 評価の例

この2つの条件の両方を満たす場合を検出成功 (図 5.2 (a)) とする。顔部品の座標が検出された領域に含まれていない場合 (図 5.2 (b)) や、検出された領域の面積が、真の顔領域の面積の 150% よりも大きい場合 (図 5.2 (c)) は検出失敗とする。真の顔領域は、左目、右目、口の座標を頂点とする三角形の外接円とした。さらに、本実験では1つの画像につき1名の人物が含まれている画像のみを用いるため、顔領域として検出された領域がない場合 (図 5.2 (d)) も検出失敗とする。

本実験における評価の指標として用いる検出成功率を以下に定義する。

$$\text{検出成功率 (\%)} = \frac{\text{検出成功の定義を満たした画像数}}{\text{評価用画像の数}} \quad (5.1)$$

5.3 性能評価実験

5.3.1 分類実験

顔検出における提案手法の性能評価のために、回転を伴う顔画像と非顔画像との分類実験を行った。分類実験の流れ図を図 5.3 に示す。分類器には **SVM** を使い、学習用画像には、回転を含まない $\theta = 0^\circ$ の画像 400 例を用いた。評価用画像には、各回転角度ごとに顔画像 400 例、非顔画像 6000 例を用いた。

実験結果を図 5.4 に示す。図の縦軸と横軸とはそれぞれ、分類に成功した画像の割合と評価画像に加えた回転角度とを示す。同心円形領域分割を用いた場合では、全ての回転角度において 94.3%以上の分類成功率が得られた。セクターによる領域分割では 91.8%以上の分類成功率が、同心円形 + セクターによる領域分割手法では 93.5%以上の分類成功率がそれぞれ得られた。セクター領域分割と同心円形 + セクター領域分割では、 $\theta = 30^\circ$ の結果と比較して、 $\theta = 45^\circ$ の分類成功率が増加した。従来手法を用いた分類実験では、 $\theta = 0^\circ, 5^\circ$ の画像に対しては高い分類性能が得られたが、 $\theta = 30^\circ, 45^\circ$ の画像では分類不能になった。

実験の結果から、同心円形分割による特徴抽出は、画像中の顔領域の回転に依存しない特徴ベクトルを抽出できることが確認された。セクターによる領域分割や同心円形 + セクターによる領域分割手法では、回転角度によって分類性能が低下することがわかった。この 2 手法は、回転不変な特徴ベクトルを得るために抽出した特徴ベクトルに対して **FFT** を施す。しかし、セクターを用いて領域を分割した場合、画像が回転すると各小領域が含む画素の範囲が変化するため、回転角度によっては分類性能が低下したと考える。

5.3.2 検出実験

顔検出における提案手法の有効性を示すために、**SVM** を用いた顔検出実験を行った。顔検出実験の流れ図を図 5.5 に示す。学習用画像には、 $\theta = 0^\circ$ の画像 400 枚を用いた。評価用画像には、各回転角度ごとに 400 画像を用いた。本実験では、以下に示す条件をそれぞれ変化させて顔検出処理を行い、提案手法を用いた顔検出の性能評価を行った。

