

多类 SVM 在图像艺术属性分类中的应用研究

陈小娥, 陈昭炯

(福州大学 数学与计算机科学学院, 福建 福州 350108)

摘 要:针对当前图像分类研究中,依据图像艺术风格属性进行分类的算法尚不多见的情况,实现了一种基于艺术属性的图像自动分类系统,其中主要涉及摄影作品、国画、水彩画、素描、油画等几种典型艺术风格的图像。系统采用支持向量机(SVM)作为分类器,运用分等级的分类方法,提出了一种针对艺术属性图像分类的特定 SVM 二叉树多类分类算法;而后通过对各类图像艺术风格特征的分析,分别提取了有代表性的、区分度好且易于计算的特征;最后针对各级分类特性和分类器总体特性进行了实验分析,实验结果表明,系统具有良好的分类性能。

关键词: 支持向量机;二叉树多类分类算法;图像艺术属性

中图分类号: TP391 **文献标识码:** A **文章编号:** 1673-4785 (2009) 02-0157-06

An application of multi-class SVM in the classification of artistic attributes of images

CHEN Xiao-e, CHEN Zhao-jiong

(College of Mathematics and Computer Science, Fuzhou University, Fujian 350108, China)

Abstract: In image classification, few current classification algorithms classify images by their artistic attributes. An automatic image classification system based on artistic attributes was developed for classifying images in typical artistic styles such as photographs, Chinese paintings, watercolors, sketches, oil paintings, and so on. The system employed a support vector machine (SVM) as a classifier. Using a classification method at various levels, an SVM binary tree multi-class classification algorithm for image classification with respect to different artistic attributes was proposed. By analyzing the images with respect to the different artistic styles, some easily computed representative characteristics with good discriminability were extracted at each classification level. Experiments on a variety of characteristics at various levels and the total characteristics of classifiers were designed to evaluate the proposed classifier. Experimental results showed that the system has good classification performance.

Keywords: SVM; multi-class classification algorithm based on binary tree; image artistic attributes

随着互联网技术的飞速发展,网络图像资源越来越丰富,如何快速、有效的进行图像检索或分类就成为非常有挑战性的课题,图像的自动分类是其中一项重要的研究内容。

图像自动分类具有广阔的应用前景。在一些专业图片公司里,很多员工每日都要对各种图片进行归类、标注,撰写描述信息,工作量十分庞大。如果计算机能够像人类一样看懂一幅图片,就可节省大量的人力、物力、财力,而分类正是促进智能识别图像的一种有效方法。其次,图像的自动分类对图像管

理、Web 图像搜索都有积极的意义。如果在搜索之前先进行分类的话,就可以缩小图像的搜索范围,加速图像搜索的速度,提高搜索效率,不失为缩小图像的低级特征和高级概念之间语义间隔的一个可行方法。同时,分类也符合人们对图像搜索、多媒体内容分析与理解、数字图书馆等应用领域的多种需求,诸如信息资料共享中文档图像分类^[1]、情感心理搜索^[2]等。

之前有关图像自动分类的研究主要是根据若干图像的属性来分类。Athitsos 和 Swain 等人^[3]提出了将图片分为普通摄影图片和由计算机生成图片的分类方法,文中主要先分析这 2 类图片的特点,然后根据这些特点提取相应的分类特征(如颜色直方图、主

收稿日期: 2008-05-12

基金项目:福建省自然科学基金资助项目(A0710006)。

通信作者:陈昭炯。E-mail: chenzy@fzu.edu.cn

导颜色数目、图片大小等特征),并采用多决策树分类方法进行分类;Szummer^[4]等人通过计算图像 Ohta 空间的颜色直方图和分辨率自回归纹理模型等特征将图像分为室内画和室外画,分类过程主要采用 K 近邻方法形成子分类器,再运用 stacking 思想对各分类器的输出进行处理;Vailaya^[5]等人将图像分为城市画和风景画,提出了颜色聚合向量和边缘方向直方图等区分度较好的分类特征,采用 K-NN 分类器进行分类;还有一些学者提出了其他的分类属性,诸如按图像主题进行分类^[6]、对自然图像的场景属性进行分类^[7],分成落日、海洋和原野等。

