

基于显著图的花卉图像分类算法研究

周 伟,武港山

(南京大学 计算机科学与技术系,江苏 南京 210093)

摘 要:在计算机视觉领域,图像分类已成为最近几年的研究热点,取得了很大的发展。然而目前的研究大多基于开放领域,分类粒度较粗,不能很好地满足花卉图像精细分类的需求。传统的图像分类算法都是基于分割后的图像进行的,较为依赖分割效果的好坏,不太适用于花卉这一类拥有复杂背景的图像。因此结合花卉图像的自身特点,提出了一种新的基于显著图的图像分类算法,将显著图融入到图像特征的提取过程中,从而避免对图像进行分割,增强了算法的适应性和可靠性,随后又对基于SVM的多特征融合方法进行了简单的介绍。通过在花卉图像库进行的实验,证明了算法的有效性。

关键词:显著图;特征提取;特征融合;图像分类

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2011)11-0015-04

Research on Saliency Map Based Flower Image Classification Algorithm

ZHOU Wei, WU Gang-shan

(Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China)

Abstract: Recently, the technology of image classification has been well developed. However, most of the current study is based on open fields, for coarse-grained classification, which cannot meet the demand for automated flower classification task. Traditionally, the feature extraction process is done on the segmented images. But for the flower images with complex background, the image segmentation is not always reliable. So proposed to use the saliency map to assist the feature extraction process, which could improve the robustness of the system. And also discuss the method of combining different features using SVM classifier. Finally, executed experiments on flower image data sets, and comparative results show that the algorithm is effective.

Key words: saliency map; feature extraction; feature combination; image classification

0 引言

在计算机视觉领域,图像分类已成为最近几年的研究热点,取得了很大的发展。随着电子技术和移动技术的发展,人们可以越来越便捷地获取各种图像(比如花卉、宠物等),但是在具体种类未知的情况下,根本无法获取更多的信息。诸如此类新兴需求的增长,提出了一个新的图像分类问题:基于特定领域的精细图像分类,然而传统的图像分类算法,主要应用于开放领域的粗粒度分类,并不能很好地解决这类问题。文中之所以选择花卉图像作为研究对象,主要基于以下两点观察:1、不同种类的花卉相似性很高,拥有相同的颜色、相似的花瓣等等;2、相同种类的花卉差异性很高,处于不同花期、容易形变、花瓣残缺等等。因此,笔

者认为花卉图像分类在精细分类问题中具有很好的代表性,对于解决其他精细分类问题也有一定的指导意义。

目前关于花卉图像分类的研究还比较少^[1],主要的研究成果来自于牛津大学的VGG小组^[2,3],他们把花卉图像分类任务大致分为以下几个步骤:1、图像分割;2、特征提取;3、特征融合;4、图像分类。笔者认为这种先分割后处理的方法存在着一定的局限性,整体的效果过于依赖图像分割的精度,分割效果不好时,丢失的信息比较多,因此算法适应性不强,对应用场合要求比较苛刻。单就花卉图像而言,背景中的泥土、杂草和落叶等对象的存在,光照下花卉、叶子间的相互投影等因素都会造成花卉区域分割的准确率下降。

文中的主要贡献在于将显著图引入到花卉图像分类算法中,避免了常规的图像分割步骤,增强了算法对于复杂图片的适应性。文中主要介绍了基于显著图的颜色直方图和SIFT特征的提取方法,一方面颜色特征和局部特征对于花卉的区分能力比较强,另一方面二

收稿日期:2011-04-25;修回日期:2011-07-28

基金项目:国家自然科学基金项目(60975043,61021062)

作者简介:周 伟(1986-),男,江苏盐城人,硕士研究生,主要研究领域为计算机视觉、图像分类;武港山,博导,CCF会员,主要研究领域为多媒体技术。

者从不同方面对花卉进行描述,可以起到一定的互补效果,通过后续的特征融合可以达到比较好的效果。最后在花卉图像库进行了实验,取得了比较好的效果,验证了方法的有效性。

1 系统架构概述

为了让读者对系统有一个较为整体而清晰的认识,首先简单介绍一下大致的算法流程。这里不再区分训练阶段和测试阶段,因为前期的特征处理阶段都是一致的,只是在利用 SVM 分类器进行训练和分类的时候有所区别。首先,计算出一副图片的显著图,然后运用本方法分别提取出图片的颜色特征和局部特征,之后利用核方法对两种特征进行融合,从而最终完成分类任务。整体算法流程如图 1 所示。

