

[ポスター講演] 自動選定された意味的属性に基づく物体認識

戸川 恵里[†] MartinKlinkigt^{††} 井上 勝文^{††} 黄瀬 浩一^{††}

[†] 大阪府立大学工学部

^{††} 大阪府立大学大学院工学研究科

〒 599-8531 大阪府堺市中区学園町 1-1

E-mail: [†]{togawa,klinkigt,inoue}@m.cs.osakafu-u.ac.jp, ^{††}kise@cs.osakafu-u.ac.jp

あらまし 本稿では、物体認識手法の中でも注目されている、意味的属性（属性）を用いる手法について述べる。本研究では、属性を用いる手法によって画像から特定物体を認識することを目的とする。しかし、属性を手作業で選定すると膨大な時間がかかるという問題がある。そこで本稿では問題解決の第一段階として、物体認識に有効な属性を自動的に得る、2つの手法を提案する。第1の手法は、DBpedia を用いて狭義で正確な属性を選定するものである。第2の手法は、WordNet を用いてより意味的に広義な属性を選定するものである。これらの提案手法を用いて選定した属性で実際に特定物体認識を行い評価した。

キーワード 特定物体認識, 意味的属性, DBpedia, WordNet

1. はじめに

物体認識の手法の一つとして、意味的属性（属性）を用いる手法が注目されている [1] [2]。この手法では、あらかじめ学習用画像をもとに、属性の有無を識別する属性識別器を学習し、クエリ画像の属性識別結果から物体を推測する。ここで属性とは、物体の色や形などの見た目や、学問的分類などの概念といった、物体に関わるすべての言語表現を意味する。本研究ではこの属性を用いる手法によって、画像から特定物体を認識することを目的とする。

属性を用いる既存手法の多くでは、物体認識に有効と思われる属性を手作業で選定している [1]。しかし、これには膨大な時間がかかるという問題がある。この問題を解決するため、本研究では、属性を用いた特定物体認識のための第一段階として、認識に有効な属性を自動的に得る2つの手法を提案する。1つは DBpedia^(注1)を用いて狭義で明確な属性を選定する手法 (DB法) である。もう1つは、WordNet^(注2)を用いてより広義な属性を選定する手法 (WN法) である。

2. 手 法

本節では、まず属性候補を自動で得る DB法と WN法について述べ、次に候補の中で画像から得られる属性を用いる物体認識法について述べる。

まず DB法について述べる。DBpedia は、Wikipedia からグラフ構造化されたデータを取り出した知識ベースである。DBpedia によって、【Kinkaku-ji】という物体が『Buddhist_temple_in_Kyoto_Prefecture』のようなカテゴリに所属

することが分かる。DB法では、これらのカテゴリのうち、2つ以上の物体を含むものを属性候補とする。DBpedia のカテゴリは物体との対応関係が非常に明確であるという利点がある。しかし、分類が細かく狭義であるため、物体認識に用いると、視覚的な特徴が共通する物体が異なるカテゴリに属してしまう点に注意が必要である。そこで本研究では、有名なシソーラスである WordNet を用いて DBpedia のカテゴリ名から広義な単語を検索し、これを属性とする WN法も提案する。WN法ではまず、複数の単語から成り立つ DBpedia のカテゴリ名から単語 N -gram を取り出し ($N \in \{1, \dots, n\}$, n は単語数), そのうち WordNet に登録されているものを属性候補とする。次に、これらの属性候補の最大5段階までの上位概念を WordNet で検索することにより、さらに広義な属性候補を追加する。追加の条件は、 N -gram から得た複数の属性候補に対して共通する上位概念で、そのうち最も下位の概念であることである。

次に、属性候補の中から、物体認識に有効なものを選択する。このために、データセットを学習用と検証用に分割し、画像特徴からどの程度属性を識別できるのかを調べる。属性識別器の学習のために、まず学習用画像から PCA-SIFT 特徴ベクトル [3] を抽出する。次に各属性について、属する特徴ベクトルと属さないものとの識別境界を χ^2 カーネル関数の SVM によって学習し、属性識別器を生成する。そして、検証用のデータセットを用いて、属性識別率が 60% 未満である属性を除く。以上のようにして残った属性の有無の組み合わせによって、物体を2値ベクトル表現する。クエリ画像が与えられた際には、各属性の有無を属性識別器によって判定し、クエリ画像を属性の有無によって2値ベクトル表現する。そして、データセット中にある物体と比較し、ハミング距離が近い順にランクづけしたものを物体認識結果とする。

(注1): <http://wiki.dbpedia.org/>

(注2): <http://wordnet.princeton.edu/wordnet/>

表 1 クエリ画像に対する属性識別率と物体認識システムの評価

	DB 法	WN 法
属性識別率	54.58 %	68.53 %
AUC	0.76	0.67

3. 実験と考察

本実験では提案手法による物体の精度を 2 つの指標で評価した。本研究では、DBpedia 上で『Visitor_Attractions_in_Japan』という単語に関連する 828 個のカテゴリと、1,111 物体、46,526 枚からなる画像データセットを用いた。属性識別器の学習のためにデータセットの半分を用い、残りを検証に用いた。クエリ画像として、データセットにある 19 物体について、web 上から新たに 165 枚の画像を収集した。そして、DB 法で 174 個、WN 法で 64 個の属性を選定し、クエリ画像の識別を行った。評価のために、2 つの指標を用いた。1 つは属性識別率の平均で、クエリ画像からどの程度属性を識別できるかを知ることができる。もう 1 つは、横軸に物体認識結果のランク、縦軸にランクまでの正解物体数をとって正規化した ROC 曲線の Area Under Curve(AUC) である。AUC は 1 に近いほどシステム全体の性能が良いことを表し、属性識別率が良いときには、識別した属性が物体認識に有効であるかを知ることができる。

表 1 は物体認識の結果であり、属性識別率は WN 法が約 14 % 上回った。このことから、WN 法の属性が、DB 法の属性と比べてより画像特徴に関連が深いと言える。しかし、表 1 の AUC の値は、WN 法が 0.09 下回った。このことから、WN 法では、正しい属性が識別されているものの、それらの属性から物体を認識するのは難しいことが分かる。逆に DB 法の属性は個々の物体との対応が明確である、即ち特定性が高いため、属性識別が出来れば正解の物体がより正確に推測できた。WN 法では、複数の単語から成り立つ DBpedia のカテゴリ名を分解した際にその特定性が失われたと考えられる。

今後の課題としては、DBpedia と WordNet による属性を併用して物体認識することが挙げられる。

謝辞 本研究の一部は日本学術振興会科学研究費補助金基盤研究 (B) (22300062) の補助による。

文 献

- [1] C. H. Lampert, et al.: "Learning to detect unseen object classes by between-class attribute transfer", Proc. of CVPR, (2009).
- [2] G. J. Qi, et al.: "Towards cross-category knowledge propagation for learning visual concepts", Proc. of CVPR, (2011).
- [3] Y. Ke, et al.: "Pca-sift: A more distinctive representation for local image descriptors.", Proc. of CVPR., (2004).