

基于互补特征和类描述的商品图像自动分类

贾世杰^{①②} 孔祥维^① 付海燕^① 金光^①

^①(大连理工大学电子与信息工程学院 大连 116023)

^②(大连交通大学电气信息学院 大连 116028)

摘要: 实现电子商务中的在线商品自动分类是电子商务智能化的迫切要求。该文研究如何运用图像特征和分类算法对商品图像的一些具体信息进行自动分类,如长袖衬衫和短袖衬衫、圆领T恤与V领T恤等。图像特征采用了具有互补特性的塔式梯度方向直方图(PHOG)和塔式关键字直方图(PHOW)相结合;在分类器设计方面提出了基于图像类描述的改进最近邻分类算法。实验结果证明本算法能使2类和3类商品图像分类正确率达到70%~99%,且能够实现快速实时分类,相对于现有方法有了明显提升。

关键词: 图像处理;商品分类;互补特征;类描述

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2010)10-2294-07

DOI: 10.3724/SP.J.1146.2009.01391

Auto Classification of Product Images Based on Complementary Features and Class Descriptor

Jia Shi-jie^{①②} Kong Xiang-wei^① Fu Hai-yan^① Jin Guang^①

^①(College of Electrical & Information Engineering, Dalian University of Technology, Dalian 116023, China)

^②(College of Electrical & Information, Dalian Jiao tong University, Dalian 116028, China)

Abstract: Auto-classification of online goods is a great need for intelligent e-commerce. Some valuable information of product, such as long sleeve T-shirt vs short one, round collar vs V-neck collar, can be tagged based on the image features and classification algorithms. As two complementary features, PHOG (Pyramid Histogram of Orientated Gradients) and PHOW (Pyramid Histogram Of Words) are adopted to extract and describe the features of product-images. An improved nearest-neighbor classifier based on the class descriptor of images is proposed. Experimental results show that the accuracies have been achieved between 70% to 99% on 2 or 3 real-time classification task, which is markedly improved compare to the existent result.

Key words: Image processing; Product classification; Complementary features; Class descriptor

1 引言

随互联网的普及和发展,电子商务逐渐进入了一个全新的时代,电子商务网站的数量急剧增长,出现了一批国内外知名的电子商务网站,如Amazon, ebay, 淘宝等。电子商务网站需要通过在线销售商品进行标注以方便用户进行搜索。目前情况下,这些标注仅仅说明商品的基本信息(元信息),如商品的名称、产地、尺寸、价格等,难以反映商品的完整特征。如女士皮鞋是圆头还是尖头, T恤衫是圆领口还是V型领口,休闲鞋鞋带是尼龙搭扣型还是细鞋带型等;这些特征都是用户可能感兴趣的潜在信息,但因为缺少进一步的标注,用户

只能通过浏览商品图片才能获得这些信息。如果在网站中设置图片分类过滤器,无疑能方便用户进行浏览。但要通过人工完成这些潜在兴趣信息的标注,对于商品数量和品种规模都很大的电子商务网站来说,无疑是非常费时费力的。如何通过图像内容特征完成在线商品的自动分类,是当前电子商务领域的迫切需求和前沿研究课题。

基于内容的图像分类(content-based image classification)是根据图像的视觉特征对图像进行语义分类。近几年基于内容图像分类的研究焦点是自然图像的场景分类(scene classification)^[1-3]和物体分类(object classification)^[4,5],主要采用有监督学习方法,通过对底层特征建模和中间语义分析来实现分类。目前研究文献中常用的测试图像数据库 Caltech 101^[1]和 Caltech 256^[2]已经达到101类和256类。与这些库中的自然图像不同,电子商务网站上

