

論文の内容の要旨

論文題目 Local and Global Invariants for Object Recognition 物体認識のための 局所的／大域的不变特徴

氏 名 長 尾 健 司

概要

本論文では画像を用いた物体認識のための不变特徴抽出の問題を扱う。幾何学的、或いは、フォトメトリックな物体認識、さらには、統計的パターン認識において、不变特徴を導入し、それらの適用可能性と有効性について議論を与える。最初に、幾何学的な不变特徴であるアフィン不变特徴を導入し、それを用いた平面物体の認識アルゴリズムについて述べる。続いて、幾何学的特徴の抽出過程において失われた、物体表面の反射特性(カラー)の情報を利用し、幾何学的不变特徴との組み合わせによって、適用可能な物体クラスの拡大を試みる。また、顔画像認識において、照合する画像間の変動を取り除くための新しいパターン認識法を導出する。さらに、扱える画像変動のクラスをシステムティックに拡大していくための新しい枠組みを導く。剛体認識や顔画像認識の実験を通して、提案された不变特徴の有効性を検証する。

1 はじめに

画像を用いた物体認識 [3] の中心的なテーマの一つは、物体を記述するための不变特徴の選択とその適用の方法にある。本論文では、既存の、或いは、ここで新規に提案する幾つかの不变特徴の物体認識への応用と、不变特徴の採用基準となる適用可能性(或いは適用範囲)や有効性に関する議論に焦点をおく。扱われる物体認識の応用問題は、剛体認識と顔認識である。

上述において、不变特徴の適用範囲、即ち、扱える物体認識問題のクラスは、不变特徴が対象とする画像変動のクラス

の範囲によって決定される。一方、不变特徴の有効性とは、そのような選択されたクラスの画像変動下において、認識の観点から不变特徴がいかに有効に機能するかを指す。筆者らは、さらに、この有効性は不变特徴の不变性(恒常性)と選択性(識別性)、或いは、計算論的効率などからなると考える。筆者らの主張は、実用的な物体認識システムの構築には、広い適用範囲と高い有効性の両方を兼ね備えた不变特徴が必要であるという点である。画像変動には、幾何学的変換 [3] のようにある種の明示的な関数のクラスで表現されるものと、アピランスマッチングによる物体認識法 [7] や、顔画像認識のための Eigenface 法 [9] が扱ったような、明示的には表現しがたい画像変動がある。後者の場合は、画像変動は、同一実体を表すデータ集合の統計量(共分散)などによって表現される。変動が数学的な観点において明示的であれ、非明示的であれ、より広いクラスの画像変動に対して、物体認識の観点からより有効に機能するような不变特徴を、筆者らは大域的な不变特徴と呼ぶ。

2 剛体認識のためのアフィン不变特徴

幾何学的不变特徴の一つであるアフィン不变特徴を用いた平面物体の認識について述べる。

幾何学的物体認識 [3] の中心的プロセスは、物体の表現(モデル)である局所幾何学特徴のセットと、入力画像中から抽出された局所特徴セットの間の幾何学的変換パラメタの復元である。アプローチは、基本的に、特徴点对応の仮説と検定による網羅的探索であるため、容易に組み合わせ爆発が生じ

る（最悪時間計算量は特徴点数の6乗のオーダー、特徴点数は一般に数百～数千）。この計算論的な課題を解決するのが本章の目的である。

2.1 画像変動と不変性・選択性

幾何学的物体認識が扱う変動は、カメラと物体との相対運動を伴う。一般に、これにはフォトメトリックな変動を伴うが、これは、初期の局所特徴の抽出過程において取り除かれる。この意味で、幾何学的物体認識は、フォトメトリックな変動にも対応している。しかし、この初期プロセスは、同時に、物体表面の反射特性などの物体固有の貴重な情報を棄ててしまう。即ち、特徴の不変性は確保されるものの、選択性が犠牲となる点に注意を要する。

2.2 アフィン不変特徴と平面物体認識

ここで用いるアフィン変換の性質は、次のものである。

【定理1】 [1] 点集合 $\{X\}$ がアフィン変換 $L(X)$ によって $\{X'\}$ に写されるとする。 X と X' の共分散行列をそれぞれ白色化する（共分散を単位行列に変換する）アフィン変換 A, A' ($\det[A], \det[A'] > 0$) によって、これらの像 Y, Y' は、直交変換の自由度を残し、不変となる。□