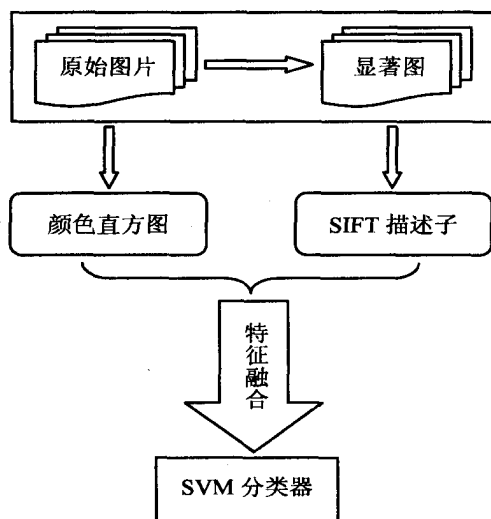


图 1 系统整体算法流程图

2 基于显著图的图像特征提取

2.1 显著图简介

Itti 等人提出了一种经典的视觉注意力模型^[4],该模型的基本思想是采用高斯金字塔对输入图像进行多尺度表示,利用灰度、颜色和方向信息来引导注意,采用中心-边缘的计算策略在每种特征图内部进行竞争,然后通过线性组合得到各个特征通道内的显著图。继 Itti 之后, Wang 等提出了一种两阶段的图像显著性检测方法^[5]; Harel 等提出了一种基于图形的视觉显著性^[6]。

文中采用的是 Achanta 等提出的 MSSS 算法 (Maximum Symmetric Surround)^[7],经过在数据集上的测试, MSSS 算法的效果令人满意,花卉和背景分别得到了很

好的增强和抑制。具体效果如图 2 所示。

以往有学者将显著图引入图像检索系统,利用显著图提取出显著性区域,然后,只对显著性区域提取颜色、纹理、形状等特征^[8]。但是显著性区域的提取和图像分割存在着同样的缺陷,往往对于复杂物体的处理效果不佳。因此,决定直接利用显著图辅助特征提取,进行一些简单的加权操作和过滤操作。

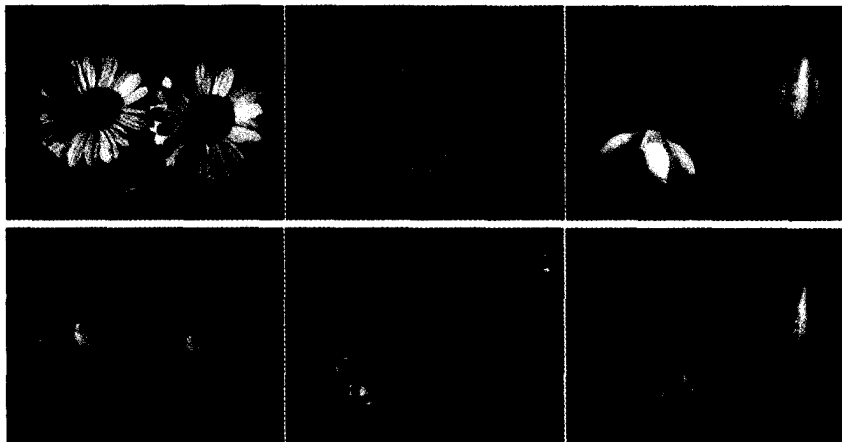


图 2 MSSS 检测效果示例(上:原图 下:显著图)

2.2 颜色特征

颜色是图像的最直接、最重要的视觉特征之一。相对其它视觉特征,颜色特征对图像内容的描述性强(颜色与图像所包含的物体或场景十分相关),含义明确,易于提取。此外,颜色特征对图像本身的尺度、旋转、视角等具有较强的不变性。颜色特征的不足之处在于它表达的是图像的全局属性,而不能很好地代表其中对象的局部特性。因此单靠颜色特征通常不能提供图像的足够信息,而需和其它方法结合使用。