2009-10-29 收到, 2010-03-03 改回

国家自然科学基金重大项目(70890083)资助课题

通信作者: 贾世杰 jiajihie@dl.cn

提供的商品图像一般是比较理想的图片, 具有较少背景干扰, 目标比较单一; 这些特点使基于内容的商品图像分类更容易获得理想的分类正确率; 但这方面的专门研究较少。目前公开发表的文献中, 只有文献[6]探讨了运用基于内容图像分类技术实现商品图像标注的问题。文献[6]主要采用了基于 sift 的分级词包模型和 K 近邻分类方法, 在 Amazon 网站上搜集的商品图像库上进行测试, 报告的分类正确率为 66%~98%, 没有说明算法的分类速度问题。在文献[6]基础上本文在以下两个方面进行了改进: (1) 在图像特征提取和描述方面, 文献[6]通过稀疏采样方式获得兴趣点, 采用 sift 特征描述形成 128 维特征向量; 这种方法没有充分利用图像的空间分布信息; 并且文献[7]已证明, 相对于稠密采样, 稀疏采样不利于分类性能的提高; 本文采用稠密采样方式, 形成了两种具有互补特性的多级塔式结构特征: PHOG 和 PHOW, 并通过线性特征融合获得最终的特征表达。这种特征描述既考虑到了图像的形状特征, 又考虑到了图像的局部分布信息, 通过图像空间多分辨率分解构成的塔式结构和特征加权融合能够更完整、灵活地描述图像特征信息, 从而提高图像分类性能。(2) 在分类器设计方面, 本文在文献[8]基础上提出了基于图像类特征描述的改进最近邻分类算法, 通过计算图像到类(而不是图像到图像)的距离来实现商品图像分类。经过与文献[6]相同图像库的分类实验测试, 分类正确率能达到 70%~99%, 比文献[6]报告的实验结果有了明显的提升, 并且能够达到实时性的要求, 说明本文采用的方法是行之有效的。

2 商品图像特征提取与描述

图像特征提取与描述是进行图像分类的第一步。根据“丑小鸭定理”^[9], 没有与“假设”无关的天生优越的特征表达。对商品图像的分类, 应该针对商品图像的特点和特定分类要求, 选择合适的最具有区分特性(discriminative character)的特征集合。从电子商务应用的角度看, 用户一般更注重商品图像的形状信息和局部特征信息, 这些信息也自然成为图像分类的重要依据。本文采用文献[10]提出的 PHOG 方法和文献[3]提出的 PHOW 方法, 这两种方法都通过图像空间多分辨率分解形成多级塔式结构表示; 前者提取与描述图像形状特征信息, 后者提取并描述图像局部特征信息, 两者形成具有互补特性的特征表示集合, 本文通过线性加权方式得到商品图像的特征描述, 其中加权系数通过交叉验证方式获得。

2.1 塔式梯度直方图(PHOG)

梯度直方图(Histogram of Orientated Gradients, HOG)是描述图像形状信息的一种有效方法。HOG 特征通过提取局部区域的边缘或梯度的分布, 可以很好地表征局部区域内目标的边缘或梯度结构, 进而表征目标的形状^[11]。构造图像 HOG 的方法是首先将子图像划分为小尺寸的单元, 将梯度方向划分为 K 个区间(bin), 计算每个单元对应的用梯度幅值加权的梯度方向直方图, 将其表示为一个 K 维的特征向量。子图像中所有单元的特征向量联结起来, 即构成子图像对应的特征向量。HOG 实际上已经考虑到图像空间位置的分布, 但没有考虑到图像不同空间尺度划分表示对分类性能的影响。为此, Bosch 在文献[8]中提出塔式梯度方向直方图(Pyramid Histogram of Orientated Gradients, PHOG), 使用空间四叉树分解形成图像的多分辨率表示, 通过联结从低分辨率到高分辨率的多级梯度方向直方图来描述图像。假设设定级数为 $L=3$, 当前级数为 $l(l=0, 1, 2)$, 梯度方向划分为 20 个区间, PHOG 描述符就由 3 个梯度方向直方图特征向量顺序联结而成。 $l=0$ 时不进行空间划分, 将整个图像作为 1 个单元计算 HOG, 其维数为 20; $l=1$ 时将图像进行四叉树划分, 将图像划分为 4 个矩形单元计算 HOG, 其维数为 $20 \times 4=80$; $l=2$ 时将图像分解为 16 个矩形单元计算 HOG, 其维数为 $20 \times 16=320$, 最终形成的直方图是 $l=0, 1, 2$ 各 HOG 直方图的顺序组合, 其维数为 $20+80+320=420$ 。用整个图像的直方图“能量”(如 L_2 范数)对特征向量进行归一化处理, 可以进一步去除光照变化的影响。PHOG 生成示意图如图 1 所示。