- (1) **SVM** のカーネル関数
- (2) フーリエ変換
- (3) 周波数スペクトルの対数変換
- (4) 特徴ベクトルの結合

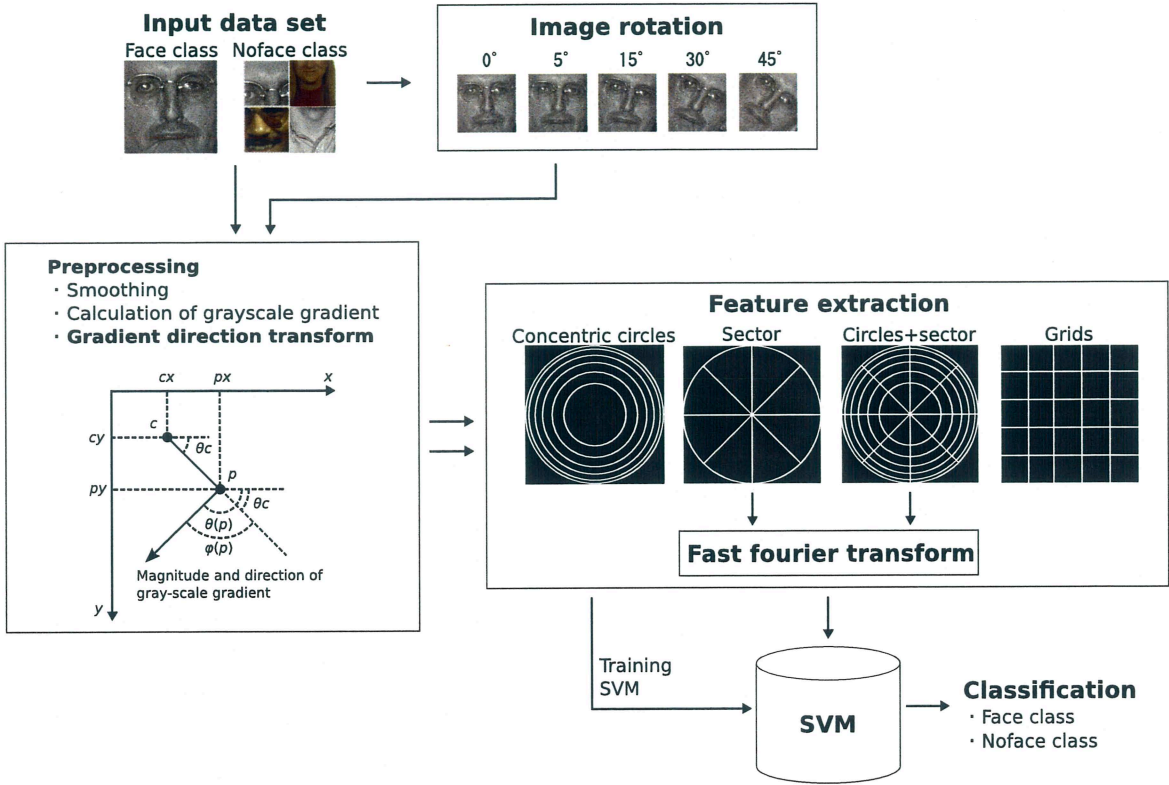


図 5.3: 分類実験の流れ

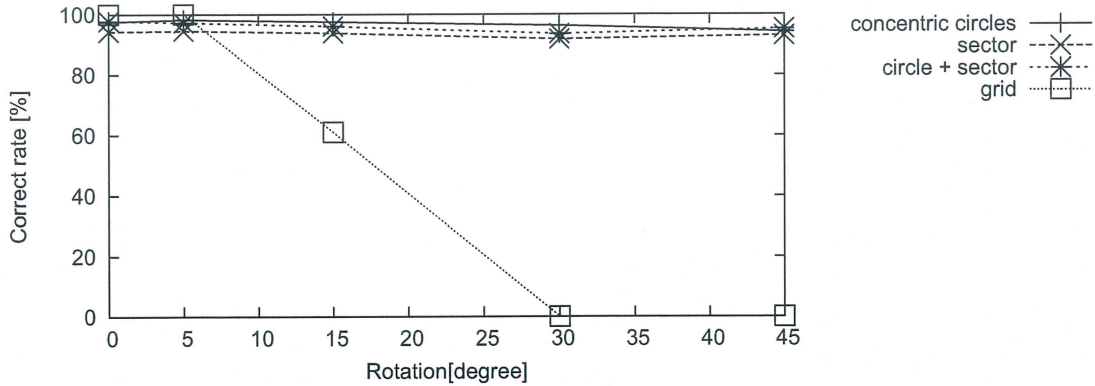


図 5.4: 分類実験の結果

classification of face class and noface class using SVM

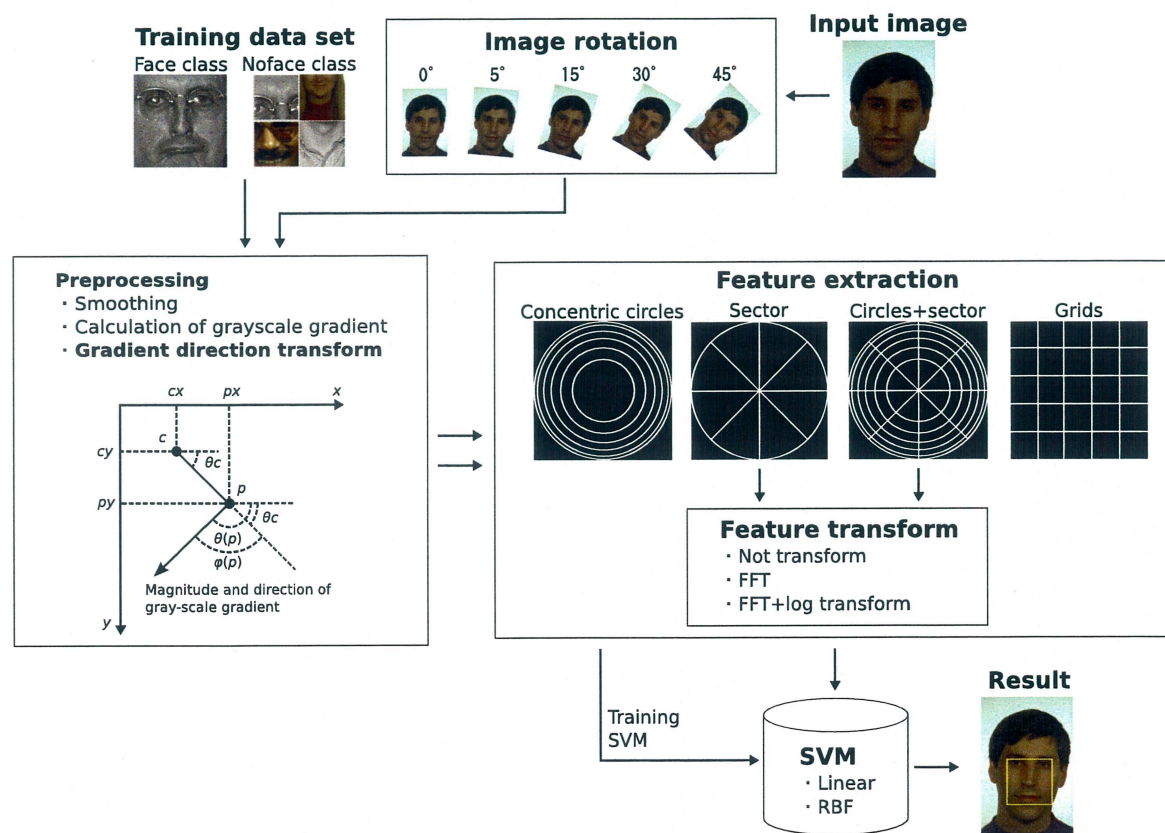


図 5.5: 顔検出実験の流れ

(1) SVM のカーネル関数

本研究で分類器として用いる SVM は、カーネルトリックと呼ばれる特徴変換手法により、原特徴ベクトルを一般的により高い次元の特徴空間に射影した後で線形識別する。ここでのカーネル関数には、様々なものが提案されているが、代表的なものとして、線形関数、RBF (Radial Basis Function) がある。線形関数、RBF はそれぞれ以下の式で与えられる。

$$K_l(u, v) = u^T v \quad (5.2)$$

$$K_r(u, v) = \exp\left(-\frac{|u - v|^2}{\sigma^2}\right) \quad (5.3)$$

本研究では、カーネル関数に線形関数、RBF を用いた SVM をそれぞれ線形 SVM と非線形 SVM と呼ぶ。本実験では、これらの SVM それぞれを用いて顔検出処理を行い、提案手法の性能を評価した。

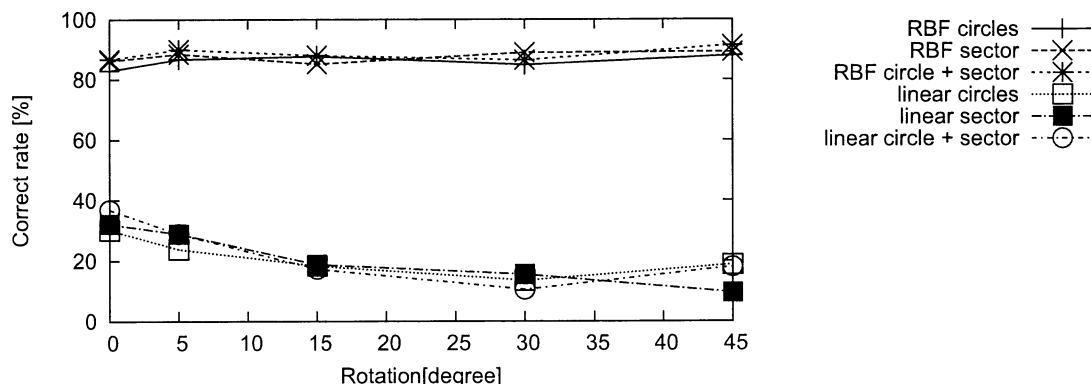


図 5.6: 実験結果 (SVM のカーネル関数の違いによる性能変化)