而当前基于不同艺术风格属性的图像自动分类算法的研究相对比较少。这里图像的艺术风格指的是图像的不同创作手法。2005 年, Cutzu 等^[8]提出了将图片分为摄影作品和绘画作品的分类方法,由不同特征形成多个分类器,再将这些分类器进行简单组合得到最终分类结果。此种分法相比上文提到的图像分类算法难度更大,因为该问题具有相对的普遍性,即不对图像的内容做任何限制。比如不像之前的分类那样根据图像内容将图片具体分为城市画和风景画、室内画和室外画。而本文提出了一个将摄影作品、国画、水彩画、素描和油画等不同艺术风格图片自动分类的算法,由于新增了多种艺术风格的图片,因此具有更大的挑战性。本文算法中,分析了不同风格属性图像的特点并提取了相应的特征。通过实验为每一级提取了有代表性的、区分度好的图像特征,其中选择了一些常用的、分类效果比较好的图像特征,如颜色数目和傅里叶频谱等。同时借鉴文献[8]并提出了一些适合不同艺术风格属性图像分类的特征,如高低灰度对比度、饱和度比值和平滑性等。对于分类器的设计,选择了支持向量机作为学习机器,并运用当前比较新的基于二叉树的 SVM 分类算法进行分类。该算法所需的分类器少,分类速度快、推广性能好。

图像的低级特征和高级概念之间存在着语义间隔。不少图像检索系统,由于对图像的领域没有限制,检索效果常常不尽人意。本文提出的基于艺术风格属性的图像分类方法,在某种意义下缩小了低级特征和高级概念之间存在的语义间隔。即通过在宽图像领域(各种不同风格属性)中先进行了粗分类,之后如果再在窄图像领域(同种风格属性)中进行细分类,这可以降低分类的难度,提高分类精度。

1 基于艺术风格属性的分类器设计

1.1 支持向量机

支持向量机(SVM)是 Vapnik 于 1995 年首先提出来的,是近年来机器学习研究的一项重大成果。其基本思想是,寻找一个最优超平面使它的分类间隙最大。

支持向量机方法建立在统计学习理论的 VC 维理论和结构风险最小原理基础上,根据有限的样本信息在模型的复杂性(即对特定训练样本的学习精度)和学习能力(即无错误地识别任意样本的能力)之间寻求最佳折衷,以期获得最好的推广能力。与传统的模式识别方法相比,支持向量机方法在小样本、高维和非线性数据空间下,在学习过程中可以充分利用多种特征提供的信息,具有较好的推广能力。正由于 SVM 具有以上优点,并且在许多实际问题中 SVM 相对于其他分类方法在某些方面也获得了较好的结果,因此,本文的分类器主要采用 SVM。

1.2 多类 SVM

传统的 SVM 是基于两类问题的,而本文所研究的基于艺术风格属性的图像分类是多类分类问题。虽然多类 SVM 的算法已经提出并得到应用,但是不同的多类 SVM 算法具有不同的复杂度和分类性能。

当前多类 SVM 算法大致可分为两大类:一是先以某种方式构造一系列的两类分类器,然后再把这些分类器以某种特定的方式组合在一起来实现多类分类,目前该分类方法比较典型的算法有 one-against-rest^[9]、one-against-one^[9]、基于二叉树的多类 SVM 法^[10];二是将多个分类面的参数求解合并到一个最优化问题中,通过求解该最优化问题“一次性”实现多类分类^[11]。该方法在分类时表面上看起来比较方便,但在最优化问题求解过程中的变量远多于前一类方法,训练速度也相对较慢,同时分类精度上也不占优势^[12]。

1.3 基于二叉树的多类 SVM

于是本文采用上文介绍的第一类方法中的基于二叉树的多类 SVM 算法。SVM 二叉树分类法是把 SVM 和二叉决策树结合起来的一种相对比较新的多类别分类方法,其本质是将原有的多类问题分解成一系列的两类分类问题。