2.2 塔式关键词直方图(PHOW)

词包(Bag of Words, BoW; 或称特征包 BOF, Bag of Features)模型的基本思想是借鉴文本分类技术, 将图像表示成一系列视觉关键词(visual words)的统计直方图。所谓视觉关键词就是由训练图像局部区域特征(如颜色、纹理、兴趣点)经过聚类形成的聚类中心, 视觉关键词的集合形成所谓词包(bags of words)。词包模型的构建分为以下几个步骤: (1) 自动检测图像兴趣点/兴趣区域或局部块; (2) 运用特征描述子(如 sift)描述局部区域; (3) 将图像描述符进行 K 均值量化后形成若干聚类中心, 即视觉关键词; (4) 计算图像包含各视觉关键词的数目, 形成视觉关键词直方图。词包模型示意图如图 2 所示^[12]。基于词包的图像局部特征在图像分类中获得了卓越的性能^[13]。

传统的词包模型忽略了图像的空间位置特征,

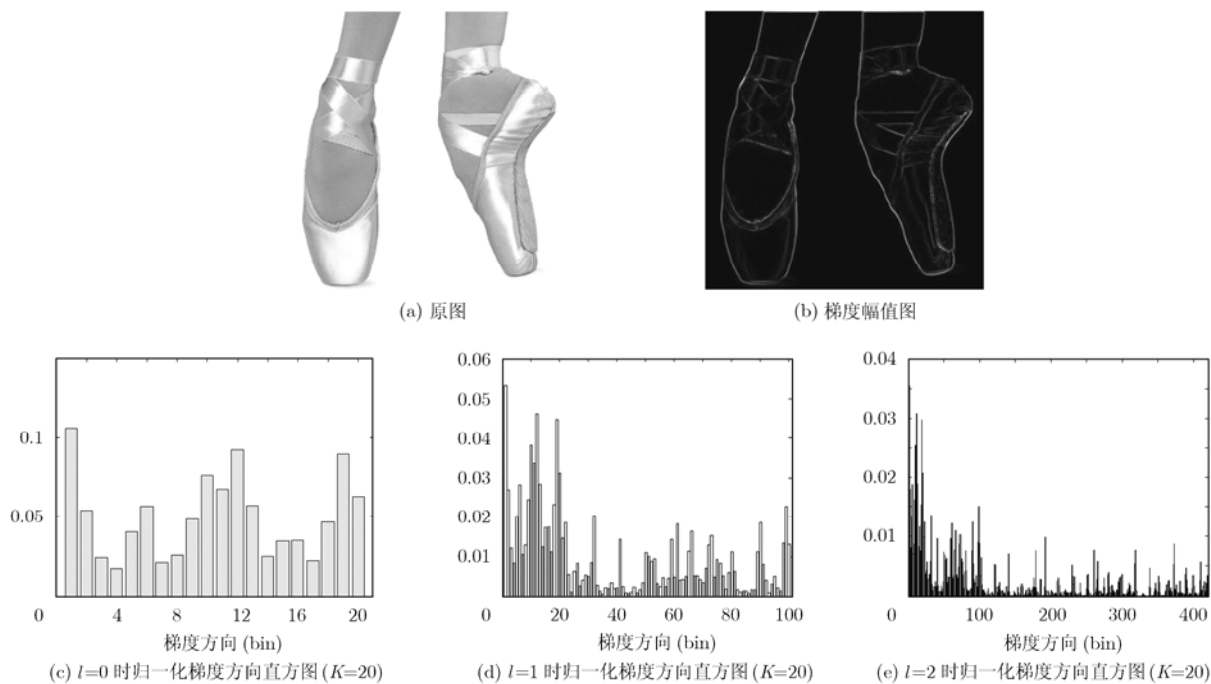


图 1 归一化 PHOG 示意图

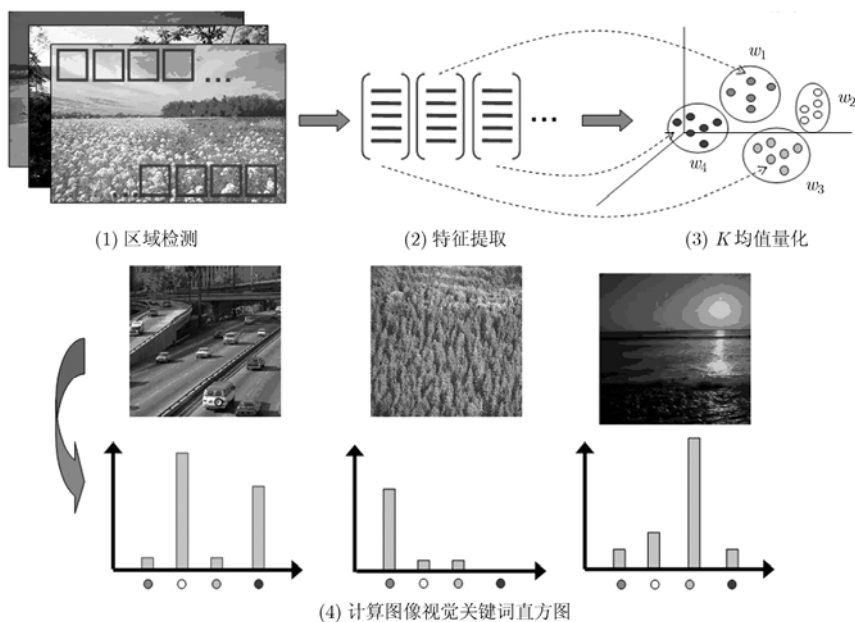


图 2 词包的形成与表示

并且采用稀疏采样方式,不利于图像结构特征的提取。由此文献[3]提出了一种基于空间塔式直方图的词包技术(PHOW, Pyramid Histogram Of Words),该方法在两个方面做了改进,(1)特征提取采用稠密采样(dense sample)方式,采样间隔设为8个像素,每个 16×16 的像素块使用sift描述符形成128维的特征向量。(2)通过对图像进行一系列空域四叉树分解,在特征空间形成从低分辨率到高分辨率表示的一系列视觉关键词直方图表示。PHOW示