Swain 和 Ballard 最先提出了使用颜色直方图作为颜色特征的表示方法^[9],经过颜色量化之后,统计落在每个颜色区间内的像素数量就可以得到颜色直方图。具体计算公式如下:

$$H(k) = \frac{n_k}{N}, (k = 0, 1, \dots, L-1)$$

其中, L 是量化后的颜色区间数目, n_k 是落在颜色区间 k 中的像素个数, N 是像素总个数。

然而花卉图像具有其自身的特殊性,大部分花卉图片的背景颜色都以绿色为主,比如草地、绿叶、树木等等。因此如果直接提取颜色直方图,这些大面积类似的背景颜色将会削弱花卉本身颜色的区分能力。一般传统做法是通过图像分割来消除背景噪声的影响。为了避免分割效果的不确定性影响,利用显著图计算加权颜色直方图,用像素的显著值来代替个数,这样既没有完全丢弃背景信息,同时降低了背景信息的权重,具有更好的鲁棒性。具体计算公式如下:

$$H(k) = \sum SM(k) / \sum_{i=0}^{L-1} \sum SM(i), (k = 0, 1, \dots, L)$$

- 1)

其中, $SM(k)$ 是落在颜色区间 k 中的像素显著值。

2.3 局部特征

SIFT 是 Lowe 于 1999 年提出的局部特征描述子, 并于 2004 年进行了更深入的发展和完善^[10]。在 Mikolajczyk 对包括 SIFT 算子在内的十种局部描述子所做的不变性对比实验中, SIFT(及其扩展算法)已被证实同类描述子中具有最强的健壮性^[11]。

SIFT 算子是一种局部特征, 它描述的是图像的兴趣区域的灰度梯度分布情况。实际上, SIFT 算法包含了两个部分: 一个尺度不变的兴趣区域检测器和一个基于兴趣区域的灰度梯度分布的特征描述符。描述符实际上是一个 3 维直方图: 其中二维是梯度位置, 被量化为 4×4 的方格 (grid); 另一维是梯度方向, 被量化为 8 个格子 (bin), 因此原始的 SIFT 描述子是一个 $4 \times 4 \times 8 = 128$ 维的向量。限于篇幅, 这里不对 SIFT 算法的技术细节展开讨论。

经过实验发现, 单张花卉图像提取的 SIFT 特征点数目高达 2400, 甚至更多。因为花卉图像的背景一般非常凌乱, 导致相当一部分兴趣点位于杂乱无章的背景区域, 比如草地、树枝等等。一来如此巨大数目的 SIFT 点将会导致运算量急剧上升, 严重影响算法性能; 二来位于背景区域的 SIFT 点, 很有可能冲淡花卉主体特征点的区分能力。与颜色直方图类似, 传统做法也是借助于图像分割技术。然而实验表明, 在分割后图像中提取出的 SIFT 点绝大多数位于分割产生的轮廓线上, 因此, 提取出的 SIFT 质量好坏将严重依赖于分割的效果。

SIFT 特征点提取效果图见图 3。



图 3 SIFT 特征点提取效果图 (左: 原图中; 原图提取 右: 分割提取)

采取的方法是仍然在原图像上进行兴趣区域检测, 对于得到的关键点通过显著图的均值进行过滤筛选, 然后仅对剩余的关键点提取 SIFT 描述子。这种方法可能会丢失部分真实边缘的特征点, 但是避免了提取出虚假的边缘特征点, 可以说有利有弊, 但是笔者认为本来存在的真实的特征点比虚假的特征点具有更强的描述能力。之后采用目前比较流行的 BoW (Bag-of-Words) 模型^[12, 13], 为每幅图像生成基于视觉词汇表的直方图作为特征向量。

3 基于 SVM 和多特征融合的分类算法

支持向量机 (SVM) 已经成为一种倍受关注的分类技术^[14, 15]。这种技术具有坚实的统计学理论基础, 并在许多实际应用 (如手写数字的识别、文本分类等) 中展示了很好的实践效用。

