

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 4 日現在

機関番号：12601

研究種目：研究活動スタート支援

研究期間：2014～2015

課題番号：26880004

研究課題名(和文) 実世界知識データベース構築のための自律撮影ロボット行動計画に関する研究

研究課題名(英文) Action planning of autonomous robots that take pictures for construction of real-world knowledge database

研究代表者

金崎 朝子 (Kanezaki, Asako)

東京大学・情報理工学(系)研究科・助教

研究者番号：00738073

交付決定額(研究期間全体)：(直接経費) 1,900,000円

研究成果の概要(和文)：実世界における有用な画像データベースを自動構築する自律移動ロボットを提案し、物体検出の基盤技術の研究開発を行った。第一に、物体検出をグラフマッチング問題として定式化した場合のパラメータを自動学習する手法を提案した。グラフマッチングパラメータを最適化する手法はいくつか知られているが、本手法は物体検出の精度を上げる目的関数を明示的に最適化する点が新しい。本研究は3DV 2014で発表した他、日本ロボット学会研究奨励賞を受賞した。第二に、未知の物体を検出する手法として、三次元空間から物体候補領域を事前知識なしに抽出する手法を提案した。本研究はIROS 2015で発表し、ソースコードを公開した。

研究成果の概要(英文)：First, we proposed an optimization method for estimating the parameters that typically appear in graph theoretical formulations of the matching problem for object detection. Although several methods have been proposed to optimize parameters for graph matching in a way to promote correct correspondences and to restrict wrong ones, our approach is novel in the sense that it aims at improving performance in the more general task of object detection. We presented this work at 3DV 2014 and also achieved 30th (2015) RSJ Young Investigation Excellence Award. Second, to detect unknown objects in the real world, we proposed a new method for obtaining object candidates in 3D space. Our method requires no learning, has no limitation of object properties such as compactness or symmetry, and therefore produces object candidates using a completely general approach. We presented this work at IROS 2015 and also published open source code.

研究分野：コンピュータビジョン, 機械学習, 物体認識

キーワード：ロボットビジョン 三次元物体認識 物体検出

1. 研究開始当初の背景

(1) 近年、インターネット上の膨大な情報、“ビッグデータ”を教師情報として機械学習に活用することにより、画像等のメディアデータの内容を機械が自動的に読み取る技術が発展してきている。特に、画像中に写っている物の名前を判定する一般画像認識技術は、1000 カテゴリーの候補の中から正解のカテゴリーをトップ5位以内に絞ったときの正解率85%程度を達成しており、実用レベルに近づいてきている [Krizhevsky et al. 2012]。機械が画像を認識して物の名前を判定することができれば、画像に対する自動タグ付けが可能となる。すなわち、キーワードによる検索が可能な画像データベースを自動的に構築することができる。

(2) インターネット上のコンテンツに対する一般画像認識技術の精度が高まってきている一方で、個人の部屋の中を撮影した写真等、雑然とした日常環境を対象とした画像の認識はいまだに困難である。日常環境下では、注目すべき物体の他にもゴミや消耗品等が散在しており、他の物体の認識を妨げている。ただし、日常環境下であっても、注目すべき物体がはっきりと写っている画像であれば、認識および画像の自動タグ付けが比較的容易である。日常環境下で撮影された写真を正しく認識するためには、個々の物体を背景から分離することが重要である。近年、ロボット研究分野では、写真だけでなく距離センサを用いて物体を背景から分離し、認識する技術が盛んに開発されている [Aldoma et al. 2013]。しかしながら、認識対象の物体群がロボットの目の前に存在するという前提がある。すなわち、ロボットがどのように動いて注目すべき物体を撮影するかという問題には焦点が当てられていないのが現状である。

2. 研究の目的

(1) 実世界における有用な画像データベースを自動構築し、キーワードにより当該画像データの検索を可能とするシステムの実現を図る。具体的には、自律移動ロボットにより周囲の環境を撮影し、取得した画像を認識および抽出する新たな方法を提案するものである。

(2) インターネット上で検索できる情報は、誰かが記述してアップロードした特定のものに限られる。また、昨今ではメモの代わりにスマートフォン等で身の回りの物品や環境・状況を撮影する人が増えてきたが、肝心なものに限って写真が無い(後で必要となる情報を自らが保存しているとは限らない)といった問題が生じる。そこで、自律移動ロボットが人にかわって周囲の有用な情報を写真に残し、人が検索可能なデータベースを自動的に構築することを目指す。

3. 研究の方法

(1) 実世界における有用な画像データベースを自動構築する自律移動ロボットの行動計画として、写真を撮るべき物体に注目する仕組みを創る。物体に注目する戦略として、本研究は色・距離センサを用いて環境中から「物体らしい」三次元パターンを検出する。