意图如图3^[3]所示。本文设定分解级数为3($l=0,1,2$),词包大小为500,则最终形成的PHOW的维数为: $500+500 \times 4+500 \times 16=82500$ 。

3 分类器设计

考虑到商品图像在线服务需要快速分类的特点,本文设计了基于图像类描述的改进最近邻分类器。这种方法不需要学习训练过程,从而大大减少分类时间。传统图像分类算法中的最近邻和 K 近邻

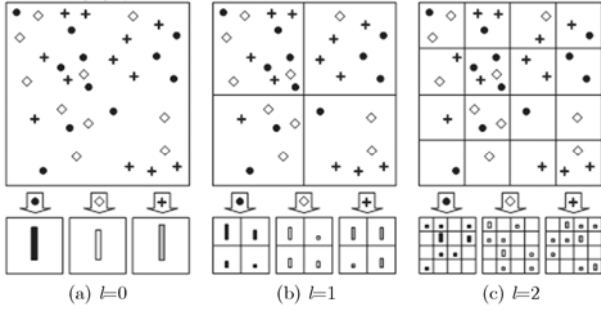


图 3 PHOW 示意图

方法都是通过计算图像到图像的距离来进行分类的。如文献[6]就采用了按图像距离加权的 K 近邻分类器, 通过计算查询图像到各标记图像类中所有图像的归一化距离, 将对应归一化距离之和最小的类别作为查询图像的分类结果。但根据文献[8]的研究结论, 采用图像到图像(image-to-image)距离的计算方法是导致最(K)近邻分类算法推广性能下降的一个重要原因。原因是相对于图像的复杂度, 标记类图像数目往往较少, 不能较完整地反映图像的类内变化的复杂性。受文献[8]启发, 本文提出基于图像类描述的最近邻分类算法。首先根据 PHOG 和 PHOW 互补特征形成每类图像的特征描述符, 然后通过计算查询图像与类特征描述符之间的距离, 得到查询图像与每个图像类之间(image-to-class)的距离, 将距离最小的图像类作为分类结果。