基于二叉树的多类 SVM,对于 K 类分类问题只需构造 $K-1$ 个分类器,分类时只需少量的分类器。

构造 SVM 二叉树可以有多种不同的方案,典型的有以下 2 种特殊结构(图 1):一种是“偏态树”结构(图 1(a)),即每个结点分类器每次只把某一类从类集合中划分出来,直到所有节点都只包含一个单独的类别;另一种是“正态树”结构(图 1(b)),即每个结点分类器把类集合均划分成 2 个子集合,直到所有的类都划分开来。显然,正态树情况下的训练时间比偏态树情况下要短,而其他情况下的训练时间则介于两者之间。

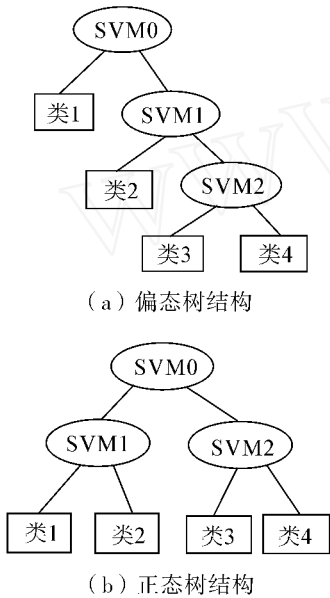


图 1 2 种典型的 SVM 二叉树分类结构

Fig 1 Two typical binary tree based SVM classification structure

1. 4 艺术风格图像分类器设计

由于二叉树本身具有层次结构,上层节点的分类错误将影响整个 SVM 分类器的性能,且在越靠近根节点的地方发生分类错误,对分类器整体性能的影响就越大。为了使“错误累积效应”尽量降低,应该使出现错分率越大的尽量远离根节点。

于是,遵循易划分的类先划分,不易划分的类晚划分的原则,提出了一种特定的针对不同艺术风格属性图像分类的 SVM 多类二叉树分类算法。分类器的设计思路:先设计一些图像的分类特征,计算不同艺术风格图像的特征值,根据不同特征值对几种风格图像的区分程度,进行分级分类;将能把这几类图像区分成两大类且区分程度最好的特征,作为第 1 级的分类特征,可以区分的类别作为第 1 级进行区分的类别;依次类推,经过实验得到算法的二叉树结构(图 2)。各层的分类器及其区分出的类别如图 2 所示,且每一层均采用特定的分类特征以及相应的

训练样本和测试集进行试验。

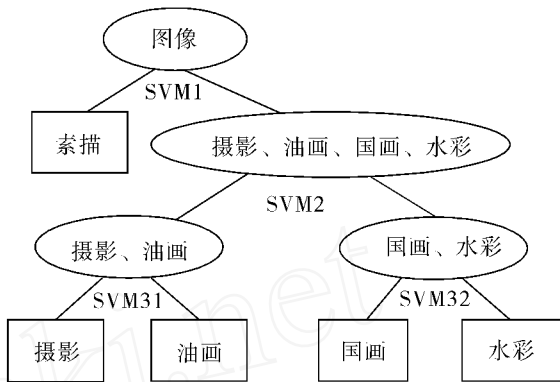


Fig 2 Binary tree based SVM multi-class classification for different artistic style images

2 分类特征选择

对图像进行分类,分类的准确率很大程度上取决于从图像中提取的特征,于是从不同方面、不同角度分析并提取了不同艺术风格图像的特征。

2. 1 颜色数目

图像的颜色是图像的重要视觉特征,图像的颜色数反映了图像色彩的复杂程度。素描是美术中最单纯的造型形式。广义上的素描,涵指一切单色的绘画。素描的颜色数相对于其他绘画作品来说比较少。于是可以采用这个特征将素描与其他风格的图片区分出来。

由于 RGB 颜色空间不直接与色调、饱和度和亮度等人的主观感觉相对应,于是先将图像由 RGB 空间转换到 HSV 空间:

$$T: RGB \rightarrow HSV. \tag{1}$$

接着对 HSV 颜色空间进行量化,综合考虑分类的准确度与算法的运行时间等因素,把 HSV 颜色空间量化成 256 个颜色,即 H 分量量化成 16 个值, S 和 V 分量量化成 4 个值,具体表示为

$$Q: HSV \rightarrow C. \tag{2}$$

式中: $C = \{c_i \mid i = 0, 1, 2, \dots, 255\}$. c_i 表示第 i 种量化颜色,由此统计每种艺术风格图像的颜色数目。实验证明,该特征可以很好地区分出素描。

