

基于核函数的高阶特征表示方法*

解晓萌¹, 梁 鹏^{2†}, 王 成³, 吴伟文¹

(1. 华南理工大学 计算机科学与工程学院, 广州 510006; 2. 广东技术师范学院 计算机科学学院, 广州 510665; 3. 仲恺农业工程学院 计算机科学与工程学院, 广州 510225)

摘 要: 为了解决高阶局部特征带来的计算复杂度提高问题, 提出一种基于核函数的高阶局部特征表示方法。通过在两幅图像的局部特征之间进行比较, 将特征空间映射到几何不变空间, 统计高阶局部特征构建核函数, 并结合支持向量机进行多类目标图像分类实验。实验结果分析表明, 该方法在提高分类准确率的同时, 所需的计算时间只与局部特征的个数呈线性增长。

关键词: 高阶局部特征; 图像分类; 核函数

中图分类号: TP301.6

文献标志码: A

文章编号: 1001-3695(2012)11-4350-03

doi:10.3969/j.issn.1001-3695.2012.11.090

High order local feature representation method based on kernel function

XIE Xiao-meng¹, LIANG Peng^{2†}, WANG Cheng³, WU Wei-wen¹

(1. School of Computer Science & Engineering, South China University of Technology, Guangzhou 510006, China; 2. School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou 510665, China; 3. School of Computer Science & Engineering, Zhongkai University of Agriculture & Engineering, Guangzhou 510225, China)

Abstract: In order to solve the computation increase caused by high order local features, this paper proposed a novel high order local feature representation approach based on kernel function. It first compared local features extracted from two difference images. Then transformed the features in the two images to the geometry invariant spaces. Next, it constructed kernel function by counting high order features. Finally, it performed the multi object classification experiment by using support vector machine. The experimental results show that proposed method has better precision and its computation time is linear to the number of local features in images.

Key words: high order local feature; image classification; kernel function

0 引言

目标图像分类的难点在于, 在几何变化、局部遮挡和光照变化等情况下从不同的类别中分类出不同的目标物体。局部特征对于几何变换有着较好的鲁棒性。在过去的几年里, 局部特征成功地在许多目标图像分类系统里面得到了应用^[1,2]。这种图像局部特征包括关键点、角点, 通过提取这些局部特征, 构建图像的中层语义, 从而达到图像分类的目的。这方面的代表性工作有 Li^[3] 的 bag-of-words 和 Lazebnik^[4] 的 spatial pyramid。然而, 随着样本数量的不断增加和目标物体的多样化, 对分类模型的分类能力提出了更高的要求。近年来, 针对提高分类模型的分类能力提出了一系列的方法, 但是分类能力的提高都使得分类算法的计算复杂度和空间复杂度被显著地提高。

Bag-of-words 模型只使用了目标的表面信息, 而丢弃了目标的空间结构信息, 因此对图像的表达非常高效、简洁, 只需较少的计算时间, 空间复杂度较低。然而, 目标图像的空间信息对分类来说是非常重要的, 丢弃了空间结构信息会减少分类模

型的分类能力。近年来许多方法尝试在 bag-of-words 模型中加入空间结构信息, 但都是以提高计算时间和空间复杂度实现的^[5-9]。Constellation models 方法^[10] 利用联合高斯表示物体的空间分布, 这种方法需要尝试所有的局部特征的结构表示形式, 因此需要大量的计算时间。此外, 该方法还需要给出目标物体的所在位置, 属于半监督的分类方法。另一种方法是采用高阶的局部特征来加入空间结构信息^[11,12], 高阶的局部特征是指不同数量的局部特征对之间的空间分布。这种方法的优点在于, 高阶的局部特征保留了局部特征表示的几何不变性, 无须特别指出目标物体的所在位置, 是一种无监督的分类方法。然而, 直接寻找高阶的局部特征是非常困难的。因为随着阶数 n 的增加, 高阶局部特征的数量会急剧地增多, 这为后续的特征聚类和分类带来了计算上的困难。因此, 为了减少计算复杂度, 许多方法都只使用了二阶的局部特征。