3.1 图像直方图之间的距离

在计算图像直方图距离之前, 首先对图像直方图进行归一化处理, 然后选择合适的直方图距离计算方法。计算图像直方图之间距离的方法有: 直方图相交法, 余弦距离法, chi-square 距离法等。在文献[10]中已证明相对于直方图相交法和余弦距离法, chi-square 距离法是一种性能较好的相似度计算方法。chi-square 距离法计算公式如式(1)所示

$$d(s_1, s_2) = 1 - \exp(-\chi^2(s_1, s_2)), \chi^2(s_1, s_2) = \sum_j \frac{(s_1[j] - s_2[j])^2}{s_1[j] + s_2[j]} \quad (1)$$

其中 $d(s_1, s_2)$ 表示两个直方图 s_1 与 s_2 之间的 chi-square 距离。

3.2 图像类特征描述符

根据 2.1 节和 2.2 节, 每幅图像都可以表示成 PHOG 和 PHOW 特征的集合; 而 PHOG 和 PHOW 又各有 L 种特征描述($l=0, 1, \dots, L-1$)这样每幅图像都可以表示成 $2L$ 个特征描述符, 本文取 $L=3$, 故共有 6 种特征描述符。

查询图像 Q 的特征 F_Q 可表示为

$$F_Q = \{F_{QPHOG}, F_{QPHOW}\} = \{\underbrace{f_{Qphog0}, f_{Qphog1}, \dots, f_{QphogL}}_L, \underbrace{f_{Qphow0}, f_{Qphow1}, \dots, f_{QphowL}}_L\} \quad (2)$$

假设某图像类 C 标记的图像数是 N_c , 下面构造图像类 C 的类特征描述符 F_C :

$$F_C = \{F_{CPHOG}, F_{CPHOW}\} = \{\underbrace{f_{Cphog0}, f_{Cphog1}, \dots, f_{CphogL}}_L, \underbrace{f_{Cphow0}, f_{Cphow1}, \dots, f_{CphowL}}_L\} \quad (3)$$

设 f_{Qphog}, f_{Qphow} 分别表示查询图像 Q 的 PHOG 和 PHOW 特征描述符, f_{Cphogj}, f_{Cphowj} 分别表示图像类 C 中第 j 幅图像的 PHOG 和 PHOW 特征描述符, 则 f_{Cphog}, f_{Cphow} 应满足以下条件:

$$d(f_{Cphogi}, f_{Qphog}) = \min_{j=1, \dots, N_c} d(f_{Qphog}, f_{Cphogj}), \quad i = 0, 1, \dots, L-1 \quad (4)$$

$$d(f_{Cphowj}, f_{Qphow}) = \min_{i=1, \dots, N_c} d(f_{Qphow}, f_{Cphowj}), \quad i = 0, 1, \dots, L-1 \quad (5)$$

3.3 图像到类(image-to-class)的距离

PHOG 特征能够较好地描述图像的形状信息而对图像的局部分布信息区分性差; PHOW 则能很好地描述图像的局部分布信息, 而对图像的形状信息区分性差; 两种特征具备一定互补性, 进行特征融合可进一步提高分类性能。本文采用线性组合的方式进行特征融合, 如式(6)所示:

$$d(F_Q, F_C) = \alpha \cdot d(F_{QPHOG}, y_{CPHOG}) + (1 - \alpha) d(F_{QPHOW}, y_{CPHOW}) \quad (6)$$

