

複数特徴量を統合したパーティクルフィルタによる物体の検出と追跡

Muhammad Attamimi† 水谷 了† 中村 友昭† 長井 隆行†

船越 孝太郎‡ 中野 幹生‡

†電気通信大学大学院電気通信学研究科

‡(株)ホンダ・リサーチ・インスティテュート・ジャパン

1 はじめに

情景中の物体を認識するためには、まずその物体を検出する必要がある。物体検出を行う際に要求されることは、対象物を見逃さずに検出することである。また動き回るロボットによる物体検出では、高速な処理が要求される。そこで本稿では、対象物を含む候補領域を高速に検出する手法の開発を目指す。複雑な背景の中から物体を検出する際には、次の2つの問題点が存在すると考えられる。1つ目は、単一の特徴量を用いて検出を行うのが困難であることである。例えば、色が豊富な物体であれば色情報を用いた検出が可能であるが、そうでない場合、他の情報を利用する必要がある。2つ目は、実環境の中に置かれた物体の姿勢は任意であることである。この場合、回転も考える必要がある。計算量が増大してしまう。これらの問題に対して本稿では、色情報、テクスチャ情報、3次元情報など複数の特徴量を用いたパーティクルフィルタによる物体検出手法を提案する。また提案法では、検出する物体を複数の矩形領域に分け、それぞれを独立に検出することで2つ目の問題点を解決する。

従来、パーティクルフィルタを用いた物体検出・追跡は多く提案されている [1]–[3]。しかし、多くは色などの単一の特徴を用いており、それらの適応的な重み付けは考えられていない。また、3次元情報を用いた物体の抽出手法も検討されていない。

2 提案手法

2.1 検出システムの構成

本稿では、文献 [4] で著者らが提案した2台の CCD カメラと赤外線 TOF カメラを組み合わせた視覚センサを用いることを前提とする。このセンサを用いることで、シーンの色情報と奥行き情報が位置合わせされ、10fps 程度の速度で取得できる。従って、物体を検出するために利用できる情報は、色情報、テクスチャ情報、3次元情報である。ここで、これらの情報をどのように組み合わせるかが問題となる。物体を検出する際にどの特徴が有効であるかは、物体の性質や探索する環境に依存する。例えば、カラフルでテクスチャの多い物体は、色やテクスチャ等、どのような特徴を用いても検出可能であるが、白い食器のような色味のない単色の物体は、色やテクスチャで検出することが困難である。さらに、これは探索する環境によっても影響を受ける。黄色いぬいぐるみを検出することは、色情報を利用することで容易であるが、背景が同色の場合や照明条件が悪い場合には困難となる。また、色味のない単色の物体であっても、背景がカラフルであれば、色味のないことが検出するための情報として意味を持つことになる。従って、どの情報を使うべきかを事前に決めることは困難である。そこで提案手法では、パーティクルフィルタにおける尤度計算において、その状

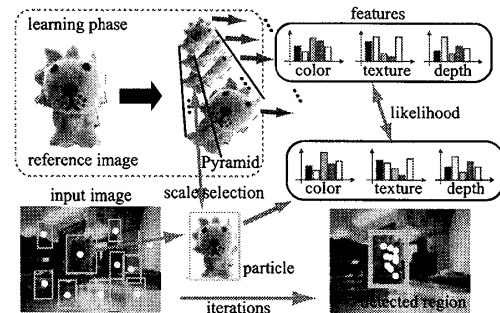


図 1: 検出システムの構成

況に応じた重みを計算することで、適応的に特徴が選択される仕組みを考える。

提案する物体検出システムの構成を図 1 に示す。物体の学習は、文献 [4] と同様に人が物体を手で把持し、ロボットに見せることで行う。この際、対象物体の特徴量である色情報、テクスチャ情報、3次元情報の特徴量を抽出し記録する。また、スケール変化に対応するために、物体画像のスケールを変化させ特徴抽出を行う。検出時は、パーティクルをシーン内にばらまき、リサンプリングを繰り返した結果パーティクルが集中した位置を候補領域として検出する。追跡は、検出された位置を初期位置とする点が異なるだけであり、基本的な処理の流れは検出と同様である。