RBF SVM v.s. linear SVM

実験結果を図 5.6 に示す．本実験では，非線形 SVM を用いた場合の検出成功率が，線形 SVM を用いた場合と比べて，すべての手法で上回った．非線形 SVM を用いた顔検出処理で，全ての角度の画像から顔領域を検出した場合の成功率は，同心円形領域分割で 86.1%，セクター領域分割で 87.6% あった．同心円 + セクター領域分割を用いた場合には，本実験で最も高い 88.5% の検出成功率が得られた．一方，線形 SVM を用いた場合の検出成功率は，どの手法でも約 21.4% 程度であった．また，実験の結果から，どちらのカーネル関数を用いた場合でも，提案手法を用いた顔検出処理は画像中の顔の回転に対して不変性を持っていることがわかった．

実験の結果から，以降の実験はすべて非線形 SVM を用いて顔検出処理を行った．

(2) フーリエ変換

セクター領域分割と同心円 + セクター領域分割では，回転不変な特徴量を抽出するために濃度こう配ヒストグラム作成後，フーリエ変換を施している．フーリエ変換で特徴ベクトルの回転不変性が確保されることを示すために，この 2 つの領域分割手法について検証を行った．分類器には，RBF カーネルの非線形 SVM を用いた．

実験結果を図 5.7 に示す．実験の結果，どちらの領域分割手法でもフーリエ変換を行うことで，回転不変なこう配特徴が抽出できていることが確認された．フーリエ変換なしの場合，無回転の画像に対しては高い検出成功率が得られるが，回転角度が 30°, 45° の顔領域を検出することは困難であることがわかった．

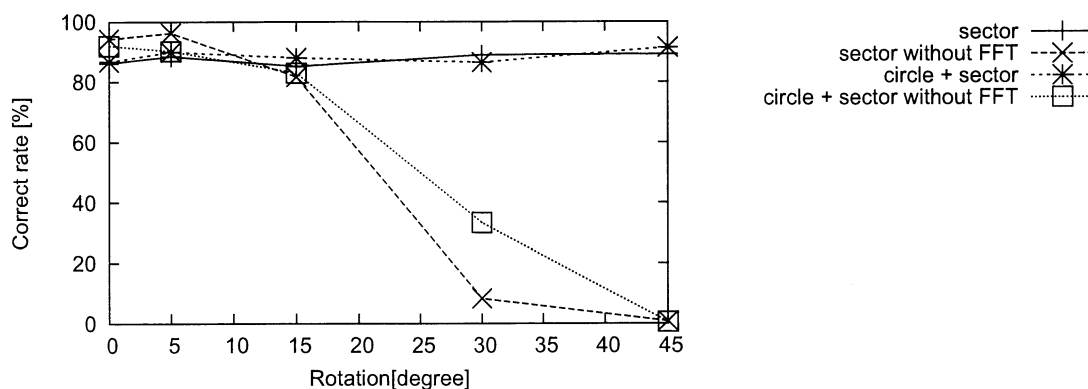


図 5.7: 実験結果 (フーリエ変換)

evaluation of fast Fourier transform for rotation-invariant gradient features

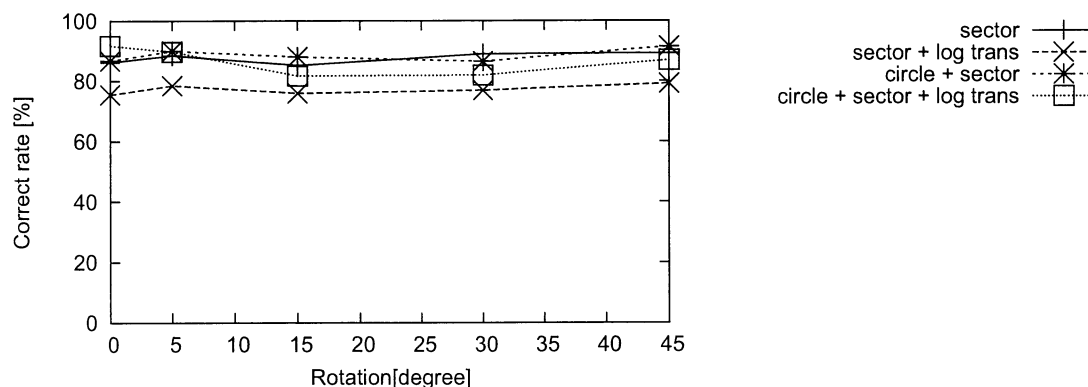


図 5.8: 実験結果 (対数変換)

evaluation of log transformation for sector method and circles + sector method

(3) 周波数スペクトルの対数変換

3.4 における回転に対する不変性の評価から、顔領域の濃度こう配ヒストグラムには、各方向ごとの強度でばらつきがあることがわかる。作成された濃度こう配ヒストグラムに対して対数変換を施すことで方向ごとの強度のばらつきを抑え、無回転の顔と回転した顔との特徴間の相関がより強まることが期待できる。

実験結果を図 5.8 に示す。対数変換を施した場合、対数変換を施さない場合に比べて検出成功率が全体的に低下することがわかった。この時、セクター領域分割で 10.4%、同心円 + セクター領域分割で 4.1% 検出成功率が低下した。