其中 $d(F_{QPHOG}, y_{CPHOG})$ 和 $d(F_{QPHOW}, y_{CPHOW})$ 分别表示以 PHOG 特征和 PHOW 特征计算的查询图像 Q 与图像类 C 之间的距离, $d(F_Q, F_C)$ 表示进行特征融合后查询图像 Q 与图像类 C 之间的距离。 α 的取值通过多重交叉验证的方法来确定。通过 α 的选择, 获得最具区分能力的特征表示。

另外, 不同分辨率的直方图对分类性能有不同的影响, 所以计算直方图距离时应该设置不同的权重系数。一般说来, 相对于低分辨率直方图, 高分辨率直方图对分类性能的影响更大一些。参照文献[3], 本文将 PHOW 和 PHOG 第 l 级的权重设为 $1/2^{L-l}$ ($l=0, 1, \dots, L-1$, L 是最高分解级数)。

$$d(F_{QPHOG}, F_{CPHOG}) = \sum_{l=0}^{L-1} \frac{1}{2^{L-l}} d(f_{Qphog}, f_{Cphog}) \quad (7)$$

$$d(F_{QPHOW}, F_{CPHOW}) = \sum_{l=0}^{L-1} \frac{1}{2^{L-l}} d(f_{Qphow}, f_{Cphow}) \quad (8)$$

最终的分类结果 \hat{C} 由式(9)确定:

$$\hat{C} = \min_C d(F_Q, F_C) \quad (9)$$

4 实验

4.1 图像库

为了同文献[6]的分类结果进行比较, 本文仍采用文献[6]的 5 类图像库。这些图像全部是从 eBay 和 Amazon.com 网站下载的商品图像, 分辨率在 280 × 280 左右。表 1 给出了要区分的商品图像的种类和示例图像。

表 1 商品图像库^[6]

种类	示例图像
尖头鞋及圆头鞋	
尼龙搭扣及鞋带	
短袖及长袖	
帆船鞋, 棒球鞋及芭蕾舞鞋	
圆领, V 形领及套衫	

4.2 性能评价

评价分类性能最主要的指标是分类正确率。由于图像测试库中每类图像数目可能有较大的差异, 使用总体分类正确率 (Overall Accuracy, OA) 的计算方法(正确分类图像数占全部图像数的比率)会导致图像数目较少的类别占用的权重较小; 所以本文采用文献[7]提出的类大小调整正确率(Class-Size-Adjusted Accuracy, CSAA)的计算方法, 如式(10)所示:

$$CSAA = \frac{1}{C} \sum_{i=1}^C \frac{P_i}{N_i} \quad (10)$$

其中 C 表示图像类别数, P_i 表示第 i 类正确的分类数, N_i 表示第 i 类图像的总数。如在短袖上衣与长袖上衣的分类中, 假设 100 幅长袖上衣有 90 幅分类

正确, 而 50 幅短袖上衣中有 30 幅分类正确, 则总体分类正确率 $OA = (90 + 30) / (100 + 50) = 80\%$; 而类大小调整分类正确率 $CSAA = 1/2 \times (90/100 + 30/50) = 75\%$ 。

考虑到商品图像在线分类的应用特点, 分类速度也是一个重要的性能指标。本文采用平均分类测试时间(Average Classification Test Time, ACTT)去描述分类速度。由于各标记图像类的特征提取及描述可以以离线方式完成, 计算平均分类测试时间将只考虑在线测试过程, 即查询图像的特征提取、描述及类描述符的形成与匹配。

4.3 实验结果及分析

本文实验全部在配置了 Intel Pentium CPU 2.66 GHz, 1 GB RAM, 运行 Windows XP 操作系统和 MATLAB7.1 软件的计算机上进行。为进一步提高分类速度, 本文使用图片批量编辑工具 Batch Image Resizer 2.88^[14] 将所有的测试图片的分辨率变为 100 × 100, 平均每幅图片的转换时间只有 38 ms。(实际上, 文献[15]已经证明 32 × 32 是能够进行物体分类识别的最低彩色图像分辨率。)实验结果如图 4、图 5 及表 2 所示。图 4(a), 4(b), 4(c)分别给出了在不同标记样本数情况下基于 PHOG, PHOW, PHOG&PHOW 特征的分类正确率, 其中 α 的值是通过五重交叉验证方式获得。表 2 给出了不同种类的最高 CSAA 分类正确率与文献[6]相应结果的对比。图 5 给出了在不同标记样本数情况下的平均分类测试时间。

