

動画像を用いた一般物体のカテゴリ識別に関する検討

A Study on Category Classification of Generic Objects Using Video Images

中村 彰吾¹ 出口 大輔¹ 高橋 友和² 井手 一郎¹ 村瀬 洋¹
Shogo NAKAMURA Daisuke DEGUCHI Tomokazu TAKAHASHI Ichiro IDE Hiroshi MURASE

名古屋大学¹
Nagoya University

岐阜聖徳学園大学²
Gifu Shotoku Gakuen University

1 はじめに

実世界の画像に含まれる物体を計算機が一般的な名称で認識することを一般物体認識と呼び、現在注目されている研究分野である [1]。一般物体認識のアプローチの一つに、画像を局所特徴量のヒストグラムで表現する Bag of Features [2] を用いた手法がある。本発表では、動画像から得られる撮影角度の異なる複数枚の画像をこの手法の入力とすることによって、画像 1 枚のみで識別するよりも識別率が向上することを示す。その結果、動画像を用いて一般物体のカテゴリを識別する手法が有効であることを報告する。

2 提案手法

本手法は入力される動画像として、物体の撮影角度が徐々に変化するものを想定している。また、出力は物体のカテゴリである。本手法は、以下の学習段階と識別段階で構成される。

学習段階: 全ての学習動画像に含まれる各画像から得られる SIFT 特徴を k-means クラスタリングによりベクトル量子化し、各画像をクラスタの重心である visual word のヒストグラムで表現する。その後、このヒストグラムを特徴量として SVM 識別器を構築する。

識別段階: 入力動画像中の各画像を、学習段階で得られる visual word のヒストグラムで表現する。その後、学習した識別器によって各画像のカテゴリを識別し、識別結果を入力動画像中の全画像で多数決することにより、最終的な識別結果を得る。

3 実験および考察

ミニカー（株式会社タカラトミー、トミカシリーズ）のカテゴリ識別を行った。識別対象のカテゴリは、セダン、バス、トラック、ハッチバックの 4 種類とした。表 1 に、各カテゴリに含まれる物体の例を示す。学習と入力に使用した台数の内訳は表 2 の通りであり、学習と入力に同一のミニカーを含まない。

今回は予備実験のため、動画像として、ターンテーブルを用いて水平方向の角度を 0° から 355° まで 5° ずつで撮影した画像系列を使用した（図 1）。学習動画像は、この動画像をそのまま用いた。入力動画像は、ミニカー毎に動画像を 8 分割（0° ~ 40°, 45° ~ 85°, ..., 315° ~ 355°）したものをそれぞれ用いた。また従来手法として、入力を画像 1 枚のみ（各ミニカーの 0°, 45°, ..., 315°）にしたものと比較した。

Visual word の数を 100 として実験を行った結果、提案手法と従来手法の識別率はそれぞれ 70.5%, 61.1% と

表 1 カテゴリに含まれる物体の例

カテゴリ	セダン	バス	トラック	ハッチバック
物体の例				
				
				

表 2 学習と入力に使用した台数の内訳

カテゴリ	セダン	バス	トラック	ハッチバック	計
学習	11	6	11	20	48
入力	10	5	10	19	44



図 1 実験で用いた画像の例

なった。この結果から、入りに 1 枚の画像を用いる場合よりも、動画像から得られた複数枚の画像を用いる方が識別率が高いことが分かる。これは、1 枚の画像からカテゴリを識別するよりも、多数の画像の識別結果を統合した方が安定した結果が得られるためと考えられる。

4 むすび

本実験により、動画像から得られる複数枚の画像を物体のカテゴリ識別の入力として用いることによって、画像 1 枚で識別するよりも高い識別率が得られることを示した。その結果、動画像を用いて一般物体のカテゴリを識別する手法の有効性を確認した。今後の課題として、撮影角度情報の利用による識別精度の改善が挙げられる。

謝辞

本研究の一部は文部科学省科学研究費補助金によった。

参考文献

- [1] 柳井, “一般物体認識の現状と今後,” 情処学論, Vol.48, No.SIG 16(CVIM 19), pp.1-24, 2007
- [2] G. Csurka, C.R. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual categorization with bags of keypoints,” Proc. ECCV2004 Workshop on Statistical Learning in Computer Vision, pp.1-22, 2004