これは、平面物体の場合、上記白色化変換によって、それぞれ対応していた点が回転変換によって重なることを意味する。この性質の重要性は、他のアフィン不変特徴の導出可能性に関わる以下の定理によって確認される。

【定理2】 もし利用可能な情報が点集合の2次以下の統計量であるならば、直交変換の自由度を残し不変性を与える唯一のクラスの線形関数が、白色化変換である。□

物体モデルには、輪郭上の高曲率点、変曲点を初期特徴として抽出し、上記白色化変換後に、回転不変のクラスターリング基準を用いて得られた特徴点クラスターの重心のセットを用いる。認識時は、同様に対象画像から抽出された特徴セットに対して、白色化変換、クラスターリングを経て、クラスター重心を得たのち、モデルとの間でアラインメント（変換パラメータの推定と検定）を行う。

図1に、平面物体モデルをテスト画像に重ねた結果を示す。非常に精度よく、パラメータが復元されていることがわかる（cf. [2]）。処理速度は、SparcWS-IPX を用いた場合、数十ミリ秒（従来法では数十分から数時間）と、画期的な改善を確認した。

3 フォトメトリック不変特徴

幾何学的不変特徴にフォトメトリック不変特徴を組み合わせ、選択性の意味で有効性を向上させる方法を述べる。

3.1 フォトメトリックな画像変動と不変特徴

フォトメトリックな画像変動を引き起こす要因には、光源変動、物体の傾きの変化、カメラセンサーの違い、物体表面の形状や反射特性の経時変化などがある。これらの要因が引き起こす画像変動を明示的に表現することは一般には容易ではない。従って、不変特徴を導くために、様々な仮定を導入するが、これは、不変特徴の適用範囲の限定に他ならない。一方、そのような仮定を導入しない場合、不変性は得られても、特徴の選択性が犠牲となることもある。

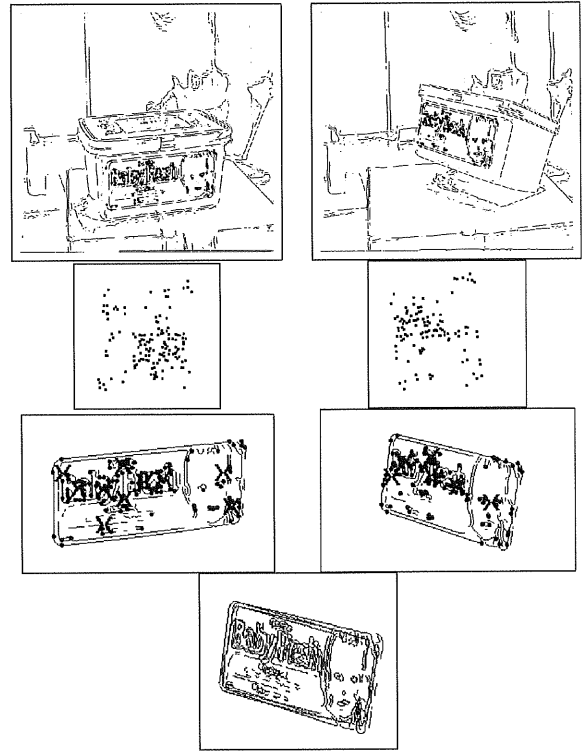


Figure 1: 平面物体の認識結果

最上段の画像は”Baby-Wipe 箱”のエッジマップで、黒丸が特徴点（正面サイドのみ）。第2段は、白色化された特徴点マップで、回転変換（約180度）を除いて、分布が一致している。第3段で、クロス (X) は抽出されたクラスター重心。最下段は、復元されたアフィン変換パラメータを用いて、上段左の画像を右画像に変換後、重ねたもの。

3.2 カラー不変特徴と拡張不変特徴

異なったカラーバンドの出力比をもとに幾つかの仮定の下導出される不変特徴について述べる。画像内の点 x における k 番目のセンサーチャネルのレスポンスを $\rho_k(x)$ とし、以下のような仮定（モデル）を用いる。**AS1**:周波数帯域選択フィルタモデルに従うセンサー、**AS2**:狭帯域センサー、**AE1**:物体表面全体に渡って一貫した入射光、**AE2**:センサーの反応帯域において緩やかな強度変動を持つ光源、**AE3**:物体表面上の局所領域で一貫した入射光。

