

SVM 用于基于块划分特征提取的图像分类

邢慧强, 王国宇

(中国海洋大学信息学院电子工程系, 山东 青岛 266071)

摘要: 在基于内容图像检索中, 图像的底层视觉特征和高层语义概念之间存在着较大的语义间隔。使用机器学习方法学习图像特征, 自动建立图像类的模型成为一种有效的方法。本文提出了一种用支持向量机(SVM)实现自然图像自动语义归类的方法, 基于块划分聚类得到特征向量作为 SVM 训练样本, 实现语义分类器。由于参与聚类的是某类图像所有块的特征, 提取的特征更能反映某一类图像特征。实验证明这种方法是有效的。

关键词: 图像划分; 特征矢量聚类; 支持向量机(SVM); 图像分类; 图像检索

中图分类号: TP391 **文献标识码:** A

Partition-Based Image Classification Using SVM

XING Hui-qiang, WANG Guo-yu

(Department of Electronics and Engineering, Ocean University of China, Qingdao, Shandong, China 266071)

Abstract: In the approach of content-based image retrieval, there exists a semantic gap between low-level visual features and high-level concepts. Using machine learning method to learn image features and to automatically construct models for image classes is a promising way. In this paper, support vector machines are trained for nature image classification. We propose an image representation method based on image partition and region clustering. The mapping which maps an image to its representation does not really depend on that image alone but on the entire collection of images from which the region groups have been built. The experimental results are promising.

Key words: image partition; feature vector cluster; support vector machines(SVM); image classification; image retrieval

1 引言

现代技术运用各种手段大量的采集和产生各种类型的信息数据, 其中, 图像信息作为多媒体信息中量最大最主要的一种正在飞速的增长, 如数字照片、医学图像、卫星图像等。因此如何有效地、快速地从大规模的图像数据库中检索出需要的图像是一个急需解决的重要问题, 基于内容的图像检索(CBIR) [1] 和基于语义的图像检索[2]技术正是解决这一问题的有效途径。基于内容的图像检索由图像自身的视觉内容, 如颜色、纹理, 形状等进行索引, 能充分表达图像固有特征。然而, 基于内容的图像检索的最大困难就是: 系统提取出的图像低层内容特征与用户检索时具有的高层语义之间无法对应起来, 也无法更好的满足普通用户的查询习惯。也就是说CBIR无法真正支持基于语义的图像检索, 其关键的局限性在于只涉及了图像表层特征, 而没有利用知识进一步挖掘图像的语义内容。

本文提出一种基于机器学习的图像语义自动分类方法, 通过对图像数据库中不同语义类别训练样本的学习, 使系统可以将新的图像自动归为恰当的语义类别中。在检索时可以实现更符合用户要求的语义图像检索。本文第2节给出描述图像所使用的特征向量; 第3节介绍我们采用的学习机器, 支持向量机(SVM); 第4节是实验的设计和结果; 第5节给出总结以及进一步的工作。

2 提取图像特征向量

2.1 图像块划分

* 基金项目: 山东省中青年科学家奖励基金资助(2001SD521)。

这里我们采用基于图像块划分聚类[3]的方法得到低层特征向量。首先对图像进行预处理——划分。划分克服了图像分割的困难，通过划分把空间信息嵌入到一个个栅格或是小块中。然后提取每个栅格的颜色直方图，这种特征既有颜色信息又有空间信息。大小不同的块扑捉不同的尺度信息，通常较小的块揭示局部的图像内容信息，例如具有高频空间的边缘和区域；较大的块描述相邻子块的相关信息，反映全局变化。综合权衡两方面因素，我们在实验中采用5×5的窗口分割图像。

2.2 颜色空间的选择及量化

由于研究的侧重点不同，人们定义了多种颜色空间。HSV颜色空间是基于人类对颜色的感知定义的，能够准确、定量地描述颜色特征。在我们的研究中用到的是HSV颜色空间。HSV颜色空间的三个颜色特征是色调、饱和度以及亮度值。如果我们在光照变换的环境中检测某种颜色，我们通常只用其H和S分量而舍弃V分量[4]。原始图像都是在RGB空间存储的，所以首先进行空间变换。

给定RGB颜色空间的值(r, g, b), $r, g, b \in [0, 1, \dots, 255]$, 则转换到HSV空间的 h, s, v 值计算如下:

$$\text{设 } v' = \max(r, g, b) \text{ 则 } v = v' / 255 \quad (1)$$

$$s = \frac{v' - \min(r, g, b)}{v'} \quad (2)$$

$$h' = \begin{cases} (5 + b'), & \text{if } r = \max(r, g, b) \text{ and } g = \min(r, g, b) \\ (1 - g'), & \text{if } r = \max(r, g, b) \text{ and } g \neq \min(r, g, b) \\ (1 + r'), & \text{if } g = \max(r, g, b) \text{ and } b = \min(r, g, b) \\ (3 - b'), & \text{if } g = \max(r, g, b) \text{ and } b \neq \min(r, g, b) \\ (3 + g'), & \text{if } b = \max(r, g, b) \text{ and } r = \min(r, g, b) \\ (5 - r'), & \text{otherwise} \end{cases}$$

$$h = h' / 6 \quad (3)$$

其中 r', g', b' 定义为

$$r' = \frac{v' - r}{v' - \min(r, g, b)}, \quad g' = \frac{v' - g}{v' - \min(r, g, b)}, \quad b' = \frac{v' - b}{v' - \min(r, g, b)} \quad (4)$$

