

氏名（本籍）	関 真規人（香川県）
学位の種類	博士（工学）
学位授与番号	甲第41号
学位授与日付	平成23年3月25日
専攻	システム工学専攻
学位論文題目	部分情報の統合に基づいた画像解析方法に関する研究
学位論文審査委員	(主査) 教授 和田 俊和 (副査) 教授 河原 英紀 教授 呉 海元

論文内容の要旨

デジタルカメラの発展は、誰もが簡単に撮影できるという手軽さと、写真をデジタル画像データとしてそのまま保存できるという便益をもたらした。これにより、我々の身の回りにはたくさんの画像データが存在するようになってきている。その多くは日常的なシーンを対象に撮影されたものであり、従来の画像認識の研究で扱われてきたような特定条件下で撮影された画像とは異なる。

こうした一般的な画像を計算機が自動的に認識するためには、前景物体と背景物体とを区別したり、画像に映る空間構造を認識したりして、画像が撮られたシーンを理解することが重要である。その基本となる画像解析技術の研究は古くから行われており、大きくわけて2種類のアプローチがある。

ひとつはアピアランスベースアプローチである。この方法では、認識対象の見掛けのアピアランス情報をそのまま用いて対象を認識する。例えば、Muraseらは、認識対象の姿勢変化を考慮したアピアランス集合を固有空間中の多様体として学習し、入力画像から対象を認識している。このような方法は、認識対象が映った学習画像さえ用意できれば適用可能であり、実際の3次元形状情報などを用意することなどに比べれば現実的なアプローチといえる。しかし、認識対象あるいは画像全体を特徴として用いているため、一部が隠蔽されるような場合には対応できないという問題がある。また、人のように変形する物体や木のように形状が定まらない物体を対象とする場合には、アピアランスとしてのバリエーションが増え、記述そのものが困難となる。

これに対し、フィーチャーベース、あるいはパートベースと呼ばれるアプローチがある。この方法では、画像中の特徴点などから認識対象についての部分的な情報を抽出し、それらを統合することによって対象を認識する。例えば、Schmidらは、特徴点の輝度値やその微分値を抽出し、その集合によって画像の認識を行っている。また、Loweも、対象のスケールや回転に不変なSIFT (Scale Invariant Feature Transform) と呼ばれる特徴点記述法を考案し、対象認識を実現している。このような部分情報を扱うアプローチは古くからあったが、新たにSIFTのような高い記述力をもった特徴量が出現したことによって、さまざまな用途で高い認識性能が得られるようになってきている。また、もとより、本アプローチは、部分情報を統合するため、対象物体の変形や隠蔽に対応できる柔軟性も持ち合わせている。

ここで部分情報の統合方法について考えてみる。部分情報の統合方法には、部分情報が得られた特徴点間の空間構造（隣接関係など）を一切使わないものと、逆に積極的に活用するタイプのものがある。前者はハフ変換やBags of Keypointsと呼ばれる方法などが該当する。後者はMarkov Random Fields (MRF) や constellation model という空間構造モデルを使う方法などが該当する。

ハフ変換は直線などの図形を検出するための方法であり、画像中の特徴点を与えられたとき、その特徴点を含む図形候補に対して投票を行い、多くの特徴点から投票された候補を正しいものと考えて図形の検出を行うものである。そのため、ハフ変換では、各候補に対しての投票度数を蓄積したヒストグラムを生成する。また、Bags of Keypointsでは、SIFT等によって記述された多数の特徴量を量子化し、量子化特徴量の出現頻度をヒストグラム化する。そして学習モデルにおけるヒストグラムとの類似度から対象を認識する。これらから分かるように、特徴点間の空間構造を利用しないタイプのものは、部分情報の集合、すなわち、ヒストグラムを生成するものが多いと言える。しかしながら、ヒストグラムの生成過程やその使い方において少なくとも3つの課題が存在する。

- ヒストグラムにおけるセルとして、図形候補や特徴量の種類を表すパラメータの最適な量子化が必要である。例えば、量子化が細かすぎると1つのセルに入るデータの個数は高々1つとなってしまい、個々のデータを扱うと変わらなくなってしまう。逆に、量子化が粗い場合は、情報のロスが大きくなる。そのほか、量子化における区切りの位置が適切でない場合は、本来1つにまとめなければならない情報が分散してしまうという問題が発生する。

- ヒストグラムの次元が高い場合、効率的で安定な最適解探索方法が必要である。例えば、高次のヒストグラム空間を全探索するとすれば計算時間が問題となる。また、山登り法によって最適解を探索する場合には、初期値に依存して誤った局所解に陥ってしまうという可能性がある。
- ハードウェアリソースの制限などにより、特徴量を表すパラメータの次元に対してヒストグラムの次元を高く設定できない場合には、情報ロスが大きくなり、ヒストグラムを用いた対象の検出感度は低下することになる。そのため、何らかの方法によってヒストグラムの先鋭化が必要となる。

