

【研究ノート】

生活支援システム構築のための複数局所特徴量による一般物体認識

Generic Object Recognition with Multiple Local Features for Daily Life Support System

小 渡 悟
Satoru ODO

1. はじめに

近年、高齢者や障害者の生活支援を目的として、ロボットを活用した自立歩行支援や移動補助の研究 [1], 室内にパン・チルト・ズームが可能なカメラを設置し、ジェスチャ認識技術を用いて部屋全体を知能ロボット化する研究 (インテリジェントルーム) [2] が行われている。これからの高齢化社会を考えるに将来的には介護や子守といった分野にまでロボットが使用されることを想定すると、なによりも安全であるということが求められる。その中で重要な技術要素としてロボットが人間と同等な環境認識が行えることがあげられる。

これまで我々は、センサの装着やコントローラの使用なしに、ヒトの自然な動作にて、コンピュータ、通信機器、家電製品、ロボットなどの身の回りのものをジェスチャ推定により操作するマン・マシンインタフェースシステムについて提案を行ってきた [3]-[5]。これまでのシステムでは操作対象となる機器に識別用マーカを貼りつける必要があり、またそれ以外の環境にある生活用品等を識別できないという問題があった。しかし、「電気スタンド」といった同一カテゴリにおいても見た目の形状の違いが大きく、特徴抽出、モデルの構築が困難であった。このような物体の認識のことは一般物体認識と呼ばれ、画像認識の分

野における困難な課題の一つとして考えられている [6]。

本研究では生活支援システムへの利用を想定し、識別用マーカを使わずに制御対象機器の推定、ならびに生活用品の推定を行うために複数の局所特徴量を用いた一般物体認識について検討を行った。このような画像からの一般物体認識が自動で行えれば、単なるマーカレスによる制御対象の識別だけでなく、ロボットの行動を決定するために重要な周囲の環境を定量的に認識することができ、ロボットを安定に稼働させることも可能となる。

2. 提案手法

2.1 使用特徴量

本研究では SURF(Speeded Up Robust Features) 特徴量 [7] と HOG(Histograms of Oriented Gradients) 特徴量 [8], Edgelet 特徴量 [9] を用いた。

(1) SURF 特徴量

SURF 特徴量は SIFT 特徴量 [10] のキーポイント検出と特徴量記述の処理を Box フィルタと積分画像を用いることで近似している。Box フィルタとは重みが一定となる平均化フィルタであり、Box フィルタを複数用いることにより、ガウシアンフィルタの二次微分の近似フィルタを得ること

ができる。特徴量の記述では、キーポイントを中心とした正方形の領域を 4×4 のグリッド領域に分割し、各グリッド領域について Harr-like を計算することで輝度勾配を求める。この輝度勾配の値をもとに 16 領域 \times 4 次元 = 64 次元のベクトルを得ることができる。

(2) HOG 特徴量

HOG 特徴量は SIFT 特徴量と同様に局所領域における輝度の勾配方向をヒストグラム化した特徴量である。どちらも局所領域における輝度の勾配方向をヒストグラム化した特徴量であるが、SIFT 特徴量が特徴点に対して特徴量を記述するのに対し、HOG 特徴量は一定領域に対する特徴量の記述を行う。HOG 特徴量は一つの局所領域内におけるエッジ方向ごとのエッジ強度に着目する。局所領域はブロックとして表現され、ブロックは複数のセルのエッジ方向ごとのエッジ強度ヒストグラムを結合した多次元ベクトルで表現される。

(3) Edgelet 特徴量

Edgelet 特徴量は、エッジの部分的繋がりをとらえる特徴量である。定義された形状パターンと入力画像の局所領域間でのエッジ方向の差異を基に特徴量を算出する。形状パターンは、直線的な繋がりや、円弧、およびそれらの対称性をとらえる。Edgelet 特徴量は、 K 個の画素から構成されており、各画素は位置と 6 方向に量子化されたエッジ方向からなる。