仮定 **AS1&AE1&AE2**、或いは、仮定 **AS2&AE1** の下、カラーチャンネル i, j のレスポンスの比 $\gamma_{ij}(x) = \rho_i(x)/\rho_j(x)$ は物体表面全体に渡って一貫した、未知の固定スケールファクタ C の自由度を除いて不変特徴である。さらに、仮定 **AS1&AE2&AE3**、或いは、**AS2&AE3** のもと、物体上の隣接点に対応する画像位置 x_l, x_m における $\gamma_{ij}(x)$ の比をとることで、スケールファクタが取り除かれ、 $\phi_{ij}^{lm} = \gamma_{ij}(x_l)/\gamma_{ij}(x_m)$ が完全な不変特徴を与える。ここで、 $\phi_{ij}^{lm} = [\rho_i(x_l)/\rho_i(x_m)][\rho_j(x_l)/\rho_j(x_m)]$ と書き直されることから、点 x_l と x_m の選択には注意を要する。なぜなら、これらの2点が、類似した周波数反射特性を持つ点から選ばれてしまうと、 ϕ_{ij}^{lm} は不変性は有しても選択性に欠けるからである。

上で導入されたフォトメトリック不変特徴を、前章で述べた幾何学的特徴と組み合わせ、クラスター重心を抽出することで、拡張されたマッチング特徴を得る。カラーチャンネル比不変特徴の選択性の効果で非常に安定で効率のよい特徴抽出が

実現できる。

4 パターンとその変動の弱直交化法

この章以降では、画像変動の明示的表現が非常に困難な場合を扱う。ターゲットとする応用問題は顔認識、特に、経時変化を伴うケースに焦点をあてる。

4.1 問題設定

同じ人物の集合を含む、2つの顔画像セット $I^{(1)}, I^{(2)}$ 間での、潜在的対応を推定する問題を考える。図2に扱う顔画像のサンプルを示す。スキャンライン順に5番目まで(順にサブセットA, B, C, D, E)、即ちセット $I^{(1)}$ の画像は、様々な照明条件の下、被験者をビデオ画像で直接撮影したものである。一方、6番目(サブセットP)、即ちセット $I^{(2)}$ は、同一人物の免許証の写真を、再度撮影して獲得したもので、撮影時期は最大3年前に遡る。図に見られるように、同一人物にも拘わらず、画像間には極端な変化が見られる。これほど異なった条件で獲得された顔画像間での認識は従来扱われていない。

4.2 振動特徴空間：画像変動の表現

画像の各画素が直交座標系を構成する高次元画像空間において、異なった人物の顔画像サンプルの集合は、比較的低次元(例えば数百次元)の部分空間(正確には多様体)によって表現されることはよく知られている[9]。この空間は、対象物体(人物顔)の違い、即ち、選択性を強調する空間であるため、しばしば顔(特徴)空間と呼ばれる。

一方、前節で述べたような照合の対象となる画像間に存在する、撮影時期の隔たりや画像生成プロセスの違いに起因する個人の顔画像の変動を表現する空間を、筆者らは、振動特徴空間と呼ぶ。これが、顔画像認識における画像変動の表現にあたる。

4.3 顔パターンとその変動成分の弱直交化法

顔特徴空間から、前節で導入した振動特徴の成分を取り除き不変特徴を抽出するための弱直交化法について述べる。これは、1970年代前半に、Fukunagaら[6]によって開発された直交部分空間法を拡張したものである。

【定義：弱く直交化された部分空間】

2つの確率分布 $\{X_1\}$ and $\{X_2\}$ の共分散行列 Σ_1, Σ_2 それぞれの固有ベクトルが完全に共有され、対応する固有値の順位が互いに逆転しているとき、これらの分布は弱直交化されているという。□

このような弱直交化された分布の導出は、2つの分布の共分散行列に関わる次の数学的性質を利用している。

【性質1】 2つの共分散行列の同時対角化変換によって、変換後の共分散行列は、固有ベクトルが共有される。

【性質2】 変換後の空間に対して、さらに適切なスケール変換を適用することで、2つの共分散行列は、固有ベクトルを共有し、対応する固有値 $\lambda^{(1)'}, \lambda^{(2)'}$ には以下の関係が成立する。

$$\lambda^{(1)'_i} + \lambda^{(2)'_i} = 1 \quad (1)$$

$$\lambda^{(1)'_i} \cdot \lambda^{(2)'_i} = 1 \quad (2)$$

式(1)において、 $n = 1$ は、Fukunagaらの結果[6]に等しい。式(2)は、変換後の共分散行列が、互いに逆行列になることを意味する。

以上のように、同時対角化変換と対角成分の適切な再変換によって、弱直交化された部分空間が得られる。この操作を、顔特徴分布と振動特徴分布に適用し、変換後の顔特徴空間の主軸(大きい固有値に対応)によって張られる空間を抽出することによって、不変性と選択性を併せ持つ不変特徴空間が得られる。

