

共起特徴と k 近傍法による追加学習型オブジェクト認識システム

中居 友弘[†] 伊藤 聡[†] 柴田 智行[†] 窪田 進[†]

[†] 株式会社 東芝 研究開発センター 〒 212-8582 神奈川県川崎市幸区小向東芝町 1

E-mail: [†]{tomohiro.nakai,satoshi13.ito,tomoyuki1.shibata,susumu.kubota}@toshiba.co.jp

あらまし オブジェクト認識において、事前に収集した学習用画像だけでなく、認識時に与えられた画像を使って辞書を更新することができれば、カテゴリの追加や認識精度の向上が可能になり、より高い利便性を実現することができる。本発表では、ユーザが自由にカテゴリの登録や画像の追加の可能な追加学習型オブジェクト認識システムを紹介する。学習に用いる画像数が少なくても十分な認識精度が得られるよう、記述力の高い共起特徴を画像特徴として用いる。また、追加学習時の処理負荷を軽減するため、認識手法として k 近傍法を用いる。直感的に利用可能なインターフェースにより、ユーザビリティの高いシステムを実現した。

キーワード オブジェクト認識、追加学習、共起特徴、k 近傍法

1. はじめに

近年、カメラ付きモバイル端末の普及と高性能化に伴い、カメラで撮影した画像に含まれるものを認識（オブジェクト認識）し、情報の検索を行うサービスが提案されている [1] [2]。これらのサービスは、従来のキーワードを用いた情報検索とは異なり、ボタンひとつで検索が可能であるため、より手軽に利用できるという利点がある。また、現実存在するオブジェクトすべてを検索のキーとして用いることができるため、非常に応用範囲の広い技術である。

オブジェクト認識では、あらかじめ収集した学習用画像を用いて識別器を学習し、与えられたクエリ画像がど

のカテゴリに属するかを判定して認識結果を求める。しかし、撮影環境の変化や認識対象のバリエーションにより、学習用画像とクエリ画像の見えが異なる場合、正しく認識できないことがある。学習用画像を追加し、識別器を再学習すれば認識精度を改善することができるが、識別器の学習に大きな計算コストを要する学習手法では、ユーザが気軽に学習データを追加することは難しい。

本発表では、認識手法に k 近傍法を用いることで、学習データの追加を容易にしたオブジェクト認識システムを提案する。k 近傍法では学習データから明示的に識別境界を求める必要はなく、辞書データベースに学習データを追加するだけでよい。そのため、学習データを逐次的に追加して学習する追加学習が可能になる。また、表