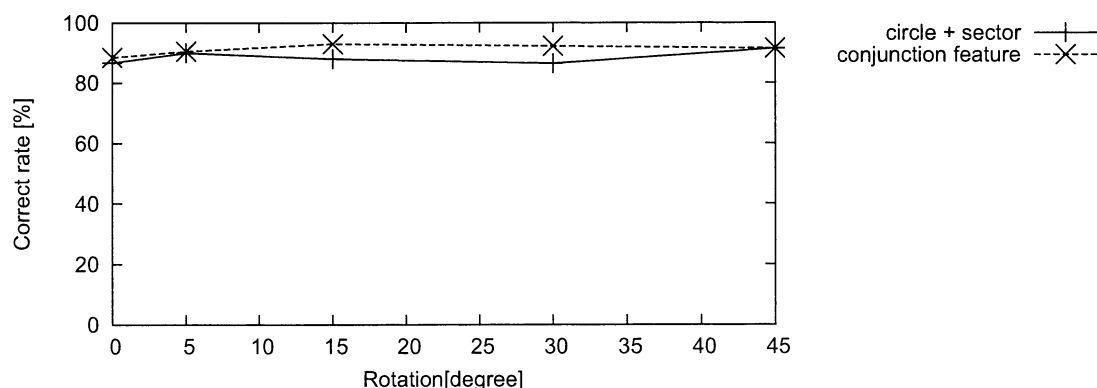


図 5.9: 実験結果 (特徴ベクトルの結合)

comparison circles + sector method to sum circles method and sector method

(4) 特徴ベクトルの結合

予備実験より、同心円 + セクター領域分割では、局所領域を 64 よりも細かく分割した場合に検出性能が低下することがわかっている。これは、各小領域に含まれる画素が分割数に反比例して少なくなるため、各ヒストグラムで表現できる特徴が低減したためであると考えられる。対象領域をより表現する特徴ベクトルを得るために、同心円形領域分割とセクター領域分割とのそれぞれで得られたこう配ヒストグラムを結合した回転不変濃度こう配特徴を用いて検出実験を行い、同心円 + セクター領域分割を用いた場合との性能比較を行う。便宜上、こう配ヒストグラムの和を用いた特徴ベクトルを **Conjunction feature** と呼ぶ。

実験結果を図 5.9 に示す。Conjunction feature を用いた顔検出処理の性能は、すべての回転角度の画像を用いた場合 91.1%であった。実験の結果から、Conjunction feature の顔検出性能は、同心円 + セクター領域分割の検出性能と比較して、約 2.6%高いことがわかった。

5.4 考察

対数変換の効果を検証する実験では、期待した効果が得られないことが確認された。対数変換を施した場合、特徴内での強度のばらつきを抑えることが可能であるが、同時に対象領域を表現するための特徴を低減した可能性がある。その結果、顔領域の特徴と非顔領域の特徴との相関が強まり、検出成功率が低下したのではないかと考える。

また、結合した特徴ベクトルを用いる実験では、同心円 + セクター領域分割の結果と

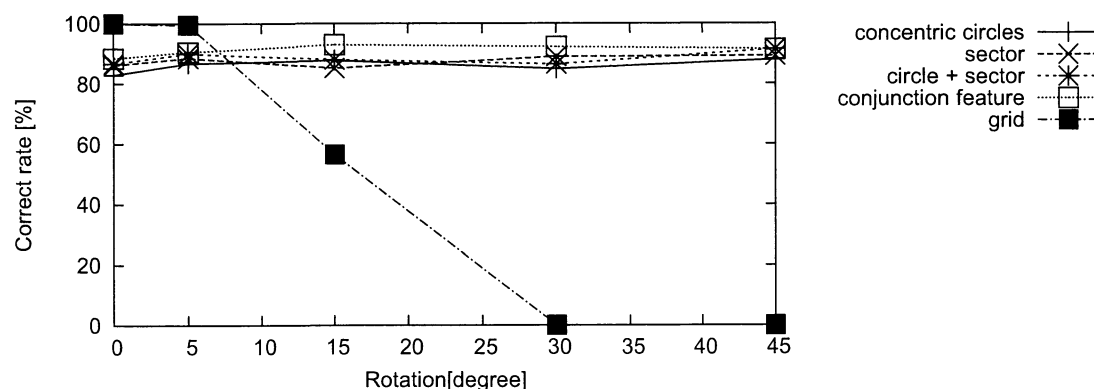


図 5.10: 各手法における検出成功率

correct rate of face detection for each methods

比較して、**Conjunction feature** は約 2.6% 高い 91.1% の検出成功率を得た。同心円形領域分割で得られたこう配ヒストグラムとセクター領域分割で得られたこう配ヒストグラムとの和を特徴ベクトルとして用いることで、同心円 + セクター領域分割を用いる場合よりも、対象領域を表現する特徴が得られたと考えられる。

各領域分割手法を用いた顔検出処理で、最も高い検出成功率を得られた実験結果を図 5.10 に示す。グラフから、提案手法を用いた顔検出処理は画像中の顔の傾きに頑健であることが確認できた。このとき、**Conjunction feature** を用いた顔検出処理の性能が最も高く、検出成功率 91.1% を得られた。一方、従来の格子形領域分割を用いた顔検出処理は、無回転の顔を検出する性能は高いが、回転した顔を検出することは困難であることがわかった。

また、提案手法を用いて無回転の顔を検出する性能は、従来手法と比べて、約 11.5% 低いことが確認された。回転に不変な濃度こう配特徴を抽出する手順の中で、対象を表現する特徴を低減してしまった可能性がある。

第 6 章

結言

6.1 まとめ

本研究は，回転に不変な濃度こう配特徴を抽出する手法を提案した．濃度こう配特徴はこれまでに対象物体の拡大，縮小に対して不変な特徴抽出方法が提案されてきたが，画像平面上の回転に対して不変な抽出方法は十分に検討されていなかった．本研究は，濃度こう配特徴の応用と回転不変特徴を抽出手法とを新たに提案した．提案手法を用いて無回転の画像と回転を含む画像とのそれぞれから抽出した特徴ベクトルの相互相関が強いことから，提案手法を用いて回転不変な濃度こう配特徴が得られることがわかった．

また，本研究では提案手法を顔検出へと応用した．実験の結果から，提案手法を用いた顔検出処理は，画像中の顔領域の回転に依存しないことが確認された．提案手法の実現は，自動視点顔検出などの回転による姿勢変化を含む物体認識，検出の研究分野へのアプローチの簡単化に寄与すると考える．

6.2 今後の課題

回転を伴う対象として図面中の文字認識や一般物体検出などの分野に提案手法を応用することである．また，他の回転不変特徴 [3] [4] [5] との比較をすることで，提案手法の有効性を示したい．

付録 A

特徴抽出におけるパラメータの決定

5.3 における濃度こう配の向きの量子化数と各手法ごとの領域分割数は，予備実験により求めた最適値に設定されている．この章では，パラメータ決定のための予備実験の概要と結果について述べる．