2. 2 高低灰度对比度

通过实验观察,得到不同风格的图像具有的高低灰度对比度不一样。高低灰度对比度,是指一幅图片在一定范围内的低灰度值的像素个数与高灰度值

的像素个数的比值.具体计算步骤如下:将图像转换为灰度图像:

$$T: \text{rgb_image} \rightarrow \text{gray_image} \quad (3)$$

计算高低灰度对比度:

$$\text{grayratio} = \frac{\#\text{pixels gray_image} < \tau_1}{\#\text{pixels gray_image} > \tau_2} \quad (4)$$

式中: τ_1 、 τ_2 分别为低高灰度阈值.在实验中,分别取 60 和 200,当高灰度值像素个数为 0 时,将其自动设置为 1.而 $\#\text{pixels gray_image} < \tau_1$,表示满足在图像 gray_image 中灰度值小于 τ_1 的像素个数.以下公式的意义以此类推.

通过计算得到,通常情况下摄影作品和油画的 grayratio 值均大于 1,而国画和水彩画的 grayratio 值均小于 1.

2.3 傅里叶变换的频谱分布

频域变换分析有着空域分析所不具备的优点,能够通过频域分量与纹理粗细简单的对应关系,直观而清晰地对纹理进行分析.如果图像中的目标形状或排列呈现某种方向性,那么具有较高值的频谱也呈现出与图像目标方向正交的方向性分布.

而经过变换后,图像的能量主要集中在低频区域,高频区的幅值比较小.对大多数无明显颗粒噪声的图像来讲,低频区集中了 85% 的能量.于是通过计算图像的傅里叶变换,对变换后的幅值图像保留其低频分量,舍弃高频分量.利用其低频分量作为图像的纹理特征,以此区分各种不同风格的图像.实验得到,摄影作品和油画的低频分量的中心值比较接近,而国画和水彩画的低频分量的中心值比较接近.

2.4 饱和度比值

由观察得到,油画包含的高饱和度的像素个数所占的比例大于其他风格的图片,此不同点对于油画和摄影作品来说区别更为突出.饱和度比值^[8]的计算方法如下.

首先,计算图像各个像素的饱和度值.

$$\max = \max(R, G, B), \quad (5)$$

$$\min = \min(R, G, B), \quad (6)$$

$$S = \begin{cases} 0, & \max = 0; \\ (\max - \min) / \max, & \text{其他.} \end{cases} \quad (7)$$

接着,定义饱和度直方图:

$$h(s_k) = n_k \quad (8)$$

式中: s_k 是饱和度 S 在区间 $[0, 1]$ 内的第 k 级饱和度, n_k 是饱和度级为 s_k 的图像中的像素数.并将其进

行归一化:

$$p(s_k) = h(s_k) / n \quad (9)$$

式中: n 为图像中的像素总数,即可知 $p(s_k)$ 表示饱和度和级 s_k 出现的频数.

最后,计算饱和度比值(实验中, k 取 5,即 $k_{\max} = 5$, $k_{\min} = 1$,当然可根据实际情况调整参数)

$$\text{Saturation_ratio} = p(s_{k_{\max}}) / p(s_{k_{\min}}). \quad (10)$$

2.5 平滑性

水彩画的画面有清沏透明之感受.用水调色,发挥水分的作用,滋润流畅、韵味无尽.而国画是以线存形的,通过线勾出轮廓、质感、体积来运用无穷变化的笔锋,发挥笔趣和墨彩,产生奇妙的效果.针对这些特点,可以采用图像的平滑性^[8]对这 2 类图像进行区分.其具体计算过程如下:首先,对图像 R 、 G 、 B 3 个通道的值进行归一化.

$$R_m = R / I, G_n = G / I, B_n = B / I \quad (11)$$

式中:

$$I = 0.3R + 0.6G + 0.1B. \quad (12)$$

接着,对归一化后的各个通道图像进行 3×3 拉普拉斯掩模卷积.用式 (13) 分别求得三通道的卷积图像 $\nabla^2 fR$ 、 $\nabla^2 fG$ 、 $\nabla^2 fB$,

$$\nabla^2 f = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)] \quad (13)$$

最后,计算图像的平滑性.

$$\text{smoothness} = [(\#\text{pixels} / \nabla^2 fR / < \epsilon) + (\#\text{pixels} / \nabla^2 fG / < \epsilon) + (\#\text{pixels} / \nabla^2 fB / < \epsilon)] / 3 \quad (14)$$

式中: ϵ 为接近 0 的参数,实验中取 0.001.