针对这一问题, 文中提出一种基于核函数的高阶局部特征表示方法。该方法在文献^[11] 的基础上, 将特征空间映射到几何不变空间, 可以在几何不变空间内统计任意阶的局部特征, 有效地保留了局部特征的空间结构信息。

收稿日期: 2012-04-15; **修回日期:** 2012-05-25 **基金项目:** 国家自然科学基金资助项目(60273064)

作者简介: 解晓萌(1973-), 男, 陕西长安人, 博士, 主要研究方向为图像处理、模式识别; 梁鹏(1981-), 男(通信作者), 讲师, 博士, 主要研究方向为图像处理、模式识别(cs_phoenix_liang@163.com); 王成(1977-), 男, 讲师, 博士, 主要研究方向为图像分割、模式识别; 吴伟文(1984-), 男, 博士, 主要研究方向为图像分割、模式识别。

1 基于核函数的高阶局部特征表示

高阶局部特征表示物体的局部结构空间信息,可以有效地减少特征匹配中的错误匹配。本文通过将局部特征进行位移对比,投影到几何不变空间。对应的局部特征对具有相同的位移,因此在几何不变空间内,对应的局部特征对会落到同一个区域内,如图 1 所示。

本文通过构造两个图像之间的内积核函数来实现从图像空间到几何不变空间的映射。给定图像特征空间 $\varphi: X$, 映射到位移空间 F , 内积核函数则表示为

$$K(I_1, I_1) = \langle \varphi(I_1), \varphi(I_2) \rangle \quad \forall I_1, I_2 \in X \quad (1)$$

SIFT 局部特征具有尺度不变性、旋转不变性和位移不变性等特征,被广泛地应用于图像检索和目标识别中^[3,4]。文中采用了 SIFT 作为局部特征,则图像 I 表示为

$$I = \{(w_1, x_1, y_1, s_1, \theta_1), \dots, (w_m, x_m, y_m, s_m, \theta_m)\} \quad (2)$$

其中: w 表示经过聚类后的 SIFT 特征描述符; x 和 y 分别为特征所在的位置; s 为 SIFT 特征的尺度; θ 为 SIFT 特征的方向。

根据文献[11]的定义,一阶局部特征表示为一对匹配的局部特征,二阶局部特征表示为两对匹配的局部特征, n 阶局部特征则表示为 n 对匹配的局部特征。第 n 阶的内积核函数表示为

$$K_n(I_1, I_1) = \langle \varphi_n(I_1), \varphi_n(I_2) \rangle \quad \forall I_1, I_2 \in X \quad (3)$$

其中:高阶的局部特征表示了较为稳定的结构空间信息,能够保证所匹配的局部特征既满足内容一致性,又满足空间一致性,因此本文给高阶的局部特征赋予较高的特征权重。

$$\begin{aligned} K(I_1, I_1) &= \sum_{n=1}^{\infty} K_n(I_1, I_1) \\ \hat{K}(I_1, I_1) &= \sum_{n=1}^{\infty} \alpha_n \hat{K}_n(I_1, I_1) = \\ &= \sum_{n=1}^{\infty} \alpha_n \left\langle \frac{\varphi_n(I_1)}{\|\varphi_n(I_1)\|}, \frac{\varphi_n(I_2)}{\|\varphi_n(I_2)\|} \right\rangle \end{aligned} \quad (4)$$