5 混合分布における不変特徴

本章では、さらに広範な画像変動に対応できる不変特徴抽出の新しい枠組みを導入する。

5.1 顔画像空間の混合分布モデル

まず、顔画像空間を幾つかの要素分布よりなる混合分布でモデル化する。簡潔さのために、この論文では、まず、照合する2つの画像セット(前章参照)に対応して、要素分布 $\Omega^{(1)}, \Omega^{(2)}$ を持つものと仮定する。

要素分布内の画像変動の主要ファクタは、光源変動である。このとき、ありうる顔画像サンプルの集合は、高次元画像空間において、照明錐体(illumination cone)を形成し[4]、それが線形分離可能であることも知られている[5]。一方、分布間の変動には、顔の経年変化やセンサーの色域変換、ガンマ補正など、非線形性の強い要素が含まれており、実際、実験で示される通り、Fisherの判別分析[6]では扱うことができない。

5.2 画像変動の表現：分布内/分布に跨る分散

上述のような顔画像空間モデルにおいて、画像の変動を表現するために、伝統的なパターン認識[8, 6]の基本構成要素であるクラス内分散、クラス間分散のモデルを混合分布型の空間にフィットするように拡張する。即ち、分布内クラス内分散 $WCWDS$ (Within-Class-Within-Distribution-Scatter)、超分布クラス内分散 $WCODS$ (Within-Class-Over-Distribution-Scatter)、分布内クラス間分散 $BCWDS$ (Between-Class-Within-Distribution-Scatter)、超分布クラス間分散 $BCODS$ (Between-Class-Over-Distribution-Scatter)の4つの分散を導入する。

5.3 混合分布型分離度基準と不変特徴抽出

混合分布モデルにおけるクラスの間隔度は、同一の分布内、又、異なった分布に跨って異なったクラスを区別する能力と、同一のクラスを同定する能力の2つの要素からなる。分離基準を以下のように定義する：

$$S = \text{tr}[C_3^{-1}C_1] - 2\text{tr}[(C_1 + C_2)^{-1}C_4] \quad (3)$$

ここで、 C_1 から C_4 は、それぞれ C_1 :分布 Ω_1 内の $BCWDS$ 、 C_2 :分布 Ω_2 内の $BCWDS$ 、 C_3 :分布 Ω_1 内の $WCWDS$ 、 C_4 :分布 Ω_1, Ω_2 の $WCODS$ である。基準 S の第1項 $\equiv S_w$ は、分布 Ω_1 内でのフィッシャー基準、第2項は、本論文で新たに導入された分布 Ω_1, Ω_2 に跨るクラス平均の分離度基準である(但し、 $'$ の符号あり)。ここで、 S が、分布 Ω_2 内での



Figure 2: 顔画像のサンプル例

S_w に相当する項を含まないのは、認識プロセスが Ω_1 の各サンプルから Ω_2 のクラス (平均) への対応付けであるため、これによって一般性が失われることはない。

個々の分布内での線形分離可能性を仮定し、特徴抽出には基準 S を最大化する分布毎の線形変換 $A^{(1)}, A^{(2)}$ を用いる: $Y^{(1)} = A^{(1)}X^{(1)}$; $Y^{(2)} = A^{(2)}X^{(2)}$. $X^{(1)}, X^{(2)}$ は、分布 Ω_1, Ω_2 からのサンプル、 $Y^{(1)}, Y^{(2)}$ は抽出された特徴である。 Y について分離度 S を最大化することで、選択性を有する、分布に跨った不変特徴を抽出できる。基準 $S(A^{(1)}, A^{(2)})$ の直接解を求めることは容易ではないため、 S の各項を独立に最大化し、それらの妥協点をとることで、直接計算法を得る。ある条件の下では、この妥協解は、最適解に一致することが示される。

5.4 顔画像認識実験とまとめ

弱直交化法、混合分布法の評価を、顔画像認識を通して、従来法との比較によって行った。比較アルゴリズムは、Eigenface 法 [9]、Fisherface 法 [5] である。