3 实验与结果分析

在本次实验中,所用的图像数据库由 1 960 幅图像组成,其中包括摄影图片 450 幅,国画 310 幅,水彩画 390 幅,素描 410 幅,油画 400 幅.这些图片均从网上相应的图片库中下载.图片大小没有限制,实验中对各个特征均作了相应的归一化处理.但是实验中的图片必须是不包含任何裱框的图片,同时不经过任何特殊的艺术化处理,且图片中不包含大量的文字.

在实验中,各级支持向量机均采用 RBF 核函数 $K(x, y) = \exp(-\|x - y\|^2)$,其中,核参数取 4,惩罚因子取 10.对于每一级的分类,将图像数据库中特定的图片集分为训练集和测试集;对各级训

练集分别进行训练,形成 4 个子分类器,并分析各个子分类器独立测试集上的准确率;再由各个子分类器组成特定的基于不同艺术风格属性图像的基于二叉树的 SVM 多类分类器,从而得到一个总的分类模型.最后,通过对总测试集的测试,分析整个分类模型的性能及推广能力.

3.1 各级图像分类

第 1 级分类,构造分类器 SVM1,利用颜色数目先将素描从 5 种风格的图像中区分出来.对于训练集,在图像数据库中任取素描 300 幅,类标号为 1;同时,为了尽量平衡两类样本数目,其余 4 种风格图片每种只任取其中 100 幅,类标号均为 2,进行训练.而剩下的素描图片和其余 4 种风格的部分图片作为该级的测试集,同样素描类标号为 1,其余为 2

同理,第 2 级分类,构造分类器 SVM2,利用高低灰度对比度和傅立叶频谱把除素描外的图像分为摄影、油画一类和国画、水彩画一类.第 3 级分类,分类器 SVM31 利用饱和度比值对摄影图片和油画进行分类,分类器 SVM32 利用平滑性对国画和水彩进行分类.各级具体的训练集和测试集图片数量及其分类结果等见表 1.

表 1 各级分类实验结果

分类器	区分类别	特征	训练集	测试集	查准率 / %
SVM1	素描、其余 4 种风格	颜色数目	700	510	97.53
	摄影和油画、国画和水彩画	灰度对比度 傅里叶频谱	1 000	550	83.30
SVM2		两特征结合	1 000	550	91.67
SVM31	摄影、油画	饱和度比值	500	350	91.46
SVM32	国画、水彩	平滑性	450	250	72.22

3.2 整个分类器性能分析

把图像库中除了作为训练样本外的其他图像全部作为测试样本,来测试整个分类器的性能,得到分类器的整体分类准确率为 85.56%.由于第 3 级分类中水彩画和国画的分类的准确率相对较低,因而多少影响了整个分类器的性能,但又由于其处于整个分类器的最底层,所以对分类器的性能影响不是特别大.图 3 是得到正确划分的部分图像,包含了各种风格的图片.其中,(a)组为正确划分的素描图片,(b)组为正确划分的国画,(c)组为正确划分的

摄影图片,(d)组为正确划分的油画,(e)组为正确划分的水彩画.



图 3 分类器正确划分的部分图像

Fig 3 Some correctly classified pictures

根据提取的特征,可以推断出哪些图片比较容易被错分.如颜色比较单一的国画、黑白摄影图片、简单色彩的水彩画都可能被错分为素描;而油画通常情况下颜色都比较丰富,因此错分为素描的几率就几乎为零.一些颜色比较鲜艳的摄影图片,错分为油画的几率会比较高一些.图 4 列出了一些试验中被错分的图片,其中,(a)、(b)是被错分为素描的图片,(c)、(d)是被错分为国画的图片,(e)、(f)是被错分为摄影的图片,(g)、(h)是被错分为油画的图片,(i)、(j)是被错分为水彩画的图片,而每幅图片下面的小标题代表它们实际的风格.