最简单的特征融合方法是直接将颜色特征和局部特征拼接成单一的特征向量来表示每幅图像, 训练一个多元分类器。这种方法的缺陷在于没有区分不同特征的重要性。比如某些花的颜色比较特别, 而某些花的纹理比较特别, 所以应该采取加权组合, 而不是简单的向量拼接。

所采取的方法是, 对不同特征的核函数进行线性组合^[16], 因此最终的核函数具有如下形式:

$$K(i, f) = \sum_{f \in F} \beta_f \exp(-\mu \chi_f^2(x_f(i), x_f(j)))$$

其中, x_f 是基于特征 f 的特征向量, β_f 是特征 f 的权重。 $\chi^2(x, y)$ 是直方图 x 和 y 的 Chi-square 距离, 具体形式为: $\chi^2(x, y) = \sum_k \frac{(x(k) - y(k))^2}{x(k) + y(k)}$ 。根据文献,

$K(x, y) = \exp(-\mu \chi^2(x, y))$ 是 Mercer 核, 因此通过线性组合获得的 $K(i, f)$ 也是 Mercer 核, 参数 μ_f 被设置为所有训练图像的 χ^2 距离的平均值的倒数^[17]。至于权重 β_f , 因为不同种类的花对于两种特征的表现不一致, 所以如果设置为全局统一值并不合适, 采用的方法是针对每一类训练相应的 one-vs-rest 分类器, 学习到针对当前类别的最优 β_f 参数。

4 实验结果及分析

为了便于实验结果的对比分析, 采用 VGG 小组的

17-flower 数据集。该数据集包括 17 种不同的花卉, 每个种类含有 80 幅图片。随机从每个类别抽取 70 幅图片作为训练集, 剩余的 10 幅图片作为测试集。一共设计了 5 组实验, 分别为: 1、原始颜色直

方图 (Color); 2、原图直接提取 SIFT 特征点 (SIFT); 3、基于显著图的加权颜色直方图 (SM Color); 4、基于显著图过滤的 SIFT 特征点 (SM SIFT); 5、基于显著图的加权颜色直方图和基于显著图过滤的 SIFT 特征点相结合 (SM Color & SIFT)。具体每个类别和整体的准确率如表 1 所示。

从上述实验结果可以得出以下几个结论:

(1) 将显著图引入颜色直方图确实对分类的准确率有贡献, 说明对背景噪声起到了一定的抑制作用。

(2) 特征融合带来的准确率提高相当明显, 说明

基于颜色的全局特征和基于 SIFT 的局部特征是能够互补的,尤其在 Colts' Foot 等类别上,效果提升特别明显。

表 1 不同方法的实验结果

	Color	SIFT	SM Color	SM SIFT	SM Color & SIFT
Daffodil	33.33%	77.78%	33.33%	66.67%	88.89%
Snowdrop	44.44%	66.67%	44.44%	55.56%	66.67%
Lily Valley	66.67%	55.56%	66.67%	77.78%	77.78%
Bluebell	77.78%	100%	88.89%	88.89%	88.89%
Crocus	55.56%	66.67%	44.44%	77.78%	77.78%
Iris	77.78%	100%	66.67%	77.78%	88.89%
Tigerlily	44.44%	100%	88.89%	22.22%	88.89%
Wild Tulip	44.44%	55.56%	55.56%	100%	100%
Fritillary	77.78%	77.78%	66.67%	66.67%	100%
Sunflower	77.78%	100%	100%	55.56%	100%
Daisy	66.67%	100%	77.78%	44.44%	100%
Colts' Foot	44.44%	88.89%	44.44%	33.33%	77.78%
Dandelion	22.22%	100%	33.33%	88.89%	77.78%
Cowslip	88.89%	77.78%	55.56%	77.78%	88.89%
Buttercup	0	88.89%	33.33%	66.67%	77.78%
Windflower	77.78%	66.67%	88.89%	77.78%	88.89%
Pansy	44.44%	100%	66.67%	88.89%	88.89%
Total	55.56%	83.67%	62.09%	68.63%	86.93%