A.1 予備実験

各手法における特徴抽出のパラメータを決定するために，SVM を用いた顔検出処理を行った．予備実験の流れ図を図 A.1 に示す．SVM の学習には，無回転の人物画像 200 枚を用いた．評価は，各回転角度ごとに 200 画像を用いて計 1000 枚で行った．予備実験で用いた画像と評価基準の詳細については，それぞれ 5.1 と 5.2 を参照されたい．

A.2 実験結果

予備実験の結果を表 A.1 に示す．表中の数字は，方向量子化と領域分割数との組み合わせを用いた場合の検出成功率を示す．予備実験の結果から選択した性能評価実験（5.3）におけるパラメータの組み合わせを表 A.2 に示す．

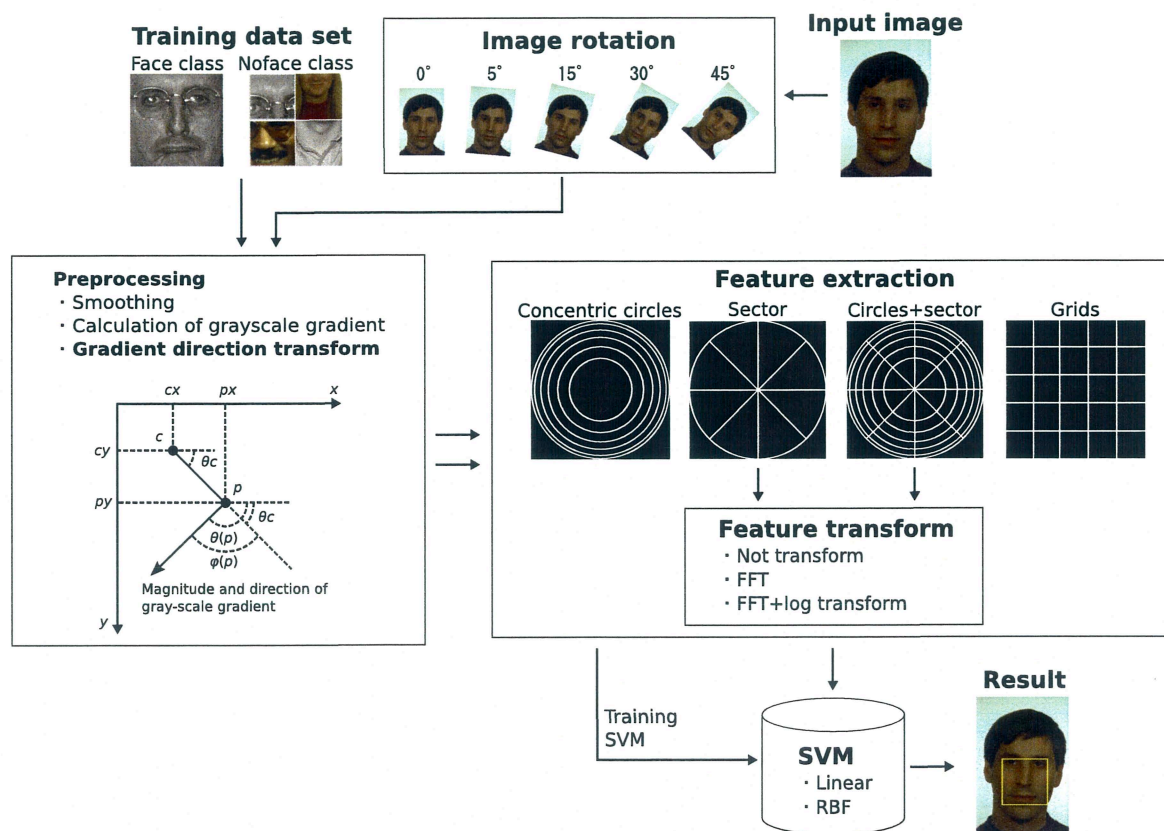


図 A.1: 予備実験の流れ

表 A.1: 予備実験の結果

(a) 同心円形領域分割

| | | number of sub-blocks | | | | |
|--|----|----------------------|------|------|------|------|
| | | 9 | 11 | 13 | 15 | 17 |
| number of direction quantization | 16 | 70.4 | 74.0 | 70.6 | 71.5 | 73.0 |
| | 32 | 77.3 | 80.8 | 80.0 | 84.6 | 85.2 |
| | 64 | 73.4 | 75.1 | 77.8 | 82.2 | 82.2 |

(b) セクター領域分割

| | | number of sub-blocks | | | | |
|--|----|----------------------|------|------|------|------|
| | | 4 | 8 | 16 | 32 | 64 |
| number of direction quantization | 16 | 56.1 | 65.7 | 75.0 | 75.8 | 73.9 |
| | 32 | 60.6 | 69.9 | 80.7 | 83.8 | 84.3 |
| | 64 | 56.5 | 70.6 | 81.3 | 85.7 | 85.8 |

(c) 同心円+セクター領域分割

| | | number of sub-blocks | | |
|--|----|----------------------|------|-------|
| | | 4x4 | 8x8 | 16x16 |
| number of direction quantization | 16 | 52.9 | 56.4 | 17.0 |
| | 32 | 60.3 | 77.9 | 51.2 |
| | 64 | 58.9 | 86.6 | 79.6 |

(d) conjunction feature

| | | number of sub-blocks | | |
|----------------------------|----|----------------------|-------|-------|
| | | 17x16 | 17x32 | 17x64 |
| num of dir quantization | 32 | 91.1 | 72.5 | 59.5 |
| | 64 | 91.2 | 83.1 | 55.5 |

表 A.2: 性能評価実験で用いたパラメータの組み合わせ

| type of sub-blocks | dir quantization | sub-block num | dimension |
|---------------------|------------------|---------------|-----------|
| concentric circles | 32 | 17 | 144 |
| sector | 64 | 64 | 1056 |
| circles+sector | 64 | 8x8 | 512 |
| conjunction feature | 64 | 17x16 | 576 |