2.2 画像特徴量

ここで特徴量に要求される性質は、画像内回転に対する不変性である。スケールに関する不変性は、奥行き情報を用いた適切な参照画像サイズの選択によって保証するため、特徴量自体で考慮する必要はない。

2.2.1 色情報

色情報としては、HSV 表色系を用いた色ヒストグラムを用いる。具体的には、色相と彩度の2次元ヒストグラムとする。量子化数はそれぞれ 32 と 10 である。

2.2.2 テクスチャ情報

テクスチャの特徴を表現する手法として、SIFT が挙げられる。SIFT はスケールと回転に対する不変性を持っているが、処理時間を考えると高速な物体検出には不向きであると言える。一方、Census Transform Histogram (CENTRIST)[5] は、高速に計算することが可能な特徴量であるが、回転に対して不変ではないという問題がある。また、8 近傍の画素値との比較だけであるため、複雑な背景の中から物体を検出するのに十分な表現力があるとは言い難い。そこで本稿では、対象画素値とその 8 近傍の画素値の差を 3 段階に分けて記述することで、特徴量の表現力を増加させる。

回転不変性への対処として、まず従来の 8 近傍との比較を、対象画素を中心とした円周上の 8 点との比較に変更する。さらに、8 点との比較で得られた 8 個の 3 段階 (0, 1, 2) の数値列を巡回させ、それによって得られる全 8 通りの値を 1 つのピンとする。最終的にテクスチャ情報は、ピンの数が 834 のヒストグラムとして表現される。

Object Detection and Tracking Using Particle Filter with Integrated Multiple Features

†Muhammad ATTAMIMI, Akira MIZUTANI, Tomoaki NAKAMURA and Takayuki NAGAI

‡Kotaro FUNAKOSHI and Mikio NAKANO

†The University of Electro-Communications

‡Honda Research Institute Japan, Co., Ltd.

2.2.3 3次元情報

3次元情報としては、各画素における距離情報をヒストグラムとして利用する Histogram of Depth(HOD) を用いることを提案する。学習時に得られた奥行情報の平均を求め、 $z_{min} = \bar{z} - k$, $z_{max} = \bar{z} + k$ を決定する。ただし k は定数であり、平均値からの最大の距離を表す。これらの値を用いて、奥行情報 z を次のようにビン b に投票する。ただし、 N はビンの数を表す。

$$b = \left\lfloor N \frac{z - z_{min}}{z_{max} - z_{min}} \right\rfloor \quad (1)$$

また、パーティクル領域中の HOD を計算する際に、背景部分が悪影響を及ぼす可能性がある。そこで、奥行情報をそのまま使うのではなく、領域内の最小の距離からある閾値内の情報のみを使用する。

2.3.1 パーティクルフィルタによる物体検出

2.3.1 複数特徴量の統合

本稿では、あるシーン中に置かれた物体に対して、一般的なパーティクルフィルタのプロセスを数回繰り返すことで物体検出を行う。 i 番目のパーティクルの尤度 P_i を、

$$P_i \propto \exp \{-w_c \lambda_c (D_c^i)^2 - w_t \lambda_t (D_t^i)^2 - w_s \lambda_s (D_s^i)^2\} \quad (2)$$

により計算する。ただし、 D_c , D_t , D_s はそれぞれ、色、テクスチャ、奥行きに対する Bhattacharyya 距離である。また、 λ_* はそれぞれの特徴に対する距離のダイナミックレンジを調整する係数であり、事前に決定しておく。 w_* は各特徴量に対する重みであり、繰り返し毎に次のように更新する。

$$w_* = \min \{1 / \min(\lambda_*(D_*^i)^2), 1\} \quad (3)$$

これは、ある特徴量が全てのパーティクルにおいて距離が大きい場合、その特徴量が対象シーンで有効に機能せずに、結果として全てのパーティクルの尤度を下げってしまうことを避ける働きをする。つまり、輝度が非常に低く色情報が使えないようなシーンでは、全てのパーティクルで距離が大きくなるため、重みは小さな値となり色情報は自動的に無効化される。逆に、物体の色と背景が同じで、どの位置でも距離が小さくなってしまいうような場合も、全てのパーティクルにおいて色に関する尤度が 1 に近い値をもつことになり、色情報が無効化されることになる。