其中: α_n 表示特征权重,随着 n 的增加, α_n 也不断增大; $\hat{K}(I_1, I_1)$ 表示对核函数进行归一化。

2 几何不变空间映射

由于提取的 SIFT 局部特征是稠密的,当两幅图像有 N 个局部特征对时,二阶局部特征有 $\frac{N(N-1)}{2}$ 个,第 n 阶局部特征则有 $\frac{N(N-1)\dots(N-n+1)}{n!}$ 个。因此文献[11,12]都只采用了二阶的局部特征,从而降低了计算复杂度和空间复杂度。但是,高阶的局部特征具有良好的结构稳定性,尤其是三阶的局部特征,具有比二阶局部特征更为稳定的重现率。然而在图像空间内对所有匹配的局部特征对进行描述是很困难的。

本文提出一种几何不变空间映射方法,通过将图像空间内的局部特征映射到几何不变空间中,具有同样几何变换的局部特征对就投影到相应的几何不变空间区域内,这样只需统计在同一个几何不变空间区域内的局部特征个数就可以得到任意阶的局部特征。更重要的是,如图 2 所示,任意阶的局部特征维度只有四个维度 (x, y, s, θ) , 因此极大地减少了高阶特征维度增加带来的计算复杂度和存储所需的空间复杂度。

在图像空间内提取的 SIFT 局部特征具有几何变换不变

性,所以从图像空间投影到几何不变空间,也同样可以保证特征的几何变换不变性。

给定两幅图像 I_1, I_2 , 则图像表示为 $I = \{(w_1, x_1, y_1, s_1, \theta_1), \dots, (w_m, x_m, y_m, s_m, \theta_m)\}$, 计算两幅图像的几何不变空间映射为

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \Delta s \begin{bmatrix} \cos \Delta\theta & \sin \Delta\theta \\ \sin \Delta\theta & \cos \Delta\theta \end{bmatrix} \begin{bmatrix} x_j \\ y_j \end{bmatrix} - \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

其中: $\Delta\theta = |\theta_i - \theta_j|$, 表示两个 SIFT 局部特征的方向夹角; $\Delta s = \frac{s_i}{s_j}$, 表示局部特征 w_i 和 w_j 的尺度比例。

通过对每个几何不变空间区域内的投影个数进行 K-均值聚类,就可以统计得到不同阶的局部特征,第 n 阶的局部特征的聚类个数为 $k = n$ 。

3 实验及讨论

高阶局部特征加入了特征的局部空间结构信息,具有更稳定的匹配能力。因此为了评价文中提出的高阶局部特征,本文将高阶局部特征用于目标分类。实验采用的配置是 3.4 GHz 处理器,4 GB 内存。数据集是采用 Caltech-101 数据集,本次实验并没有使用到数据集提供的掩模图像。因为通过高阶局部特征匹配,可以有效地过滤掉空间不一致的局部特征,是一种无监督的目标分类实验。

本次实验的局部特征提取方法采用的是 SIFT 局部特征,聚类方法采用 K-均值聚类,分类器使用的是公开 SVM 库函数。实验的分类策略是每个类型抽取前 50 张用于训练,后 50 张作为测试集进行测试。分类策略采用多个二分类器进行分类,图像分类的结果由这些二分类器来投票选取,得票最多的类型就是该测试图像的类型。根据所提取的样本数量,特征码本的大小设置为 400。

实验流程如图 2 所示。

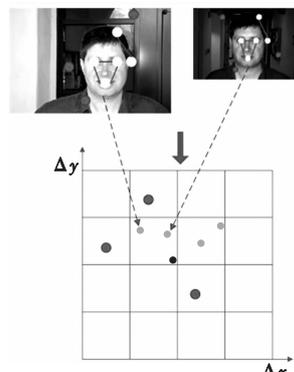


图1 文中方法的局部特征投影

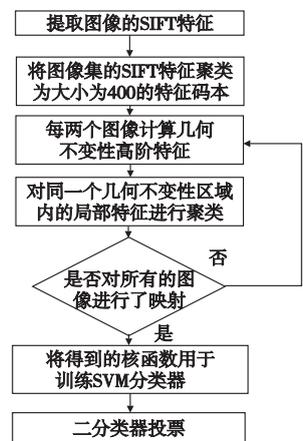


图2 实验流程

评价准则为准确率:

$$\text{准确率} = \frac{\text{正确分类的图像数量}}{\text{SVM 分类的图像数量}} \times 100\%$$

3.1 高阶局部特征对分类性能的影响

首先对文中提出的基于核函数的高阶局部特征方法对分类性能的提升进行实验,从 Caltech 数据集中提取出(蝴蝶、车辆、山羊、人脸、狗熊、步枪)六个类别的图像进行分类实验。实验分为两类分类(蝴蝶、车辆)、四类分类(蝴蝶、车辆、山羊、

人脸)和六类分类三个部分。图 3 是实验所用的部分图像。通过逐渐提高局部特征的阶数,获取分类的准确率和计算时间来评价方法的有效性。所有的实验都重复十次以获取平均准确率和分类误差。

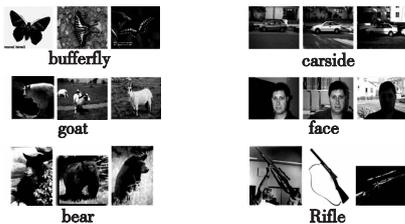
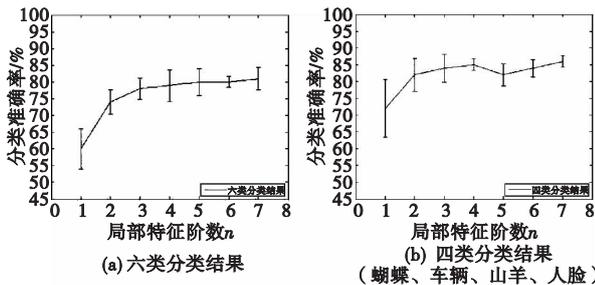
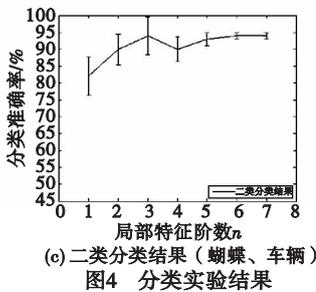


图3 Caltech-256部分示例图像

图 4 分别给出了二类、四类类和六类分类的实验结果。随着局部特征阶数 n 的增加,分类结果的准确率也在不断增加,当阶数 $n > 3$ 后,本文方法对分类准确率的提高不明显。这是因为绝大部分的空间结构可以用二阶和三阶局部特征表示。



(a) 六类分类结果 (蝴蝶、车辆、山羊、人脸)



(c) 二类分类结果 (蝴蝶、车辆)

图4 分类实验结果

此外,表 1 给出了文中方法在不同的局部特征阶数时的分类平均计算时间。给定一个图像内有 N 个特征,则对于 Constellation models 来说,三阶局部特征有 $\frac{1}{6}N(N-1)(N-2)$ 个,特征的维度增加导致计算复杂度为 $O(N^3)$ 。计算两个图像之间的特征匹配和映射到几何不变空间的时间为 $O(P)$,本文方法在匹配二阶局部特征时的最差计算复杂度是 $O(PN^2)$,其中 $O(N^2) > O(P)$ 。值得注意的是,局部特征之间匹配是采用了一对一的策略,因此每个局部特征不会都存在匹配的特征,因此实际中的计算复杂度只是 $O(PN)$ 。更重要的是,三阶局部特征计算无须再计算二阶局部特征的匹配,三阶局部特征的最差计算复杂度为 $O(PN^2 + PN^2) \approx O(PN^2)$,并没有随着特征阶数的增加而大幅地提高计算复杂度。

表 1 分类平均计算时间

	一阶局部特征	二阶局部特征	三阶局部特征
计算时间/s	0.571 5	0.890 2	1.635 6

3.2 方法对比实验

为了更好地体现本文方法的性能,将本文方法与文献[3]的 bag-of-words 模型和文献[4]的 spatial pyramid 模型进行对比。当本文方法取一阶局部特征时,实际上就相当于 bag-of-words 模型。测试的结果如表 2 所示。