(2) 本研究ではロボットに色・距離センサを搭載し、三次元画像を取得して物体の探索に活用する。二次元の画像のみを用いる場合、遠くの物体が小さく写るため、注目するのが困難である。これに対し、三次元画像の中では物体の大きさが視点に依存しないため、遠くの物体も発見しやすい。特に本研究では物体検出をグラフマッチングの問題として定式化した手法を用い、三次元画像から得られる幾何的な情報を活用して物体検出の精度を高める。

4. 研究成果

(1) 第一に、三次元画像からの物体検出をグラフマッチング問題として定式化した場合のパラメータを自動学習する手法を提案した。本手法は特徴点マッチングに基づく物体検出を行う。特徴点マッチングは、1. 特徴点(キーポイント)検出、2. 局所特徴抽出、3. (参照画像中の特徴点群との)対応点探索、4. 誤対応点除去の4ステップからなる。4. 誤対応点除去の手法としてはRANSACが最も定番であるが、対応点集合全体に対して含まれる誤対応点の割合が大きいか場合にうまくいかない。そこで、グラフマッチングを使って二点对二点の比較、すなわちある対応点ペアと別の対応点ペアを比較して整合性を評価し、対応点集合の中で他の対応点との整合性がとれない誤対応点を除去する手法が提案されている。特に筆者らは、距離画像を用いることで二点間の三次元距離を測り、これを用いて対応点ペア同士の整合性を評価する手法を提案した(図1)。本手法は、剛体上の二点間のユークリッド距離が位置姿勢に依らないことを利用している。ユークリッド距離の代わりに測地線距離等を用いれば、非剛体物体のマッチングも可能である。

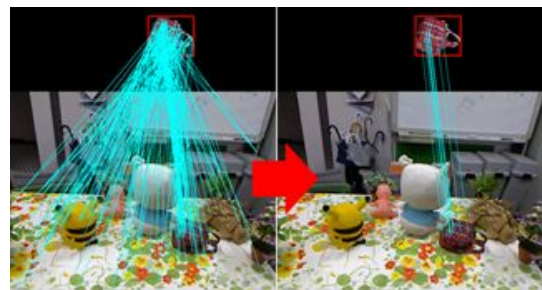


図1 グラフマッチングによる物体検出

特徴点マッチングのタスクそれ自体は物体検出タスクとは異なる点に注意されたい。

すなわち、特徴点マッチングはある観測された物体と参照用の物体とを対応付けることが目的であり、それらが同一の物体であるか否かの判断は別途必要である。筆者らは特徴点マッチングによって得られた対応点集合に対してその類似度(スコア)を計算し、物体検出タスクに用いる手法を提案した。ここでは、特徴点マッチングを Quadratic Assignment Problem (QAP)で定式化し、対応点ペア同士の整合性を示す類似度行列を用いる。この類似度行列をひとつのスコア(スカラー値)に統合する際に用いる重みを学習により最適化する。RGBD センサにより観測された剛体の実物体を検出する実験では、色によって区別された対応点ペアに対する重みを学習し、全ての対応点ペアを均一に評価する場合(ベースライン)に対して、相対的に約 15%の性能向上を確認した。また、非剛体物体検出実験では複数種類の類似度関数の統合重みを学習した。SHREC '10 correspondence benchmark を用いた評価実験では、非剛体物体のマッチングによく使用される類似度関数 1 を単独で使用した場合に Average Precision (AP)が 0.76, また別の類似度関数 2 を単独で使用した場合に 0.92, そしてこれら二つの類似度関数の出力を単純に平均した場合に 0.77, 提案手法を用いた場合に 0.98 という結果が得られ、提案する対応点集合類似度学習手法の有効性が確認できた。

(2) 未知の物体を検出する手法として、三次元空間から物体候補領域を事前知識なしに抽出する手法を提案した。2014 年以前の一般物体検出にはスライディングウィンドウが最もよく使われていた。スライディングウィンドウは Exhaustive Search とも呼ばれるように画像中のあらゆる領域から物体を探すため、計算コストが高い。HOG 等のヒストグラムベースの特徴量であれば比較的高速な処理が可能であるが、2012 年以降ブームとなったディープニューラルネットワークの CNN 特徴を単純に用いようとすると、ウィンドウ毎に特徴抽出を行う必要があり、計算時間は膨大となる。そこで近年は、前処理として各領域の物体らしさ(Objectness)を計算し、このスコアが高い領域のみに注目して特徴抽出と認識処理を行うアプローチが流行している。現在の物体検出手法のデファクトスタンダードである R-CNN は Selective Search により物体候補領域のウィンドウを列挙し、各ウィンドウから CNN 特徴を抽出して物体検出を行う。Selective Search は、第一に superpixel セグメンテーションにより画像全体を細かく領域分割し、第二に隣接する領域を類似度の高い順に階層的に結合していくことで、様々な大きさの物体候補領域を生成する。ここで、セグメント領域の隣接関係は二次元画像上で判断される。筆者らは距離画像を用いて二次元画像を三次元空間へと