2.3.2 距離に基づくパーティクルサイズの決定

各パーティクルの窓サイズは、パラメータの 1 つとしてサンプリングすることが一般的であるが、提案するセンサ構成では精度の高い奥行き情報を取得することが可能であり、窓サイズを物体のサイズとパーティクルの窓中心の距離から決定することができる。従って、パーティクルフィルタのパラメータは、 (x, y) の 2次元である。また、特徴量がスケール不変ではないため、距離によって決定したパーティクルの窓サイズに最も近いスケールを事前に計算した参照画像のピラミッドから選択し、そのスケールにおける特徴量を使って尤度計算を行う。

2.3.3 参照画像の矩形領域への分割

使用する特徴量は回転不変であるが、パーティクルフィルタにおいて回転の探索を行わないため、たとえばペットボトルのような縦長の物体が横に置かれていた場合の検出が困難になると考えられる。そこで、参照画像を複数の正方形領域に分割することで対処する。また、回転不変性と物体の境界領域の誤差を考慮して、ガウス窓によって重み付けをする。このような、参照画像の分割は回転探索の問題だけでなく、オクルージョンによって一部が隠れている場合などの探索への対処ともなる。

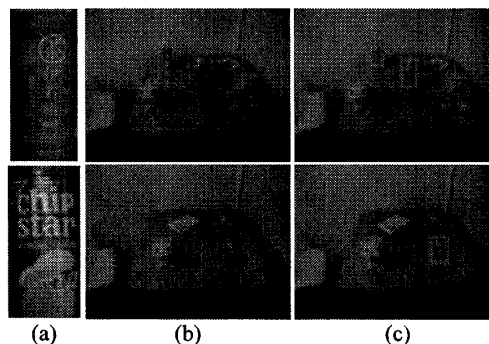


図 2: 検出の例 (上) 複数抽出の例 (下) 姿勢の異なる例



図 3: 変化する照明条件における追跡の例

2.3.4 候補領域の決定

パーティクルフィルタで計算した N 個のパーティクルの尤度 w_i を利用して、

$$P(x, y) = \frac{1}{P_{max}} \sum_{i=1}^N \frac{w_i(x, y) - w_{min}}{w_{max} - w_{min}} \quad (4)$$

より確率マップを求める。この確率マップを用いて、確率の値がある閾値を超えたところを候補領域とする。

3 実験

提案手法を用いて、実際のリビングルーム環境において物体の検出と追跡を行った。図 2 に検出の例を示す。図 2 上は、シーン中に対象物が 2 つある例である。ここでは、両方を検出できていることが分かる。また、図 2 下は、学習時と検出対象の姿勢が異なる例である。この場合も正しく領域が検出されている。一方、図 3 は追跡の例である。この例では、追跡の途中で部屋の照明を落とし、照明条件を大きく変化させている。提案手法では、このような条件でも依然として追跡に成功していることが分かる。

4 まとめ

本稿では、シーン中の高速な物体検出システムを目指し、パーティクルフィルタを用いた手法を提案した。提案手法では、色情報、テクスチャ情報、3次元情報を適応的な重み付けによって統合しており、様々な環境で物体を検出することができる。今後の課題としては、物体の様々な視点方向からの情報に対応することが挙げられる。

参考文献

- [1] K.Okuma *et al.*, "A Boosted Particle Filter: Multitarget Detection and Tracking", in Proc. of ECCV04, pp.28-39, 2004
- [2] J.Czyz *et al.*, "A Particle Filter for Joint Detection and Tracking of Color Objects", IVC(25), no.8, pp.1271-1281, 2007
- [3] 小畑ほか, "パーティクルフィルタを組み込んだパラメトリック固有空間法による三次元物体の検知・姿勢推定・追跡の同時実行", 電学論 C, vol.129, no.5, pp.838-845, 2009
- [4] 水谷ほか, "赤外線 TOF カメラと CCD カメラのキャリブレーションによる 3 次元センサの実現と画像処理への応用", 電気学会研究会資料, IM-09-60~74, pp.69-74, 2009
- [5] J.Wu *et al.*, "Visual Place Categorization: Problem, Dataset, and Algorithm", in Proc. of IROS09, 2009