結果を図 3 のグラフに示す。横軸は画像セットの組み合わせ、縦軸は一位認識率 (被験者は 108 人) を示す。全ての場合において、混合分布法 (DSFE)、続いて、弱直交化法 (Worth) の順に優れている。DSFE 法は、サブセット P を実験に含むか否かで、性能の変化がほとんど見られなかった。サブセット P を含まない限り Fisher 法も優れているが、P を含むと 60% 以上の性能低下が見られる。今回の実験では、Eigenface の性能は一貫して低かった。

6 結論

本論文では、画像の不変特徴の物体認識への応用を扱った。実用的な物体認識システムの構築に向けて、不変特徴の適用可能な画像変動の範囲の拡大と、識別能力や計算論的な効率の向上に焦点をあてた議論を行った。

最初に、幾何学的な不変特徴であるアフィン不変特徴を導入し、それを用いた平面物体の認識アルゴリズムについて述べた。続いて、幾何学的特徴の抽出過程において失われた、物体表面の反射特性 (カラー) の情報を利用し、幾何学的不変特徴との組み合わせによって、適用対象とする物体クラスの拡大をはかった。また、顔画像認識において、照合する画像間の変動を取り除くための、新しいパターン認識法を弱直

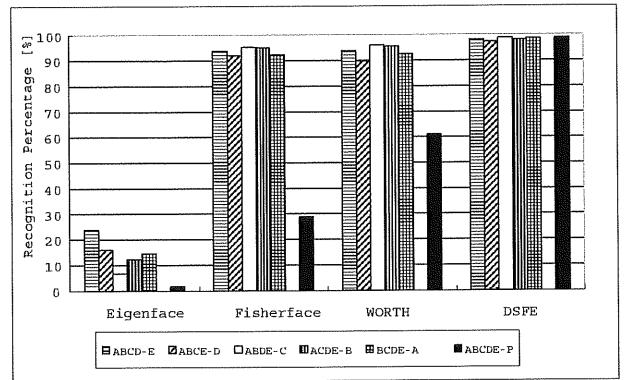


Figure 3: 認識性能比較実験の結果

異なった撮影条件の下獲得された 6 つの画像サブセットを様々に組み合わせ、2 つのセット $I^{(1)}, I^{(2)}$ が作られた。図中 ABCD-E の標記は、サブセット A, B, C, D がセット $I^{(1)}$ に用いられ、一方サブセット E が、セット $I^{(2)}$ に用いられたことを示す。A~E のサブセットでは、それぞれ各個人 3 枚を学習に 2 枚をテストに用いた。学習とテストでサンプルの重複はない。P は各人一枚の画像のみからなるため、学習とテストは同一の画像である。

交化法導出し、さらに、扱える画像変動のクラスをシステムティックに拡大していくための新しい枠組み混合分布型不変特徴法を導いた。剛体認識や顔画像認識の実験を通して、以上の提案された不変特徴の有効性を検証した。

残された課題には、以下のものが含まれる。

(1) 混合分布型の不変特徴抽出法において、2 つの要素分布から一般の混合分布モデルへの拡張が必要である。特に、画像空間の要素分布への分割と、それらの要素分布に跨った不変特徴の抽出法の開発が必要である。

(2) 弱直交化法と混合分布型不変特徴抽出法のフレームワークの幾何学的、あるいは、フォトメトリックな物体認識への応用をはかる。

References

- [1] Kalle Åstroöm, "Affine Invariants of Planar Sets," In *Proc. IAPR SCIA '93*, pp. 769-776, 1993.
- [2] D. P. Huttenlocher and S. Ullman, "Recognizing Solid Objects by Alignment with an Image," *Int. J. Comp. Vision*, 5:2, pp. 195-212, 1990.
- [3] W. E. L. Grimson, *Object Recognition by Computer*, MIT Press, 1991.
- [4] N. Belhumeur, P. Hespanha, J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," In *Proc. ECCV'96*, Vol 1, pp. 45-58, 1996.
- [5] N. Belhumeur, P. Hespanha, J. Kriegman, "What is the sets of all possible images under variable illuminations," In *Proc. CVPR 96*, Vol , pp. 45-58, 1996.
- [6] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press 1972.
- [7] H. Murase, S.K. Nayar, "Visual learning and recognition of 3-D objects from appearance," *International Journal of Computer Vision*, Vol. 14 ,pp. 5-24, 1995.
- [8] E. Oja, "Subspace Methods of Pattern Recognition," Research Studies Press LTD.
- [9] M. Turk, A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991.