

## 自然シーンにおける対象物の認識に関する研究

## 学位論文内容の要旨

近年、特に1990年代に入ってから計算機の進歩はめざましいものがあり、高速な中央演算処理装置、十分な容量の記憶装置を持つ計算機システムを、かつては考えられないほど容易に手にいれることができるようになってきている。このような計算機の性能向上に伴い、我々が処理を要求するデータの規模も次第に巨大なものになってきている。特に音声情報、静止画像情報、動画像情報といった、いわゆるマルチメディアデータを対象とした処理の需要が増えてきている。マルチメディアデータは一般的に人間が使用することを目的としたデータであるため、これらのデータを解析する際には、処理対象の名称や人間の感覚に基づいた修飾語句などの、数値データへの変換が難しい入力に対する処理を実現する必要がある。ここに、数値や文字列を扱う一般的な大規模データベースとは違った実現の難しさがある。マルチメディアデータの中でも、特に画像データは容量が大きいので、大量の画像データや動画像データを解析する場合には、解析手法が十分に高速であることが重要となる。中でも、画像検索のように、すべての画像データについて情報を抽出する処理には、実用に見合う高速な手法が要求される。

静止画像データベースや一連の動画像データの中から、ある特定の条件を満たす画像をユーザの指示で検索する処理を考えた場合、建物や車などといった物理的な対象が検索キーになることがある。これらの対象の切出しには、画像中の物理的な対象を表わすであろう領域の境界を定める、領域分割という手段を用いる。しかし、異なる物理的な対象に対応しているが隣接する二つの領域が類似した色である場合や、画素値勾配が緩やかな場合に正しい分割ができないという問題がある。動画像の場合は、切出す対象の動きを解析し、オプティカルフローと呼ぶ、動きベクトルをクラスタリングによって分類することなどで、一つの対象を他の対象から分離することができる。しかし、切出した領域から得られる特徴のみでは十分な情報を得られず、これらの領域をパターン認識手法で正しく認識するのは難しい。結果として、パターン認識の結果を、元の画像における対象に関する知識に照らし合わせて、整合性を判断することが有効な方法となる。この整合性を高めるように、対象に関する知識とシミュレーテッドアニーリングによってラベルを修正することで、認識率を改善する方法がこれまで提案されている。しかし、この方法ではラベルの修正に時間がかかるという問題があり、大量の画像を処理する目的には向かない。一方、動画像から動きを抽出する手法の大部分は、物体の輪郭部分において精度が高くない、あるいは、物体の細かな動きに追従しにくいといった問題がある。本研究では、まず、屋外シーンを対象とした静止画像の認識手法について、大量の画像が対象となる場合に、処理時間を短縮することと認識精度の向上のために、知識を有効に利用する手法について検討を行った。また、動画像中の物体間の分離精度向上が期待できる、物体の輪郭部で精度の高いオプティカルフロー抽出手法を提案する。

本論文は対象を静止画像と動画像に分け、それぞれの場合に、対象物の抽出、対象物の動き抽出に関して、新しい手法を提案するとともに、検討を行った。内容は二部構成をとり、その概要は以下の通りである。

第一部では、領域分割によって得られた領域に対象を表すラベルを適切に割当て、屋外シーンを認識する手法について述べる。

第1-2章ではラベリングによる屋外シーン認識について、その背景と従来法の概要、および問題点について述べる。多くの従来手法は、最適なラベルを得るためにシミュレーテッドアニーリングを用いたラベル改善を行うので、処理に時間がかかり、大量の画像を処理する際の障害となっていた。また、領域の特徴量として考慮する情報が色に偏りがちであり、見かけの色といった、物体の外見の変化に対応できないという問題を持つ。

第3-4章では提案手法の構成について述べ、実験を行い結果を考察する。提案手法では、シミュレーテッドアニーリングに要する処理時間を短縮するために、ローカルヒルクライミングによってラベルの改善を行う。また、対象物体に関する知識を積極的に利用することによって、ローカルヒルクライミングによる局所解への落ち込みを回避する。さらに、色以外に領域形状やテクスチャの特徴量の導入によって特徴量の色への偏重を補い、認識結果の高精度化を計る。実験では、提案手法と従来手法を実際の屋外シーンに適用して認識結果と処理時間を比較する。この実験により、提案手法が従来手法に較べてはるかに高速にラベルを割当てられることを示す。

第5章では第一部の結論を述べる。

第二部では、動画像からの物体の動きを、オプティカルフローとして抽出する手法について述べる。

第6-7章では、動き抽出の背景について述べ、微分方程式とブロックマッチという抽出に関する異なる二つの基本原理について示し、これらの基本原理に基づく従来手法の問題点を明らかにする。従来手法は、原理的に物体の輪郭部においてオプティカルフローが抽出できない、もしくはオプティカルフローの抽出単位が大きく、細かなフローの抽出に向かないという問題がある。

第8-9章では、提案手法の構成について述べ、実験を行い結果を考察する。提案手法では、二つの基本原理に基づく従来手法の問題点に対し、微分方程式に基づく手法の輪郭部での抽出不能問題をブロックマッチに基づく手法を用いることで回避する。また、ブロックマッチに基づく手法の問題点に対しては、画像の各部分について動き抽出に用いる適切なブロックのサイズを求め、各ブロック単位で抽出したオプティカルフローを補間することで、細部のフローを求める。実験ではブロックマッチに基づいた従来手法との比較により、提案手法の有効性を示すとともに問題点を挙げる。

第10章では第二部の結論を述べる。

最後に、二部のまとめを述べ、将来解決すべき問題を挙げる。

# 学位論文審査の要旨

主査	教授	新保	勝
副査	教授	伊達	惇
副査	教授	宮腰	政明
副査	教授	北島	秀夫
副査	助教授	工藤	峰一

## 学位論文題名

### 自然シーンにおける対象物の認識に関する研究

近年、計算機の性能向上に伴い、音声情報や静止画像情報、動画像情報といった、いわゆるマルチメディアデータを対象とした処理の需要が増えてきている。中でも画像データは容量が大きいため、大量の画像データや動画像データを解析する場合には、解析手法が十分に高速であることが望まれる。特に画像検索のように、すべての画像データについて様々な情報を抽出する場合には、高い精度と共に実用に見合う高速性が要求される。

本論文では、このような背景から、屋外自然シーン画像の異なる対象に領域分割と対象に関する知識を用いてラベルを割り当てる手法の高速化と、動画像中の移動物体の動きをオプティカルフローで抽出する手法の高精度化を試み、その有効性を検証したものであり、成果は次の点に要約される。

(1) 屋外自然シーンのラベリング処理において、ラベリング結果の評価関数に従来よりも自然で強い仮定を導入し、評価関数の極小解を減じた。これにより従来用いていたシミュレーテッドアニーリングの代りに高速なローカルヒルクライミングを使うことが可能となった結果、精度を維持しつつ高速化が計れることを示した。

(2) 屋外で自動車が動いているような動画像において、対象の動きを表すオプティカルフローの抽出問題を扱い、階層的な手続を用いることにより、高速で解析精度を落さずにオプティカルフローを抽出できることを示した。

(3) 同様の問題において、幾何学的な変換を考慮したブロックマッチングを導入することにより、オプティカルフローを精度よく抽出できることを示した。

これを要するに、著者は、自然シーン画像中の対象物の認識に関して、高速なラベリング手法と高精度な動き抽出手法を提案し、その手法の有効性を検証したものであり、数理工学ならびに画像処理工学の分野に寄与するところ大なるものがある。よって著者は、北海道大学博士(工学)の学位を授与される資格あるものと認める。