一方、MRF は多数の変数とその変数間の相互作用からなるグラフィカルモデルであり、グラフに沿って情報を相互に連鎖させることができる。例えば、Zhang らは、画像上における個々の画素値が独立に生成されたものではなく、互いに関連しあって意味を成したデータとなっていると考え、画素の空間的な構造に MRF を当てはめ、その MRF のもとで情報を統合し領域分割を行っている。また、constellation model も特徴点の位置関係が扱えるモデルであり、特徴点周りの特徴量、それらの相対位置、スケールの同時確率によって対象を認識することができる。しかし、各特徴点に割り当てべきラベルとしてエネルギー関数を最小/最大にする組み合わせを求めたり、画像から得られた特徴点群とモデルがもつノード群とのさまざまな組み合わせを考えたりする必要があり、

- 部分情報を統合していくための組み合わせ最適化計算に時間がかかる

という問題がある。

そこで、本研究では、一般的な画像における対象認識を目指し、部分情報の統合方法についての上記 4 つの課題をとりあげ、解決策を提案する。

最初に、直線検出用 $\gamma-\omega$ ハフ変換を題材として、パラメータ空間の量子化問題を取り上げる。ベースとなる $\gamma-\omega$ ハフ変換では、パラメータ空間を均一なセルで量子化しているが、そのために 1本のデジタル直線に含まれる画素集合からの投票がパラメータ空間中の複数のセルに分散し、デジタル直線を構成する画素の数が投票度数として正しく捉えられないという問題が発生している。これは、 $\gamma-\omega$ ハフ変換が、画像空間中に存在する全てのデジタル直線を検出対象として記述できていないことに起因する。ここでは、 $\gamma-\omega$ パラメータ空間の最適な量子化方法として、画像空間中に存在する全てのデジタル直線と 1対1に対応したセル配置を求める。また、そのセル配置に対する妥当な投票方法を明らかにする。

次に、時系列画像からの消失点検出を題材として、局所解に陥らずに効率的に最適解を探索できる方法を提案する。一般的な従来手法では、直線検出ハフ変換のパラメータ空間において、消失点を構成するはずの直線候補群に相当する複数のセルの投票度数の総和を統合する手法が採られている。この方法は、全ての消失点候補を評価することになるため、画像中に消失点を構成しない直線が存在するような場合には誤った結果に至る可能性がある。ここでは、パーティクルフィルタを利用して、予測される消失点の確率密度に従って複数の仮説を生成し、各仮説に対して求められる尤度の期待値から消失点を検出する方法を示す。

そして、3つ目の課題については、背景差分法を題材として、ヒストグラムの先鋭化について述べる。従来の背景差分法では、背景画像の学習サンプルから画素や部分画像ごとに画素値のヒストグラムを生成して、その度数に応じてダイナミックレンジの調整を行っている。このため、ダイナミックレンジの大きい部分では前景物体の検出感度が落ちてしまうという問題があった。ここでは、局所領域間にある依存関係、すなわち画像変化の共起性を利用して周辺から着目画素の状態を推定することでヒストグラムを先鋭化し、ダイナミックレンジの幅を動的に狭める新しい背景差分法を提案する。

4つ目の課題については、MRF を用いた領域分割法を題材として、計算時間の削減方法を示す。従来の領域分割法では、MRF のもとで画像空間での連結性評価と特徴空間でのクラスタリングを反復的に行い、各画素に最適なクラスタ中心値を割り当てるといった方法が用いられている。ここでは、連結性評価とクラスタリングを階層的に行うことで高速化を図る。具体的には、特徴空間にクラス中心点を与え、ポロノイ分割することで画素値分布のクラスタリングを行う。このとき、階層的にクラス中心点を増やすことで、ポロノイ分割を更新する。また、クラス中心点を適宜シフトすることでも最適化を図る。また、これに合わせて、画像空間の連結性評価も階層的に行う。

最後に、領域分割をベースとして、各領域から得られる部分情報の統合に基づいた対象認識手法を提案する。

論文審査の結果の要旨

学位論文、申請書類、公聴会に基づいて審査を行った。関君の当該研究内容は、視覚情報処理の分野において古くから研究されてきた「部分情報の統合問題」に対するオリジナルな研究をまとめあげた優れた論文である。この論文では様々な技法を提案しているが、個々の手法は当該分野のトップレベルの国際会議、ジャーナルペーパー等で発表されており、学術的新規性や実用技術としての有効性も兼ね備えている。具体的にはHough変換、Particle Filter、背景変動の共起性を利用した背景差分、Markov Random Fieldに基づく領域分割などの技法を提案しており、論文の最終章では、これら技法を総合して、領域分割に基づくシーン認識、物体認識法の提案と実験結果が述べられている。研究業績は、本論文に関係のある掲載済み論文が5件、国際会議が2件、国内会議での発表が10件であり、必要な要件を満たしている。これらのことより、上記の論文内容は博士学位の授与に値するものであると認める。

最終試験の結果の要旨

最終試験として、公聴会において諮問を行い、さらに公聴会終了後に学力諮問を行った。本論文に関連する画像解析、仮説生成・検証型アルゴリズムに関する知識と学力は十分備わっているものと判断した。Markov Random Fieldに関しては、さらに学習を進めるべき点もあったが、総合的には博士学位の授与に値する十分な学力があると認める。