マッピングし, superpixel の代わりに色付きの三次元点群から得られる supervoxel を用いた 3D Selective Search を提案した(本手法はソースコードを公開している)。ここでは二次元画像上ではなく三次元空間におけるセグメント領域の隣接関係を評価するため、前景の物体のセグメントが隣接する背景のセグメントと誤って結合される心配がなく、より高精度に物体候補領域を求めることができる。また、Selective Search は二次元画像におけるセグメント領域の大きさを評価するが、3D Selective Search は三次元空間におけるセグメント領域の大きさを評価するため、視点から物体までの距離に依らず一定の粒度で領域が分割される。この結果として、近くに写っている物体が細かく分割されがちになる、あるいは遠くの物体が他の物体に吸収されてしまうといった問題が生じにくくなる。

オリジナルの Selective Search, BING と提案する 3D Selective Search を比較した実験結果を図 2 に示す。評価用データセットはよく知られたベンチマークデータセットである NYU depth dataset v2 を用いた。BING は物体らしさ(Objectness)をサンプルベースに学習するものであり、図 2 の凡例の()内には学習に用いたデータセットの名称を示した。物体検出率が学習に使用するデータセットに依存することが分かる。これに対し、Selective Search は学習を行わず、汎用的な基準で物体らしい領域をとらえることができる。画像 1 枚中の窓数が 100 以下のとき、提案する 3D Selective Search が最も性能がよいことが示された。

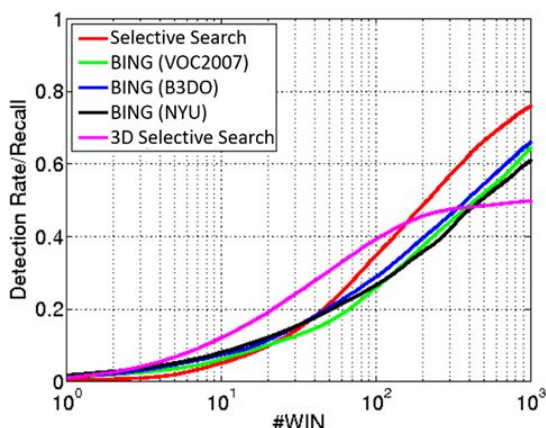


図 2 画像 1 枚中の窓数に対する物体検出率

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 8 件)

金崎 朝子. 三次元情報を活用した物体検

出の三手法 . Vision Engineering Workshop (ViEW), OS5-01, pp.284-287, 神奈川県横浜市, 2015.

Asako Kanezaki and Tatsuya Harada. 3D Selective Search for Obtaining Object Candidates. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), pp.82-87, Hamburg (Germany), 2015.

金崎 朝子, Emanuele Rodolà, 原田 達也. グラフマッチング学習を用いた RGB-D 画像からの物体検出 第 20 回ロボティクスシンポジウム, pp.432-437, 長野県北佐久郡, 2015.

Asako Kanezaki, Emanuele Rodolà, Daniel Cremers, and Tatsuya Harada. Learning Similarities for Rigid and Non-Rigid Object Detection. International Conference on 3D Vision (3DV 2014), pp.720-727, 東京都目黒区, 2014.

金崎 朝子, Emanuele Rodolà, Daniel Cremers, 原田 達也. 対応点集合類似度学習を用いた剛体・非剛体物体検出 Pattern Recognition and Media Understanding (PRMU), pp.13-18, 東京都江東区, 2014.

金崎 朝子, Emanuele Rodolà, 原田 達也. RGB-D 画像からの物体検出における対応点集合類似度の学習 the 32th Annual Conference of the Robotics Society of Japan (RSJ), 3I2-03, 福岡県福岡市, 2014.

Asako Kanezaki, Yusuke Mukuta, and Tatsuya Harada. Mirror Reflection Invariant HOG descriptors for Object Detection. IEEE 21st International Conference on Image Processing (ICIP 2014), pp.1594-1598, Paris (France), 2014.

Asako Kanezaki, Sho Inaba, Yoshitaka Ushiku, Yuya Yamashita, Hiroshi Muraoka, Yasuo Kuniyoshi, and Tatsuya Harada. Hard Negative Classes for Multiple Object Detection. IEEE International Conference on Robotics and Automation (ICRA 2014), pp.3066-3073, Hong Kong (China), 2014.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

6 . 研究組織

(1)研究代表者

金崎 朝子 (KANEZAKI, Asako)

東京大学・大学院情報理工学系研究科・助教

研究者番号 : 00738073

(2)研究分担者

なし

(3)連携研究者

なし