(3)出乎意料的是单独的 SIFT 特征取得了非常高的准确率,经过分析认为,虽然大部分花卉图像的背景颜色比较相似,但是它们的一些局部特征还是有差异的,另外不同种类的花的叶瓣也存在着差别,但是这些背景、叶片等在显著图中基本都处于被削弱的区域,因此导致了部分 SIFT 关键点的丢失,导致基于显著图的 SIFT 特征性能反而下降了。如图 4 所示,Bluebell 一般生长于草丛之中,而 Tigerlily 的叶子很有特色,因此这些类似的类别在原始 SIFT 方法中几乎都达到 100% 的准确率,在基于显著图的 SIFT 中准确率反而有所下降。



图 4 示例图片(上:Bluebell 下:Tigerlily)

(4)文献[4]中,基于前景区域 SIFT 特征和 HSV 颜色特征的准确率为 85.8%,基于轮廓 SIFT 特征和 HSV 颜色特征的准确率为 81.4%,可见结果是具有竞争力的,甚至在某种程度上超越了先分割后处理的效果,说明对于花卉分类问题,基于显著图的方法是具有

一定优势的。

5 结束语

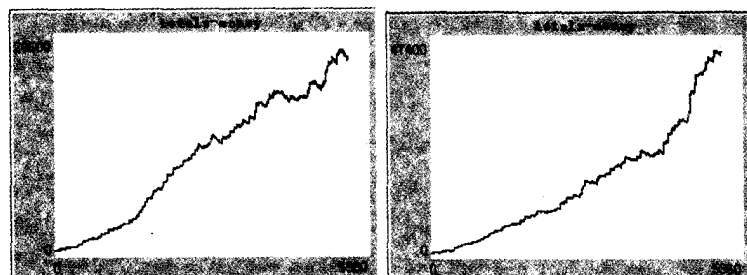
文中将显著图引入花卉图像分类体系,从而避免传统的先分割后处理的流程,并给出了基于显著图提取颜色特征和局部特征的方法,然后简单介绍了特征融合的多核学习方法,最后通过实验证明了方法的有效性。下一步,准备基于显著图提取花卉图像更多种类的特征,比如形状特征、纹理特征等等,寻找描述花卉图像的最优特征组合。

参考文献:

- [1] 柯 道,陈小芬,李绍滋. 基于多特征融合的花卉图像检索[J]. 计算机科学,2010,37(11):282-286.
- [2] Nilsback M, Zisserman A. A Visual Vocabulary for Flower Classification[C]//IEEE Conference on Computer Vision and Pattern Recognition. [s. l.]:[s. n.],2006.
- [3] Nilsback M, Zisserman A. Automated Flower Classification over a Large Number of Classes[C]//Indian Conference on Computer Vision, Graphics & Image Processing. [s. l.]:[s. n.],2008.
- [4] Itti L, Koch C. Computational modelling of visual attention[J]. Nature reviews neuroscience,2001,2(3):194-203.
- [5] Hou X, Zhang L. Saliency detection: A spectral residual approach[C]//IEEE Conference on Computer Vision and Pattern Recognition. [s. l.]:[s. n.],2007.
- [6] Harel J, Koch C, Perona P. Graph-based visual saliency[J]. Advances in neural information processing systems,2007,19:545-552.
- [7] Achanta R, Susstrunk S. Saliency detection using maximum symmetric surround[C]//IEEE International Conference on Image Processing. [s. l.]:[s. n.],2010.
- [8] 宋雁澜,张 瑞,支 琤,等. 一种基于视觉注意模型的图像分类方法[J]. 中国图象图形学报,2008,13(10):1886-1889.
- [9] Swain M J, Ballard D H. Color indexing[J]. International journal of computer vision,1991,7(1):11-32.
- [10] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision,2004,60(2):91-110.
- [11] Mikolajczyk K, Schmid C. A performance e-evaluation of local descriptors[J]. IEEE transactions on pattern analysis and machine intelligence,2005,27(10):1615-1630.
- [12] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories[C]//IEEE Conference on Computer Vision and Pattern Re-

(下转第 22 页)