付録 B

プログラムソースリスト

作成したプログラムを以下のディレクトリ

yokota/MieUFaceRecog/

に置く．ディレクトリ構造と各プログラムの概略を以下に示す．

```
|-- GFVextraction
|   |-- featext
|   # 顔, 非顔の画像と正解データを読み込み,
|   # 抽出した特徴を dat ファイルに記述する.
|
|-- assess
|   |-- assess_result
|   # 検出結果が記述された txt ファイルと正解データを読み込み,
|   # プログラム内部の正解判定に基づいて正誤を決定する.
|
|-- doc
|   # 本研究で用いたプログラムやライブラリを導入するための
|   # 手順や, 顔検出プログラムを実行するための注意事項などが
|   # 記述された README を含む.
|
|-- facedetSVM
|   |-- src
|       |-- facedetSVM
|   # 顔検出処理のメインプログラム.
```



```
|
|-- lib
|   |-- feret_process.cpp
|   # 回転不変濃度こう配特徴を抽出するコードを含む.
|
|-- libsvm-2.82
|   |-- 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
|   # 顔分類実験を行った.
|
|-- misc
|   |-- crop_part
|   # 人物画像の txt ファイル (顔部品の座標などが記述されたファイル)
|   # を読み込み, 顔領域の geometry を返す.
|
|-- script
|   # 分類器の学習, 顔検出, 正誤判定などを行うための
|   # sh ファイルを含む.
|
|-- tool
|   # 各ディレクトリで生成された実行形式ファイルを含む.
|   # MieFaceRecog/にて, make clean, make, make install
|   # と順番に実行することで, このディレクトリにコピーされる.
```

付録 C

画像ファイルソースリスト

FERET データベースの画像ファイルとその回転画像とを laputa のローカル

/home/fafb_color/

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．

```
|-- theta_000
|   |-- fafb_list_000.txt
|   | # 0° の画像リストが記述されたファイル.
|   |
|   |-- groud_truth
|   | # 画像ファイルの顔部品などが記述されたデータ.
|   |
|   |-- image
|   | # 人物画像を含む.
|   |
|   |--svm_train_img
|   | # 各ディレクトリのデータを顔検出処理用にまとめたデータ.
|   | # *.png : 人物画像
|   | # *.txt : ground_truth
|   | # *_noface.txt : 学習で抽出する非顔領域の geometry
|   |
|   |--true
|   | # 正解領域の geometry と顔部品の座標が記述されたデータ.
|
```

```
|-- theta_005
|   |--fafb_list_005.txt
|   |--ground_truth
|   |--image
|   |--svm_train_img
|   `--true
|
|-- theta_015
|   |--fafb_list_015.txt
|   |--ground_truth
|   |--image
|   |--svm_train_img
|   `--true
|
|-- theta_030
|   |--fafb_list_030.txt
|   |--ground_truth
|   |--image
|   |--svm_train_img
|   `--true
|
|-- theta_045
|   |--fafb_list_045.txt
|   |--ground_truth
|   |--image
|   |--svm_train_img
|   `--true
|
|-- experiment_database
|   # 上述したデータ内から、本実験で用いたデータ群のみを
|   # 抽出したデータベース
|
|   |-- ex_list_*.txt
|   # 本実験で用いた画像リストが記述されたファイル.
|
```

```
`--theta_*  
# 上述した svm_train_img/内のファイルから実験で用いた  
# ファイルのみを抽出したデータ.
```

付録 D

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などを御教授下さった三重大学大学院工学研究科情報工学専攻教授木村文隆先生に深く感謝します。研究に関する多くの知識や研究分野における様々な技術を御教授くださった准教授若林哲史先生に深く感謝します。資料作成からプレゼンテーションのコツなどをご指導いただきました助教大山航先生に深く感謝します。専門的な観点から深く幅広いアイデアや知識を御教授下さった三重大学名誉教授三宅康二先生に深く感謝します。

また、日頃いろいろとお世話になった田中みゆき事務、多くの有益なアドバイスを下さった研究室の同期の皆様に感謝します。研究室での生活を支え、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祖父母、両親、兄達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] 黄瀬 浩一, 岩村 雅一, 中居 友弘, 野口 和人. 局所特徴量のハッシングによる大規模画像検索. *DBSJ Journal*, Vol.8, No.1, Jun 2009.
- [2] Chang HUANG, Haizhou AI, Yuan LI, and Shihong LAO. Vector boosting for rotation invariant multi-view face detection. In *ICCV*, 2005.
- [3] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Journal of Computer Vision*, 60, 2, pp.91–110, 2004.
- [4] 松浦 武信, 森 和也. 回転不変特徴を用いた印鑑照合の一方法について. 東海大学紀要電子情報学部, Vol.3, No.1, pp.29–34, 2003.
- [5] 栗田 多喜夫, 堀田 一弘, 三島 健稔. Log-Polar 画像の高次局所自己相関特徴を用いた大きさ・回転に不変な 2 次元図形及び顔画像の認識. 電子情報通信学会パターン認識・メディア理解研究会, PRMU96–212, pp.151–158, 1997.
- [6] M. H. Yang, D. J. Kriegman and N. Ahuja. Detecting Faces in Images: A Survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.24, No.1, pp.34–59, Jan 2002.
- [7] W. Zhao, R. Chellappa, P. Phillips and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, Vol.35, No.4, pp.399–458, 2003.
- [8] Guangzheng Yang and Thomas S. Huang. Human face detection in a complex background. *Pattern Recognition*, Vol.27, No.1, pp.53–63, 1994.
- [9] K. C. You and R. Cipolla. Feature-based human face detection. *Image and Vision Computing*, Vol.15, No.9, pp.713–735, 1997.
- [10] M. Hamouz, J. Kittler, J. -K. Kamarainen, P. Paalanen, H. K”alvi”anen and J. Mates. Feature-based affine-invariant localization of faces. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol.27, No.9, pp.1490–1495, 2005.
- [11] O. Jesorsky, K. J. Kichberg 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.