图 4 部分被错分的图像

Fig 4 Some wrongly classified pictures

4 结束语

本文提出了一个特定的针对不同艺术风格图像的分类算法并实现了相应的分类系统,在相对独立测试集上的实验以及整个分类系统的测试实验都得到了较好的分类结果.当然,系统还存在着一些可以改进的地方,如进行训练试验的图像数量稍微偏少了一点,以后可以多增加一些;在实验的基础上,为

每一级分类选择了分类效果相对比较好的相同 SVM 参数,今后可以结合其他算法为每一级选择最合适的不同的 SVM 参数;只提取了一二个图像特征作为每一级 SVM 输入的分类特征,在今后的工作中可以尝试更多的特征以进一步提高每一级的分类准确度,从而提高整个系统的分类性能;另外,还可以考虑加入更多风格属性的图片,如再加上遥感图片、文档风格的图片等,进一步完善整个分类系统.

参考文献:

- [1] 毛大伟,张宏,童勤业. 利用不同尺度下复杂性的差异区分文字和照片[J]. 计算机辅助设计与图形学学报, 2005, 17(8): 1834-1838.
MAO Dawei, ZHANG Hong, TONG Qinye. Image classification using Lempel-ziv complexity difference at different scale[J]. Journal of Computer-aided Design & Computer Graphics, 2005, 17(8): 1834-1838.
- [2] 汪上飞,王熙法. 图像情感检索研究的进展与展望[J]. 电路与系统学报, 2005, 10(4): 102-110.
WANG Shangfei, WANG Xufa. Development and prospect on emotion image retrieval[J]. Journal of Circuits and Systems, 2005, 10(4): 102-110.
- [3] ATHITSOS V, SWAN M J, FRANKEL C. Distinguishing photographs and graphics on the world wide web[C]//Proc IEEE Workshop on Content-Based Access of Image and Video Libraries Puerto Rico, 1997: 10-17.
- [4] SZUMMER M, PICARD R W. Indoor-outdoor image classification[C]//IEEE International Workshop on Content-based Access of Image and Video Databases Bombay, India, 1998: 42-51.
- [5] VALAYA A, JAN A, ZHANG H. On image classification: city, vs landscape[J]. Pattern Recognition, 1998, 31(12): 1921-1935.
- [6] 孟祥增,钟义信. WWW 中图像的主题分类研究[J]. 情报杂志, 2004, 23(10): 47-57.
MENG Xiangzeng, ZHONG Yixin. Theme-based classification of images in WWW[J]. Journal of Information, 2004, 23(10): 47-57.
- [7] 付岩,王耀威,王伟强,高文. 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.
FU Yan, WANG Yaowei, WANG Weiqiang, GAO Wen. Content-based natural image classification and retrieval using SVM[J]. Chinese Journal of Computers, 2003, 26(10): 1261-1265.
- [8] CUTZU F, HAMMOUD R, LEYKN A. Distinguishing paintings from photographs[J]. Computer Vision and Image Understanding, 2005(100): 249-273.
- [9] 边肇祺,张学工. 模式识别[M]. 2版. 北京:清华大学出版社, 2004: 284-303.
- [10] SANG H O, SUNGMOON C, SOO Y L. Support vector machines with binary tree architecture for multi-class classification[J]. Neural Information Processing Letters and Reviews, 2004, 2(3): 47-51.
- [11] WESTON J, WATKINS C. Multi-class support vector machines[R]. Technical Report CSD-TR-98-04, Royal Holloway, Department of Computer Science, University of London, 1998.
- [12] HSU C W, L N C J. A comparison of methods for multi-class support vector machines[J]. IEEE Transactions on Neural Networks, 2002(13): 415-425.

作者简介:



陈小娥,女,1984年生,硕士研究生,主要研究方向为图形图像处理技术.



陈昭炯,女,1964年出生,教授,主要研究方向为图形图像处理与智能算法设计等.主持及参与多项国家和省级基金项目,已发表学术论文 50 余篇.