

Chinese Restaurant Processを用いた Visual Wordの作成

住岡 浩之 外山 託海 岩村 雅一 黄瀬 浩一

大阪府立大学大学院工学研究科

{sumioka|takumi}@m.cs.osakafu-u.ac.jp {masa|kise}@cs.osakafu-u.ac.jp

1 はじめに

物体の画像を計算機に認識させることを物体認識という。物体認識では、Visual Word と呼ばれるものを用いる手法がある。Visual Word を作成するには、まず各画像から局所特徴量を抽出し、この局所特徴量をクラスタリングする必要がある。クラスタリングに用いられる手法は、 K -means 法などが一般的である。 K -means 法は、指定した K 個のクラスターにデータ集合を分類する方法である。そのため、正しいクラスター数が分からない時、様々な K の値を試してみるしかないという問題がある。そこで、Visual Word を作成する際、局所特徴量のクラスタリングに Chinese Restaurant Process(CRP) [1] を使用することを提案する。クラスタリングに CRP を用いることで、クラスター数を自動推定しつつクラスタリングができる。本稿では、CRP を用いて Visual Word 作成し、この Visual Word を用いた物体認識の認識率を調べることが目的とする。

2 Visual Word の作成方法

Visual Word の作成方法を説明する。まず、様々なカテゴリの画像から局所特徴量を抽出する。この局所特徴量をクラスタリングする。そして、1 枚の画像から得られる局所特徴量がどのクラスターに何回属するか調べる。これによって、クラスター数と同じ次元数を持つ頻度ベクトルを 1 枚の画像に対して 1 つ作ることができる。この頻度ベクトルが各画像を表す特徴量となる。頻度ベクトルにおけるクラスターは、文書検索分野の文書ベクトルにおける単語 (word) に対応する。よって、クラスターを Visual Word と呼ぶ。

様々なカテゴリの画像全てから頻度ベクトルを作った後、この頻度ベクトルを各カテゴリで分ける超平面を SVM で計算する。そして、検索質問画像が与えられた時、この画像から頻度ベクトルを作る。このベクトルが超平面で分けられたどの空間に位置するか計算することで、検索質問画像の物体カテゴリを認識することができる。

本稿で使用するクラスタリング手法である CRP について説明する。CRP はクラスタリング対象となるデータ点が、既存クラスターと新規クラスターのどちらに分類されるかを表す事前確率分布である。CRP によると、所属データが多いクラスターに、クラスタリング対象データ点が分類される確率は高くなる。新規クラスターに分類される確率は CRP のパラメータによって設定される。クラスタリング対象データに応じてクラスター数が増えるため、クラスター数を自動推定できる。クラスター数を自動推定する他の手法として、 X -means 法 [2] がある。 X -means 法とは、 $K=2$ で再帰的に K -means 法を実行し、クラスターの分割前と分割後でベイズ情報量規準が改善しなくなるまで分割する手法である。

表 1: 手法別認識率

手法	クラスター数	正解数	正解率
CRP	127	46	27.1%
X -means	52	53	31.2%
K -means	50	44	25.9%
	100	38	22.4%
	150	42	24.7%

3 実験と考察

CRP, K -means 法, X -means 法, これら 3 手法によって作られた Visual Word による物体認識の認識率の違いを調べた。実験の学習用データとして Caltech101 の 5 カテゴリを選び、それぞれから 8 画像ずつ、テスト用データとして Caltech101 の同じ 5 カテゴリの残り 170 画像を用いた。学習用データから PCA-SIFT を用いて局所特徴量を 5,682 個抽出し、CRP, K -means 法, X -means 法でクラスタリングした。 K -means 法のクラスター数 K を 50, 100, 150 に変化させた。クラスタリング結果は、CRP で 127 クラスター, X -means 法で 52 クラスターとなった。このクラスタリング結果から Visual Word を作成し、認識実験をした。表 1 に実験結果を示す。

認識結果は、CRP は K -means 法より高い認識率となった。 K -means 法では $K = 50$ の時、最も正解率が高かった。 X -means 法が最も良い結果となった。 X -means 法のクラスター推定数である 52 クラスターは、最も正解率が高い K -means 法のクラスター数に近い。よって、 X -means 法はクラスター数を自動推定すると考えられる。

CRP のクラスタリング結果、所属データ数が 1 個のクラスターが 35 クラスターあった。CRP では、所属データ数が 1 個となるクラスターが作られやすいという傾向が見られた。これは CRP のパラメータ値によって改善できる可能性がある。CRP パラメータ値と認識率の関係を調べることを今後の課題とする。

4 まとめ

本稿では、CRP, K -means 法, X -means 法によって Visual Word を作成し、それぞれにおける物体認識の認識率を調べた。CRP を用いると、 K -means 法より高い認識率となった。

参考文献

- [1] 上田 修功, 山田 武士, “ノンパラメトリックベイズモデル”, 応用数理, pp.196-214, September 2007.
- [2] Dan Pelleg, Andrew Moore, “ X -means: Extending K -means with Efficient Estimation of the Number of Clusters”, ICML-2000.