- [12] Ying Dai and Yasuaki Nakano. Face-texture model based on sgld and its application in face detection in a color scene. *Pattern Recognition*, Vol.29, No.6, pp.1007–1017, 1996.
- [13] X. Chen, P. J. Flynn and K. W. Bowyer. Fully automated facial symmetry axis detection in frontal color images. In *4th IEEE Workshop on Automatic Identification Advanced Technologies*, pp.106–111, 2005.
- [14] P. S. Hiremath and Ajit Danti. Detection of multiple facecs in an image using skin color information and lins-of-separability face model. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.20, No.1, pp.39–61, 2006.
- [15] I. Craw, D. Tock and A. Bennett. Finding face features. In *ProceedingsofSecondEuropeanConf.ComputerVision*, pp.92–96,1992.
- [16] A. Lanitis, C. J. Taylor and T. F. Cootes. An automatic face identification system using flexible appearance models. *Image and Vision Computing*, Vol.13, No.5, pp.393–401, 1995.
- [17] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, Vol.3, No.1, pp.71–86, 1991.
- [18] H. A. Rowley, S. Baluja and T. Kanade. Neural network-based face detection. *IEEE Transactionsof Pattern Analysis and Machine Intelligence*, Vol.20, No.1, pp.23–38, 1998.
- [19] N. Dalal and B. Triggs. Histograms of oriented gradient features for human detection. In *Proc.of IEEE Conference on CVPR*, pp.886–893, 2005.
- [20] 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.77, No.10, pp.2046–2053, Oct 1994.
- [21] H. Hongo, M. Hoya, M. Yasumoto, Y. Niwa and K. Yamamoto. Focus of attention for face and hand gesture recognition using multiple cameras. In *Proc.IEEEInt'lConf.AutomaticFaceandGestureRecognition*, pp.156–161, 2000.
- [22] 藤吉 弘亘. Gradient ベースの特徴抽出 -SIFT と HOG-. 情報処理学会研究報告. CVIM, Sep 2007.
- [23] 澤 和宏, 若林 哲史, 鶴岡 信治, 木村 文隆, 三宅 康二. こう配特徴ベクトルと変動吸収共分散行列による手書き漢字認識の高精度化. 電子情報通信学会論文誌, Vol.J84–D–II, No.11, pp.2387–2397, Nov 2001.

- [24] P.J. Phillips, H. Wechsler, J. Huang and P. Rauss. The FERET database and evaluation procedure for face recognition algorithms. *ImageandVisionComputingJ*, Vol.16, No.5, pp.295–306, 1998.

回転不変濃度こう配特徴とその顔検出への応用

Rotation-Invariant Gradient Features and Its Application for Face Detection

三重大学大学院 工学研究科 情報工学専攻
ヒューマン・インタフェース研究室
横田智彦



はじめに

- 画像を用いた物体の認識・検出は、
 - 検索質問を画像とする大規模画像検索[1]
 - 特定物体の認識・検出[2]
 - 生体機能である視覚のモデル化
 など広く応用・研究がなされている。

[1] 黄漢 浩一, 岩村 雅一, 中庭 友弘, 野口 和人, 局所特徴量のハッシングによる大規模画像検索, DBSJ Journal, Vol. 8, No. 1, Jun 2009.
[2] 松浦 武雄, 森 和也, 回転不変特徴を用いた印刷照合の方法について, 東海大学紀要電子情報学部, Vol. 3, No. 1, pp. 29-34, 2003.



2

課題

- 制約のない環境下で取得された画像中から特定の物体を認識・検出する問題は、以下に示す困難さを含む。
 - 物体の位置
 - 物体の姿勢
 - 物体の個数
 - 撮影環境
 - etc



3

物体の姿勢

- 姿勢の変化は、主に物体の回転に基づく。
 - 画像中の物体の回転は、画像平面上の回転(RIP)と画像平面外の回転(ROP)とに分離できる。



無回転の人物画像



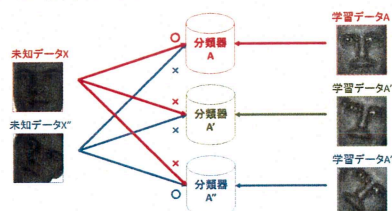
RIPを含む人物画像



4

回転を含む物体の検出・認識アプローチ

- 回転を含む物体を検出・認識するためには、複数の分類器の競合[3]や回転を前提とした学習データセットが要求される。

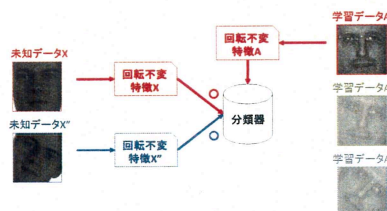


[3] Chang HUANG, Haihou AI, Yuan LI and Shihong LAO, Vector boosting for rotation invariant multi-view face detection. In ICCV, 2005.

5

回転を含む物体の検出・認識アプローチ

- 回転不変特徴の実現は、回転を含む物体の認識、検出アプローチの単純化に寄与する。



6

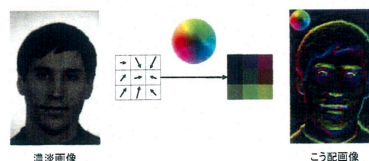
研究の目的

- 画像平面上の回転に対して不変な特徴ベクトルを抽出する手法を提案する。
- 回転を含む対象物体を、高精度に認識、検出するシステムの構築を目指す。



濃度こう配特徴

- 濃淡画像中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[4].
- 認識や検出の高精度化に寄与する特徴量[5].

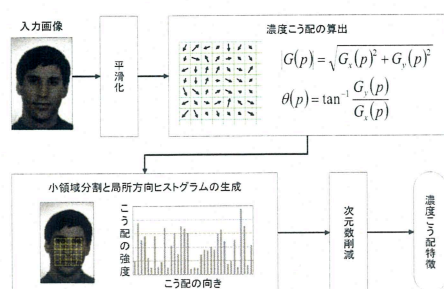


濃淡画像

こう配画像

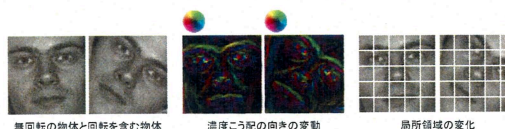
[4] 若林 智史, 鶴岡 慎治, 木村 文雄, 三宅 康二. 特徴量の次元数増加による手書き数字認識の高精度化. 電子情報通信学会論文誌 (D-II), Vol.77, No.10, pp.2046-2053, Oct 1994.
[5] 藤吉 弘直. Gradient ベースの特徴抽出 -SIFT と HOG-. 情報処理学会研究報告, CIVM, Sep 2007.

濃度こう配特徴抽出の流れ



回転に伴う変動

- 物体の回転に伴う濃度こう配の向きの変動.
- 各小領域に含まれる画素の位置の変動に伴う局所領域の変化.

