

審査の結果の要旨

氏名 柳井 啓司

近年の画像入力機器の普及や記憶装置の大容量化により、計算機への実世界画像の蓄積が容易になってきており、画像の意味内容に応じた画像検索や自動分類などの実現のために、計算機による画像認識技術への要求が高まっている。従来の画像認識の研究では、顔画像など特定の対象の認識を目的としていたり、工業部品の認識の様に認識対象が完全に既知であることを前提としていたりしており、対象を限定しない実世界の画像に対しては適用が難しいという問題点があった。これに対して本研究は、実世界画像の持つ2つの困難な問題である(1)多数の種類が存在して物体の種類によって適する認識方法やモデル表現が異なる場合、(2)単一種類の物体でも様々な個体が存在して、画像中に現れる見え方が多様である場合、に対してそれぞれシステム構成法と認識方法の観点から追究したものである。(1)に対しては、従来の画像認識の研究で多数提案されている特定種類の物体に対する認識手法と知識表現を統合して対処することとし、多数の認識手法と知識表現を統合するための認識システム構築法を提案している。(2)については、単一種類の物体の画像中での様々な現れ方に対応するために、多数の学習画像を WWW (World-Wide Web) から自動収集して自動的に画像認識のための知識ベースを構築する方法を提案している。

本論文は「実世界画像に対する画像認識の研究」と題し、10章からなる。

第1章「序論」では、研究の背景、目的、位置付けについて述べている。

第2章「実世界画像の物体認識」では、実世界画像に対する従来の研究についてまとめている。従来の研究の流れは、システム構築に関する研究と認識手法に関する研究の2つがあることを述べている。

第3章から第6章では、実世界画像に対応したシステムを構築するために、異なる多数の認識手法や知識表現の統合をマルチエージェントの考え方に基づいて実現する方法を提案している。

第3章「認識システムの構成法」では、従来の実世界画像に対する認識システムの構成法についてまとめており、従来の認識システムでは、対象を予め想定してシステム構築が行われてきたので、システムの構成が認識対象の画像の種類に依存したものになっていたり、各対象物の認識のための知識が相互に密接な関係を持っていたりするために、異なる種類の画像に対する知識を混在させることが困難であり、様々な種類の画像が存在する実世界画像の認識には適用が難しいという問題点があることを指摘している。

第4章「マルチエージェント型画像理解システムの提案」では、前章にあげた問題点を解決するためのマルチエージェントによる画像認識システムの構築法を提案している。各エージェントが1種類の物体のみを認識する独立した認識システムとし、物体毎にそれぞれに異なる知識の表現および認識手法を用いることができ、システム構築の自由度が増すことを主張している。この有効性は、実験によって室内画像と屋外画像の両方に対応でき

る認識システムの構築によって確認している。

第5章「物体間の位置関係に関する空間推論の導入の提案」では、マルチエージェントによる画像認識システムに物体間の定性的な位置関係に関する推論機構を導入することを提案し、より複雑な画像の認識を可能とする方法を示している。実世界画像においては、物体が物体の上に乗ったり、手前に位置したりして、物体が物体を隠すオクルージョンが発生する。オクルージョンのために一部分しか画像中に現れていない物体を認識可能とするためには、物体間の位置関係を利用することが不可欠であるが、従来は主に画像上での物体領域同士の2次元的な位置関係しか利用されていなかった。ここでは、物体の定性的な3次元情報を利用して定性的な3次元位置関係の推論を行うことにより、実世界画像で問題となるオクルージョンに対処する方法を提案し、室内画像に対するシステムを実現してその有効性を実証している。

第6章「多重解像度解析の導入による高解像度画像の利用の提案」では、画像中に小さくしか現れていない対象を高解像度画像によって認識可能とする方法を提案している。単純に高解像度画像を用いることは、認識時間の著しい増大を招くが、ここでは、多重解像度解析を導入することによって、効率のよい認識を実現する手法を提案し、より複雑な実世界画像が認識可能となることを実験によって示している。

第7章から第9章では、画像内容を表すテキスト情報を伴った多種多様な画像をWWWから自動収集することによって画像認識のための知識ベースを自動構築し、同一種類でも多様な個体が含まれる実世界画像を認識可能なシステムを実現する方法を提案している。

第7章「多数の学習画像を用いた画像認識」では、多数の学習画像を用いた実世界画像の認識について従来の研究をまとめており、従来の研究では学習画像を収集することが困難であったために、顔画像や自動車の画像などの限定された対象にしか実験が行われていなかったという問題点を指摘している。

第8章「WWWからの画像収集方法の提案」では、実世界画像を大量にしかも手軽に収集する方法として、WWWから自動的に大量の実世界画像を収集する方法を提案している。WWW空間中に存在する画像はその多くが画像内容を表すテキスト情報を伴っているので、テキスト情報を解析することによってユーザの望む画像をWWWから収集することが可能であるということに着目したものである。

第9章「WWWからの収集画像を用いた画像分類の提案」では、提案した画像収集法を用いて様々な実世界画像を自動収集し、それらを学習画像として、実世界画像に対する認識を行う手法を提案している。最初の学習段階ではWWWから認識したい対象画像を各種類(クラス)毎に数百枚から数千枚程度収集し、それらから色情報、テクスチャ情報などを画像特徴として抽出して各クラス毎の画像特徴に関する知識ベースを構築する。次いで、認識段階では認識対象の画像から同様に画像特徴を抽出し、知識ベースと照合を行って最も可能性の高いクラスに分類して認識するものである。実験では、この提案手法により画像に関する知識をまったく与えることなく、単語入力のみで画像分類が可能になることを示

している。

第10章「終章」では、本論文の内容をまとめ、今後の実画像認識の研究についての課題と展望を述べている。

以上、これを要するに、本研究は実世界画像に対する画像認識手法に関する提案と、その手法を実証するシステム構築と実験による評価を与えたものであり、情報工学の研究に貢献するところ大である。よって本論文は博士(工学)の論文として合格と認められる。