表 2 人体肤色识别平均准确率 /%

准确率	本文方法		bag-of-words	spatial pyramid
	二阶局部特征	三阶局部特征		
二分类	90	94	82	86
四分类	82	84	72	80
六分类	74	78	60	72

从表 2 的实验数据可以得出,本文所使用的基于核函数的高阶局部特征方法通过加入空间结构信息,能更有效地表达局部特征的信息。通过几何不变空间的映射,使得高阶局部特征的计算复杂度大大降低,能够快速应用于图像分类。

4 结束语

本文提出了基于核函数的高阶局部特征表示方法,通过将特征空间映射到几何不变空间,可以快速获取任意阶的局部特征表示,高阶的局部特征可以有效地保持目标的空间信息。在 Caltech-101 数据库上的无监督分类实验表明,文中提出的高阶局部特征表示方法与半监督的局部特征表示相比较,可以提高分类准确率。但本文方法仍存在不足之处:由于特征空间映射到几何不变空间是在两幅图像之间进行对比, N 幅图像的数据库需要进行 $\frac{N(N-1)}{2}$ 次转换,如何构建更有效率的核函数,是今后工作的重点。

参考文献:

- [1] GRAUMAN K, DARRELL T. Pyramid match kernels: discriminative classification with sets of image features [C]//Proc of IEEE International Conference on Computer Vision. 2005:1458-1465.
- [2] ZHANG J, MARSZALEK M, LAZEBNIK S, et al. Local features and kernels for classification of texture and object categories; a comprehensive study [J]. International Journal of Computer Vision, 2007, 73 (2): 213-238.
- [3] LI Fei-fei. Visual recognition: computational models and human psychophysics [D]. California: California Institute of Technology, 2005.
- [4] LAZEBNIK S. Semi-local and global models for texture, object and scene recognition [D]. [S. l.]: University of Illinois at Urbana Champaign, 2006.
- [5] KIM G, FALOUTSOS C, HEBERT M. Unsupervised modeling and recognition of object categories with combination of visual contents and geometric similarity links [C]//Proc of the 1st ACM International Conference on Multimedia Information Retrieval. New York: ACM Press, 2008:419-426.
- [6] LEORDEANU M, HEBERT M. A spectral technique for correspondence problems using pairwise constraints [C]//Proc of the 10th IEEE ICCV. Washington DC: IEEE Computer Society, 2005: 1482-1489.
- [7] LEORDEANU M, HEBERT M, SUKTHANKAR R. Beyond local appearance: category recognition from pairwise interactions of simple features [C]//Proc of CVPR. [S. l.]: IEEE Computer Society, 2007: 18-23.
- [8] 刘扬闻, 霍宏, 方涛. 词包模型中视觉单词歧义性分析 [J]. 计算机工程, 2011, 37(19): 204-206, 209.
- [9] TIAN Qi, HUA Gang, HUANG Qing-ming, et al. Generating descriptive visual words and visual phrases for large-scale image applications [J]. IEEE Trans on Image Processing, 2011, 20 (9): 2664-2667.
- [10] NIBLES J, LI Fei-fei. A hierarchical model of shape and appearance for human action classification [C]//Proc of CVPR. Minneapolis: IEEE Press, 2007: 1-8.
- [11] LIU D, HUA G, VIOLA P, et al. Integrated feature selection and higher-order spatial feature extraction for object categorization [C]//Proc of CVPR. 2008: 1-8.
- [12] ZHOU Wen-gang, LU Yi-juan, LI Hon-qiang, et al. Spatial coding for large scale partial-duplicate Web image search [C]//Proc of the International Conference on Multimedia. New York: ACM, 2010: 511-520.