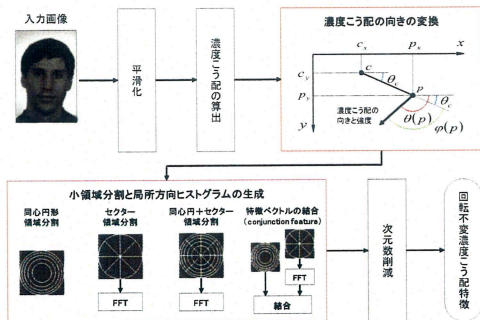


無回転の物体と回転を含む物体

濃度こう配の向きの変動

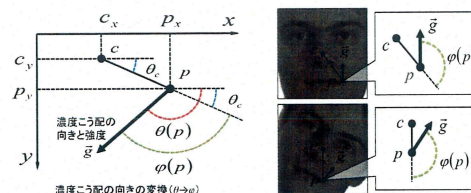
局所領域の変化

回転不変濃度こう配特徴



濃度こう配の向きの変換

- 各濃度こう配の向き $\theta(p)$ を、特徴を抽出する領域の中心点 $c(x, y)$ と各画素とを結ぶ線分を基準にした向き $\varphi(p)$ に変換する。



回転に対して不変な小領域分割

- ### 1. 同心圓形領域分割



- ## 2. セクター領域分割



FFT(一次元)



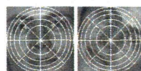
13

回転に対して不変な小領域分割

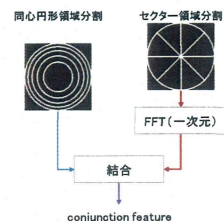
- ### 3. 同心円+セクター領域分割



FFT(一次元)



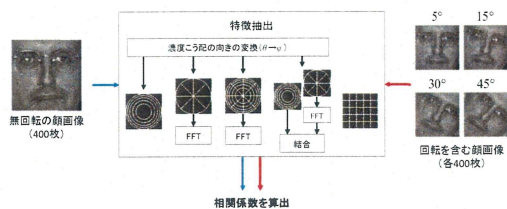
- #### 4. 回転不変特徴の結合 (conjunction feature)



14

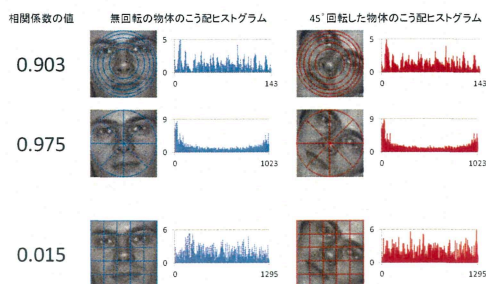
回転に対する不変性の評価

- 無回転の物体と回転を含む物体とのそれぞれから抽出した特徴量の相互相関で評価.



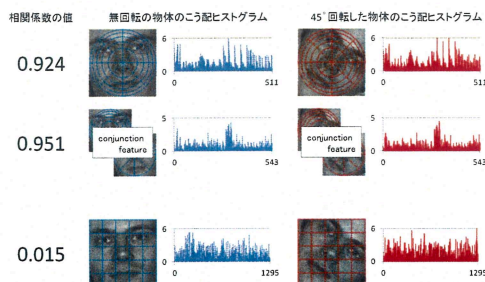
15

回転に対する不変性の評価



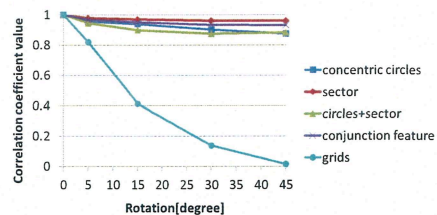
16

回転に対する不変性の評価



17

評価結果



- 提案手法により得られた特徴の相互相関はすべての手法において0.88以上の値を示した.



18

顔検出への応用

- 顔検出とは、入力画像中から顔の位置と大きさを検出することである。



入力画像



顔領域

顔検出実験

- 無回転の顔／非顔のみを学習させたSVMを用いて、回転した顔を含む画像中から顔領域を検出する。
- 評価データは、FERETデータベース[6]の人物画像を用いて作成した。

人物画像
(400枚)

0°



5°



15°



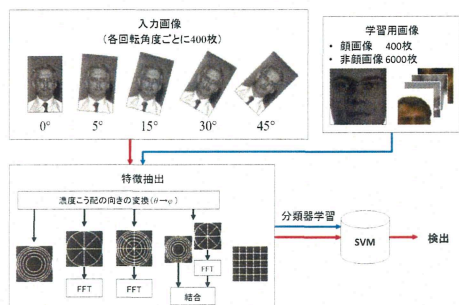
30°



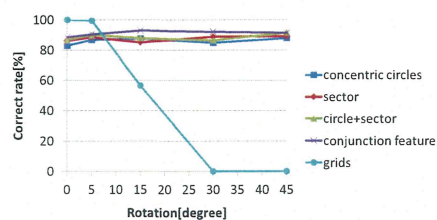
45°

回転した顔領域を含む画像
2000枚(各回転角度ごとに400枚)

実験の流れ



実験結果



- すべての回転角度を対象とした場合の提案手法を用いた顔検出成功率は、86.1%(同心円), 87.6%(セクター), 88.5%(円+セクター), 91.2%(conjunction feature)であった。

考察

- 提案手法を用いた顔検出処理は、画像中の顔の回転に依存しないことが確認された。
- 無回転の顔を検出する性能は、格子形領域分割と比べて、約9%低いことがわかった。
 - 回転不変特徴を抽出する手順の中で、対象を表現する特徴を低減している可能性がある。

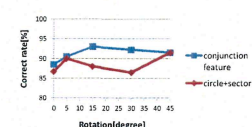
考察

- conjunction featureは、同心円+セクター領域分割と比較して、各小領域に含まれる画素が多いため、より対象を表現する特徴が得られたと考えられる。

同心円+セクター領域分割



回転不変特徴の結合 (conjunction feature)



まとめ

- 回転不変な濃度こう配特徴を抽出する手法を提案した。
 - こう配ヒストグラムの相互相関から、提案手法を用いて、対象の回転に不変な濃度こう配特徴が抽出できることを確認した。
- 提案手法を顔検出処理に応用した。
 - 実験結果から、提案手法を用いた顔検出処理は、画像中の顔の回転に依存しないことがわかった。
 - conjunction featureを用いた顔検出処理において、検出成功率91.2%を得た。



25

今後の課題

- 回転を伴う対象として図面中の文字認識や一般物体検出などの分野へ提案手法を応用する。
- 他の回転不変特徴との比較を行う。



26

御清聴ありがとうございました。

『回転不変濃度こう配特徴とその顔検出への応用』
 三重大学大学院工学研究科情報工学専攻
 ヒューマン・インタフェース研究室
 横田智彦



27