从以上实验结果中可以看出:

(1)不同分类任务分类正确率存在较大的差异。如长袖与短袖的分类, 在训练样本数为 5 时就已经达到 90%, 增加训练样本数分类正确率逐渐接近 100%; 而尼龙搭扣与鞋带的分类正确率最高不过 70%。

(2)从总体上看, 基于 PHOW 的分类正确率高于基于 PHOG 特征的分类正确率。而 PHOG&

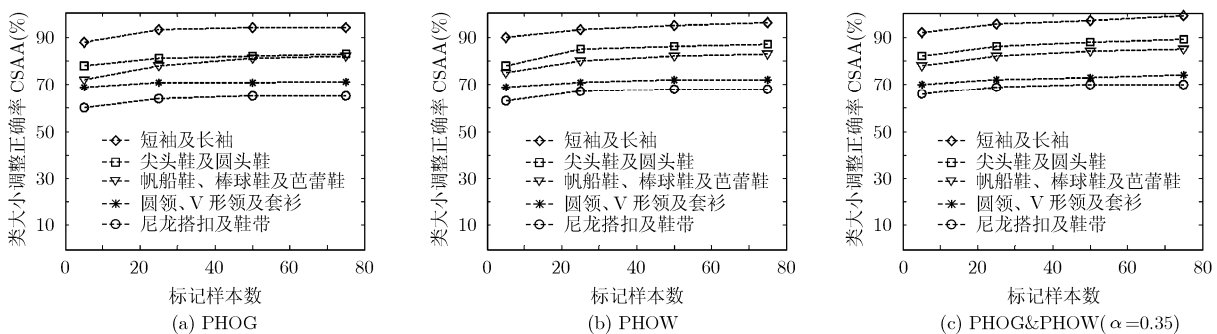


图 4 不同标记样本下的 CSAA 分类正确率

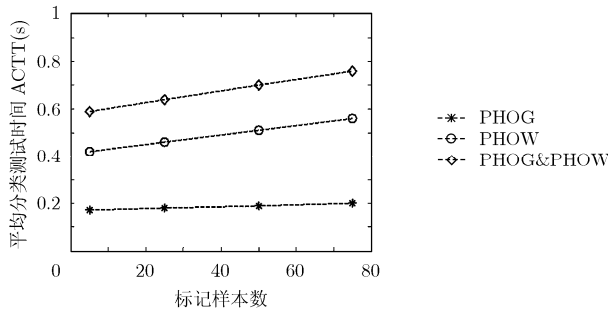


图 5 不同标记样本下的平均分类测试时间

表 2 不同种类的最高CSAA分类正确率

种类	最高CSAA分类正确率(%)			文献[6]
	PHOG	PHOW	PHOG&PHOW	
短袖及长袖	94	96	99	98
尖头鞋及圆头鞋	83	87	89	88
帆船鞋, 棒球鞋及芭蕾舞鞋	82	83	85	83
圆领, V 形领及套衫	71	72	74	67
尼龙搭扣及鞋带	65	68	70	66

PHOW 特征融合的的分类正确率又有了 1~3 个百分点的提高。

(3) 相对于文献[6], 基于 PHOG&PHOW 的分类正确率都有一定程度的提高。基于尤其是圆领、V 型领与套衫的 3 分类和尼龙搭扣与鞋带的 2 分类最高分类正确率分别由 66%, 67% 提高到 70% 和 74%。其原因是(a)与文献[6]的单一图像特征相比, 本文采用了更有区分力的互补性图像特征, (b)文献[6]采用了基于图像到图像距离的 K 近邻分类算法, 本文设计了更合理的基于图像类描述的最近邻分类器, 通过计算图像到类的距离获得更好的推广性能。