察系统的 totals-money 变化趋势如图 4 所示。



(a) 无分工

(b) 有分工

图 4 有无分工情况下(相对的)totals-money 变化趋势

从图 4 可以看出,系统在有分工和无分工的情况下,集群效益差距很大,尤其在系统分工后,系统的 totals-money 上升更快。这也从另一个层面说明分工对于集群的重要性。分工也是衡量集群发展水平的重要指标。

对仿真结果分析可知:

1) 产业集群演化网络是无标度网络,其度分布服从幂律分布;

2) 恶性竞争和政府对集群的发展影响很大,尤其在萌芽阶段;

3) 以合作为基础的产业分工的形成,是集群演化的外在表现,对集群效益影响很大。

3 结束语

针对从传统经济学角度研究产业集群方法的不足以及产业集群的复杂性,考虑到复杂性系统理论研究此问题的优越性,文中以 CAS 理论为视角,把产业集群的概念拓展并结合个体行为演化模型和 ECHO 模型提出产业集群系统构架,并在此基础上提出产业集群演化机制模型。

该模型把产业集群演化的生命周期特性和复杂系统的演化识别过程结合起来,以复杂系统的演化机制为依据,提出产业集群诞生于混沌经济环境,经过系统对环境的自适应学习、自身的分岔与突变、系统整体的涌现,达到系统暂态稳定的循环过程。其中自适应机

制主要体现在产业集群系统主体之间、主体与环境的交互,分岔与突变机制主要体现在产业集群内主体的特定性知识增加为核心的规则集更新,其外在表现为以物质流、信息流、能量流的更新为主的系统创新。

参考文献:

- [1] 傅荣,袁丽,张喜征,等.产业集群参与者交互偏好与知识网络演化:模型与仿真[J].中国管理科学,2006,14(4):128-133.
- [2] 刘友金,郭新.集群式创新形成与演化机理研究[J].中国软科学,2003,23(2):91-95.
- [3] 牟绍波,王成章.产业集群持续成长的自组织机制研究[J].科技进步与对策,2007,24(7):73-75.
- [4] 许登峰.基于自组织理论的旅游产业集群发展研究[J].广西民族大学学报,2010,32(1):122-126.
- [5] 欧阳莹之.复杂系统理论基础[M].上海:上海科技教育出版社,2002.
- [6] 付韬,张永安.我国核型集群创新网络的多 agent 仿真研究[J].计算机应用研究,2010,27(8):3002-3005.
- [7] Lomi A, Larsen E R, Freeman J H. Things Change: Dynamic Resource Constraints and System-Dependent Selection in the Evolution of Organizational Populations[J]. Management Science, 2005, 51(6): 882-903.
- [8] Zhai E, Shi Yongjiang, Gregory M. The Growth and Capability Development of Electronics Manufacturing Service (EMS) Companies[J]. International Journal of Production Economics, 2007, 107: 1-19.
- [9] 黄春萍.基于 CAS 理论的企业系统演化机制研究[D].天津:河北工业大学,2007.
- [10] March J G. Rationality, Foolishness, and Adaptive Intelligence[J]. Strategic Management Journal, 2006, 27: 201-214.
- [11] 陈莞.产业集群结构分析及其演化的元胞自动机模拟[D].福州:福州大学,2005.
- [12] 武海鹰,王绪安.分布式人工智能与多智能体系统研究[J].微机发展(现更名:计算机技术与发展),2004,14(3):80-82.

(上接第 18 页)

cognition. [s. l.]: [s. n.], 2006.

- [13] 王宇石,高文.用基于视觉单词上下文的核函数对图像分类[J].中国图象图形学报,2010,15(4):607-616.
- [14] 付岩,王耀威,王伟强,等. SVM 用于基于内容的自然图像分类和检索[J].计算机学报,2003,26(10):1261-1265.
- [15] 陈丽,陈静.基于支持向量机和 k-近邻分类器的多特征融合方法[J].计算机应用,2009,29(3):833-835.
- [16] Bach F R, Lanckriet G R G, Jordan M I. Multiple kernel

learning, conic duality, and the SMO algorithm[C]// International Conference on Machine Learning. [s. l.]: [s. n.], 2004.

- [17] 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: 213-238.