

論文の内容の要旨

論文題目	Efficient Object Discovery Based on Locality Sensitive Hashing (Locality Sensitive Hashingに基づく効率的なオブジェクト発見)
学位申請者	Gibran Fuentes Pineda

画像内のオブジェクトを認識するには、コンピュータに取り扱えるようオブジェクトをモデル化する必要がある。現在のオブジェクト認識手法の多くは、オブジェクトモデルを生成する際に、オブジェクトの切り出し、アノテーション、含まれるオブジェクトの種類数などを人が教示する必要がある。しかし、人手による教示は画像枚数が多くなるとコストが非常に大きい。そして、画像枚数が多いほど、また画像表現の次元数が高いほど、認識性能が劣化する。このため、現在のオブジェクト認識手法では近年出現するようになった大量の画像情報を取り扱うことができない。

そこで、本論文は画像から教示なしに効率的にオブジェクトを発見する手法について論じる。本研究の目的は、画像集合を与えられた時に、オブジェクトを含む画像の検索や画像内オブジェクトの位置同定に使えるオブジェクトモデルを自動獲得することである。とくに本論文では繰り返しパターンを探すことにより自動的にオブジェクト発見を行う新たなオブジェクト発見手法を提案する。提案手法は、高速なLocality Sensitive Hashing (LSH) を活用して効率的に繰り返しパターンを探索することで、大規模な画像集合を取り扱えるようにする。提案手法は、画像全体の情報ではなく局所的に存在するオブジェクトに注目して画像間の関連付けを行うものである。本論文ではこのアイデアに基づいた手法を2つ提案している。

1つ目の手法は、まず低レベルの領域分割により画像から領域を抽出し、そして、近接領域の繰り返しパターンを抽出することにより、オブジェクトを自動発見する。本手法は以下の4ステップから構成される。(1)同色の近接ピクセルをまとめ、オブジェクト部品(部品)として取り出し、(2)部品を色と大きさでクラスタリングすることにより、部品へのラベル付与、(3)空間的に近接した部品をまとめて、オブジェクト候補の生成、(4)同じオブジェクト候補が複数存在する場合に意味のあるオブジェクトとして返す。各ステップは、ユークリッド空間上

のLSHによって高速に実行される。本手法は、回転、平行移動、さらにある程度のクラス内変化に対してロバストにオブジェクトを発見できる。

2つ目の手法では、まず各画像をベクトル量子化されたアフィン共変特徴ベクトルの集合として表す。そして、何度も同じ画像に出現する特徴点群（共起特徴点群）は同じオブジェクトを構成する可能性が高いという仮定の下、共起特徴点群をクラスタリングにより統合してオブジェクトとして抽出する。本手法では共起特徴点群はMin-Hashingによって効率的に発見される。また、本手法が生成するオブジェクトモデルはオブジェクト識別能力にすぐれ、シーンの乱雑さ、オクルージョン、スケール・照明・視点の変化に対してロバストである。実験により、提案手法が人間によるアノテーションと合致するオブジェクトを発見できることを確認した。さらに、処理速度については5062枚の画像を7.5分で高速に処理でき、認識性能についてはベンチマークを用いた定量評価において提案手法の性能が他の最新手法の性能を上回った。

論文審査の結果の要旨

学位申請者氏名 GIBRAN FUENTES PINEDA

審査委員主査 渡辺 俊典

委員 古賀 久志

委員 多田 好克

委員 末廣 尚士

委員 森田 啓義

本論文では、与えられた画像集合から、それらに含まれるオブジェクトのモデルを、非教師モードで自動獲得する問題を取り上げ、その解決のための二つの方式を提案している。提案手法で獲得されるモデルは、内容にもとづく画像検索などに有用である。論文は以下の5章からなる。

第1章では、研究の動機、目標、および主たる貢献点を述べている。

近年のデジタルカメラの普及やインターネット上の画像情報の増加によって、それらの自動解析へのニーズが高まっているが、現状の画像認識技術の多くは、人手介入や高い計算コストを要することから、有効な解決策となり難いことを指摘した後、提案手法のベースとなる考えを示している。それは、画像内に繰り返し出現する情報は、たまたまの事実というよりも、意味のあるオブジェクトを表現しているとみなせるといふものである。