(4) 从分类速度上看, 随着标记样本数的提高, 平均分类测试时间有接近线性的较缓慢的增长, 说明测试时间主要取决于查询图像的特征提取过程, 类描述符的提取及匹配时间影响较小。当每类标记数达到 75 时, 基于 PHOG、PHOW、PHOG&PHOW 的平均分类测试时间分别为 0.2 s, 0.56 s 和 0.76 s, 都能够达到实时性的要求, 其中基于 PHOG 的方法在分类速度上有明显的优势。

5 结束语

实现电子商务中的在线商品自动分类是电子商务智能化的迫切要求。本文使用互补的图像特征 PHOG 和 PHOW 及基于类描述的改进最近邻分类

算法实现了 2~3 类商品图像的自动分类, 正确率达到 70%~99%, 并且能达到实时性的要求; 说明基于内容的图像分类技术在电子商务领域有潜在的应用前景和研究价值。本文所做的工作还只是初步的探索, 以后需要进一步解决的问题主要有(1)如何快速实现多类图像的快速自动分类。(2)如何实现结合同一商品多视图图像来提高分类正确率。(3)如何进一步结合电子商务应用的特点, 借鉴人眼视觉感知的研究成果进行图像特征的有效提取、描述和分类器设计。

参考文献

- [1] Li F F and Perona P. A Bayesian hierarchical model for learning natural scene categories[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, Vol. 2: 524-531.
- [2] Grin G, Holub A, and Perona P. The caltech-256. Technical report, Caltech, 2007.
- [3] Lazebnik S, Schmid C, and Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories[C]. Proceedings of the IEEE Computer Society Conference of Computer Vision and Pattern Recognition (CVPR'06), New York, USA, June 17-22, 2006, Vol 2: 2169-2178.
- [4] Nilsback M E and Zisserman A. Automated flower classification over a large number of classes[C]. Proceedings of Computer Vision, Graphics and Image Processing in Indian, 2008: 722-729.
- [5] Agarwal S and Awan A. Learning to detect objects in images via a sparse, part-based representation[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2004, 26(11): 1475-1490.
- [6] Tomasik B, Thiha P, and Turnbull D. Tagging products using image classification[C]. Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, Boston, MA, USA, 2009: 792-793.
- [7] Nowak E, Jurie F, and Triggs B. Sampling strategies for bag-of-features image classification[C]. 9th European Conference on Computer Vision in Computer Vision ECCV 2006, Graz, Austria, May 7-13, 2006: 490-503.
- [8] Boiman O, Shechtman E, and Irani M. In defense of nearest-neighbor based image classification[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08), Anchorage, Alaska, USA, June 23-28, 2008: 1-8.
- [9] Duda R O, Hart P E, and Stock D G. *Pattern Classification (2nd Edition)*[M]. New York, USA, Wiley

- Interscience, 2001: 536-539.
- [10] Bosch A, Zisserman A, and Munoz X. Representing shape with a spatial pyramid kernel[C]. Proceedings of the 6th ACM international conference on Image and video retrieval, Amsterdam, Dutch, 2007: 401-408.
- [11] Dalal N and Triggs B. Histograms of oriented gradients for human detection[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, Vol 1: 886-893.
- [12] Bosch A, Muoz X, and Marti R. Which is the best way to organize/classify images by content? [J]. *Image and Vision Computing*, 2006, 25(6): 778-791.
- [13] Jurie F and Triggs B. Creating efficient codebooks for visual recognition[C]. Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05), Beijing, China, 2005, Vol 1: 604-610.
- [14] Batch Image Resizer, <http://www.jklnsoft.com/>, 2009. 12.
- [15] Torralba A, Fergus R, and Freeman W T. 80 million tiny images: A large data set for nonparametric object and scene recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(11): 1958-1970.
- 贾世杰: 男, 1969 年生, 博士生, 副教授, 研究方向为图像分类与检索.
- 孔祥维: 女, 1963 年生, 教授, 博士生导师, 主要研究方向为多媒体信息安全、统计图像处理和识别等.
- 付海燕: 女, 1981 年生, 博士生, 研究方向为图像处理及应用.