2.2 実験条件

評価実験には K. Saenko らが公開している共通データベース [11] を用いる。このデータセットには机、椅子、本棚、電気スタンドなどのインテリアからパソコン、携

帯電話、ペン、コップなどの日常生活において使用頻度の高い全 31 種類のカテゴリで構成されている。このデータセットは amazon での商品画像をまとめたもの、デジタルカメラで撮影された画像、Web カメラで撮影された画像の 3 種類に分けられている。データセットに含まれる画像例を図 1 に示す。左側から amazon での商品画像 2 点、デジタルカメラで撮影された画像 2 点、Web カメラで撮影された画像 2 点となっている。

2.3 識別方法

事前に学習データを用いて各特徴量ごとに SVM (Support Vector Machine) の境界面を学習する。入力画像に対して各特徴量を抽出し、学習済みの SVM を用いて各カテゴリの SVM スコアを算出する。各特徴量の SVM スコアを統合し、最も高い値が得られたカテゴリを認識結果とする。

識別結果の評価方法は、認識率 = (正しく分類されたデータ数) / (分類されるべきデータ数) とする。

3. まとめ

本研究では生活支援システムへの利用を想定し、識別用マーカを使わずに制御対象機器の推定、ならびに生活用品の推定を行うために複数の局所特徴量を用いた一般物体認識について検討を行った。共通データベースを用いて複数特徴量を統合した場合の有効性を確認することができた。これより、単なるマーカレスによる制御対象の識別だけでなく、ロボットの行動を決定するために重要な周囲の環境を定量的に認識することができ、ロボットを安定に稼働させることが期待できる。

参考文献

- [1] 武富 卓三, 山海 嘉之: “ロボットスーツ HAL による脳性麻痺患者の歩行支援に関する研究”, 生体医工学会, vol.50, no.1, pp.105-110 (2012)
- [2] 入江耕太, 若村直弘, 梅田和昇: “ジェスチャ認識に基づくインテリジェントルームの構築”, 日本機械学会論文集 C 編, vol.73, no.725, pp.258-265 (2007)
- [3] Satoru Odo: “Vision-based Motion Estimation for Human-Computer Interface using Spatio-Temporal and Depth Information”, The 43rd International Symposium on Robotics (ISR 2012), p.p.144-149 (2012)
- [4] 国吉勇人, 小渡悟: “深度センサ情報を用いた小型飛行ロボットの制御”, 第22回インテリジェント・システム・シンポジウム, 2A-2, CD-ROM (2012)
- [5] 小渡悟, 富田元将, 星野聖: “マン・マシンインタフェースの応用を目的とした単眼カメラによる手指形状推定”, 第19回インテリジェント・システム・シンポジウム, pp.554-559 (2009)
- [6] 柳井 啓司: “一般物体認識の現状と今後”, 情報処理学会論文誌. コンピュータビジョンとイメージメディア, vol.48, no.16, p.p.1-24 (2007)
- [7] H.Bay, A.Ess, T.Tuytelaars, and L.Van Gool: “Speeded-up robust features(SURF)”, Comput. Vis. Image Underst., vol.110, no.3, p.p.346-359 (2008)
- [8] N.Dalal, B.Triggs: “Histograms of oriented gradients for human detection”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.886-893 (2005)
- [9] B.Wu and R.Nevatia: “Detection and tracking of multiple, partially occluded humans by Bayesian combination of edgelet based part detectors”, Int. J. Comput Vis., vol.75, Issue.2, p.p.247-266 (2007)
- [10] D.Lowe: “Distinctive Image features from Scale-Invariant keypoints”, International Journal of Computer Vision, vol.60, no.2, p.p.91-110 (2004)
- [11] K. Saenko, B. Kulis, M. Fritz and T. Darrell: “Adapting Visual Category Models to New Domains” In Proc. ECCV, p.p.213-226, September (2010)



図 1 共通データベースの画像例