図 1 オブジェクト認識システムのインターフェース。

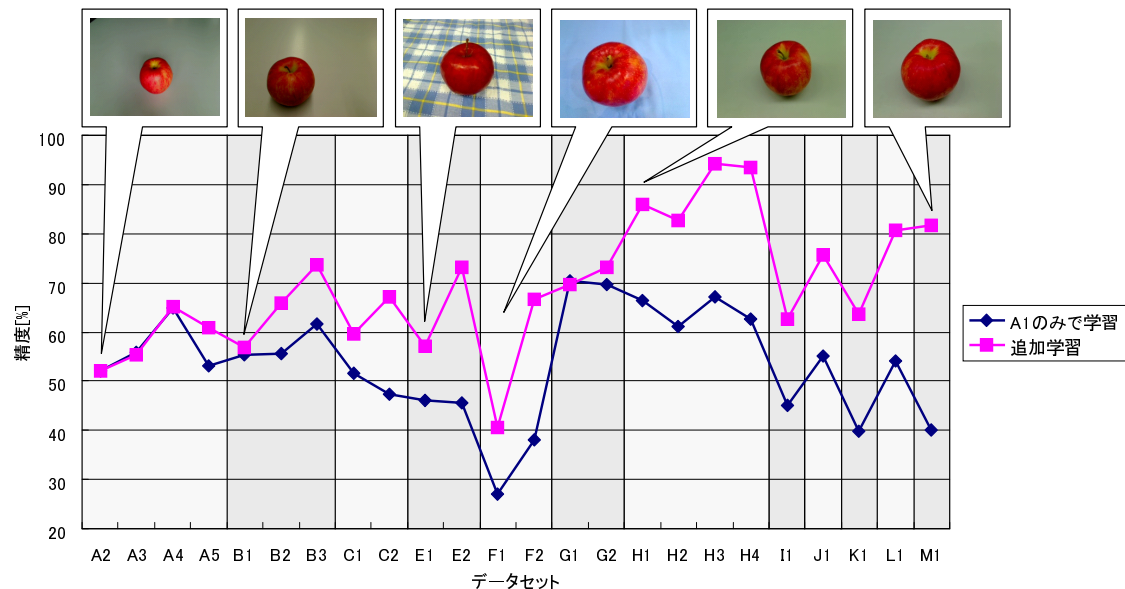


図 2 データセットの画像の例と、辞書固定時と追加学習時の精度の比較．

現力の高い共起特徴を用いることで、少ない学習サンプルでも高い認識精度が得られるようにし、新たなカテゴリを登録する際のユーザの負荷を軽減した．

2. 提案手法

画像特徴抽出処理では、認識画像から [3] で提案された画像特徴を抽出する．これは、CoHOG (Co-occurrence Histogram of Oriented Gradients) [4] の色情報を利用した拡張である color-CoHOG, CoHED, CoHD, カラーヒストグラムという複数の画像特徴を連結したもので、色、テクスチャ、形状などの異種の特徴を併用することで高い表現力を実現している．抽出される画像特徴の次元数は 1,194 次元となる．高次元特徴は表現力が高いため、認識精度の向上が見込める一方、記憶容量の圧迫や処理時間の増大を招く．本手法では、画像特徴に低次元化処理を行い、50 次元の特徴量に変換する．

類似度計算処理では、 k 近傍法を用いて辞書データベースから特徴量の近傍の学習サンプルを探す．探索では [5] で用いられているハッシュを使った近似最近傍探索によって処理を高速化する．そして、見つかった学習サンプルの属するクラスの頻度から類似度を求める．

図 1 に示される出力結果から、ユーザが正しい候補クラスを選択して確定すると、選択されたクラスを正解として特徴量の追加学習を行う．学習・認識手法は k 近傍法であるため、識別境界の計算などの処理は行わず、単純に特徴量を辞書データベースに追加登録する．

3. 実験結果

追加学習による精度向上の効果を調べるため、実験を行った．実験では、18 種類の野菜・果物を撮影した画像を、個体および撮影環境ごとにまとめたデータセットを用いた．データセットに含まれる画像数は、10~200 枚である．なお、データセット名で同じアルファベットを

もつものは、同じ個体を撮影したものである．例えば、B1 と B2 は同じ個体を異なる環境で撮影したデータセットである．ただし、数字が同じでも撮影環境は同じとは限らない．

データセットを追加学習した場合と、データセット A1 (画像数 180) のみで学習した場合の精度を比較した．追加学習した場合は、A1 で学習して A2 を認識し、次に A1 と A2 を学習して A3 を認識するというように、辞書データベースに学習データが順次追加されていく．

結果を図 2 に示す．学習データの少ない A2~A4 では精度にあまり差はないが、学習の進んだ J1~M1 では追加学習した方が 20~40% 程度高くなった．これは、追加学習によってさまざまな撮影環境の学習データが追加され、環境ロバスト性が向上したためと考えられる．

4. ま と め

追加学習の可能なオブジェクト認識システムについて述べた．認識手法に k 近傍法を採用することで、学習データの逐次的な追加を可能にした．実験により、追加学習を行うことで環境ロバスト性が向上し、環境が変化しても高い認識精度が得られることが確認された．

文 献

- [1] “Google goggles”, <http://www.google.com/mobile/goggles/>.
- [2] “Snaptell”, <http://www.snaptell.com/>.
- [3] S. Ito and S. Kubota: “Object classification using heterogeneous co-occurrence features”, ECCV 2010, Vol. 2, pp. 209–222 (2010).
- [4] T. Watanabe, S. Ito and K. Yokoi: “Co-occurrence histograms of oriented gradients for pedestrian detection”, the 3rd Pacific Rim Symposium on Advances in Image and Video Technology, pp. 37–47 (2009).
- [5] T. Shibata and O. Yamaguchi: “Local fisher discriminant component hashing for fast nearest neighbor classification”, SSPR/SPR 2008, pp. 339–349 (2008).