第2章では、伝統的画像認識技術のレビューを行ない、それらの限界を整理している。また、提案手法で用いる、事物間の類似性を計算する効率的ハッシング手法であるLocality Sensitive Hashing (LSH)法を説明している。

第3章では、最初の提案方式を述べている。

これは領域分割済みの画像を入力とし、発見したオブジェクトを出力する。アルゴリズムは以下の4ステップで構成される：

- ・色と場所がともに似ている画素を集め、その領域をオブジェクト部品とする。
- ・部品の色とサイズに着目し、部品にラベル（名辞）を付与する。
- ・異なる部品で近接するものを集めてオブジェクト候補とする。
- ・オブジェクト候補の似たものを集めてクラスタ化し、オブジェクトモデルを構築する。

それぞれのステップはハッシング手法をベースとした簡便な実現となっている。特に、最初の3ステップで距離の大小による事物の類似度判断のために、距離感受性ハッシング手法（ユークリッドLSH）を共通して利用し、簡便かつ高速な方式を実現している点が特徴的である。

本手法が、画像回転、平行移動、クラス内でのあまり大きくないオブジェクト

の変動、などに対応できることが実験で示されている。また、オブジェクト発見のための上記ステップは、オブジェクト認識にもそのまま使えるようになっている。この点も興味深いものである。

第4章では、特徴点集合Bof (Bag-of-Features) を基礎としたオブジェクト発見手法を提案している。それは以下のステップで構成される：

- ・与えられた画像のそれぞれからビジュアルワード (VW:ベクトル量子化アフィン共変量) の集合を抽出する。それらを逆引き辞書ファイルとして記憶する。数学的には、VWから出現画像リストへの写像の定義となる。
- ・これを用いて同一画像内に共起するVWを抽出し仮オブジェクトとする。
- ・最後に共通要素の多い仮オブジェクト候補をまとめて最終オブジェクトモデルとする。

Min-Hashing手法によって、これらの処理の基礎である共起VW集合の抽出とクラスタリングの両者を実現している。オブジェクトモデルも画像もVW集合として表現しているため、VW集合を与えてそれを含む画像を検索すること、すなわち画像認識も同じ枠組みで実現される。これらは、本研究の新たな特徴である。

実験の結果、画像から人が目視で抽出したオブジェクトを同じように抽出できること、しかも撮影方向、照明、隠ぺいなどの条件変化に対しロバスト (性能の著しい劣化が無いこと) であることなどが実証されている。具体的には、CPUメモリが806 MBの通常の単体PC計算機上で101,922枚の画像からのオブジェクトモデル抽出を行わせた時の所要時間は26.73分であり、大型並列計算機を用いた先端類似研究の報告が、数時間~数十日を要していることと比較して飛躍的な高速化を実現している。得られるオブジェクトモデルの妥当性検証は、Oxford建造物画像アーカイブ (人手で建物の名前がラベル付けされている5062枚からなる) を用いて実施され、自動抽出モデルによる画像分類結果が人手によるものと高い整合性を示すこと、抽出モデルによる類似画像検索の平均精度 (AP値) がすべてのクラスの画像において現状の先端研究を20%程度凌ぐことなどが示されている。また、その性能は外部指定パラメタの広い範囲で安定していることも示されている。

第5章では、結論と今後の課題が述べられている。

以上を踏まえ、本論文は以下のような意義を持つと考える。

本論文は自動画像解析への重要なステップの報告である。提案された高性能オブジェクト発見手法は、巨大画像アーカイブを、その保有内容によって組織的に記憶したり検索したりするというチャレンジングな課題への新たな解を与えている。数学的には、距離位相を持つ、あるいは持たない一群の集合の観測結果から、その共起部分集合族を発見するアルゴリズムをHashing手法の統一的利用によって実現したものである。文書コーパス集合からのトピック抽出などの他のデータマイニング問題にも広く適用できる可能性を備えている。提案手法は、与えられたデータセットからのオブジェクトモデルの学習と、得られたモデルを用いた未知データセットの認識という二つの問題の両者を同時に解決している。これは、機械学習分野において新たな興味深いスキームの提案ともなっている。

これらにより、本論文は、博士 (工学) の価値を十分に備えていると判定する。