这里, $r, g, b \in [0, 255]$, $h \in [0, 1]$, $s \in [0, 1]$, $v \in [0, 1]$ 。

一幅彩色图像中包含至少上百种颜色。颜色直方图的高维数妨碍了有效的计算和检索，占用了过多的空间。为了减少直方图的维数，在保证图像内容损失最小的情况下，颜色空间量化是必不可少的。通过实验我们发现可以将 HSV 颜色空间的每一维分量等分成 10 份，由于只选取了其中对分辨颜色较为重要的 H 分量和 S 分量，所以我们的颜色直方图量化为 100 个颜色。颜色量化大大降低了后续处理的计算复杂度，而且不会丢失过多的颜色信息。

2.3 特征向量聚类(用 Kohonen 自组织特征映射神经网络实现)

图像划分成邻接的小块后，提取每个小块的颜色直方图。为了减少计算的复杂度和更好的抽象出图像的内容，我们将所有的向量聚类。聚类的算法有很多，例如最近邻聚类算法、

k-均值算法、分级聚类算法、遗传算法和人工神经网络聚类算法等。其中 k-均值算法和 Kohonen 网络常用于大型的数据聚类，因为这两种算法的控制参数较容易设定，而且他们的执行速度较快[5]。由于用 k-均值算法聚类时，聚类结果的好坏决定于起始的聚类中心的选取，容易使聚类结果陷于局部最优，所以在本实验中用到 Kohonen 网络实现聚类。具体方法如下：

我们用前面得到的 HSV 颜色空间的直方图（即一个 100 维特征矢量）来描述每一幅图像所划分的块。如一幅 300×200 的图像分为 2400 个块。然后提取某类图像的所有块的特征矢量，用 Kohonen 自组织特征映射网络算法聚类，每类的聚合中心作为网络的输出。每一幅图像我们调整聚类数目为 200。然后将某类图像所有的聚类结果再次聚类，聚类数目仍为 200。经过聚类，得到数目较少的一些矢量（这里是 200 个 100 维向量），这些矢量包含了某类图像的语义信息，可以用来表征此类图像。我们把这些聚类后数目较少的矢量称为矢量组。为了描述某类中每个具体图像，我们要计算表示每幅图像的特征矢量中距离矢量组中每一特征矢量最近的个数。也就是说，纪录矢量组中每一个特征矢量在某一幅图像特征矢量中所占的比例。这个比例即作为描述每一幅图像的特征。这样，图像的特征不仅与此图像本身有关，还与此类别所有图像有关，可以反映某一类图像特征。我们可以将此特征作为 SVM 的输入（即一个 200 维的矢量），进行语义分类训练。对于待分类的每一个图像，在第一次聚类完成后，用同样的方法得到描述它的一个 200 维向量，作为 SVM 的输入。

3 学习机器——支持向量机 (SVM)

支持向量机 (SVM) 是 Vapnik 等人[6] 提出的一类新型通用机器学习方法。由于 SVM 方法有统计学习理论作为其坚实的数学基础，并且可以很好地克服维数灾难和过拟合等传统算法所不可规避的问题，所以受到了越来越多的研究人员的关注。该技术已成为机器学习界的研究热点，并在很多领域都得到了成功的应用，如人脸检测 [7]、手写体数字识别 [8]、文本自动分类 [9] 等。下面对 SVM 作简单介绍：

SVM 是基于统计学习理论的结构风险最小化原则的，它将最大分界面分类器思想和基于核的方法结合在一起，表现出了很好的泛化能力。由于 SVM 方法不仅考虑了对渐近性能的要求，而且在现有有限信息的条件下得到最优结果，并且能够根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳折中，从而获得最好的推广能力。

3.1 线性可分情况

SVM 算法是从线性可分情况下的最优分类面 (Optimal Hyperplane) 提出的。所谓最优分类面就是要求分类面不但能将两类样本点无错误的分开，而且要使两类的分类空隙最大。

d 维空间中线性判别函数的一般形式为 $g(x) = (w^T \cdot x) + b$ ，分类面方程是 $w^T x + b = 0$ ，我们将判别函数进行归一化，使两类所有样本都满足 $|g(x)| \geq 1$ ，此时离分类面最近的样本都满足 $|g(x)| = 1$ ，而要求分类面对所有样本都能正确分类，就是要求他满足

$$y_i[(w^T \cdot x_i) + b] - 1 \geq 0, i = 1, 2, \dots, n \quad (4)$$

式 (4) 中使等号成立的那些样本成为支持向量 (Support Vectors)。两类样本的分类空隙 (Margin) 的间隔大小：

$$\text{Margin} = 2 / \|w\| \quad (5)$$

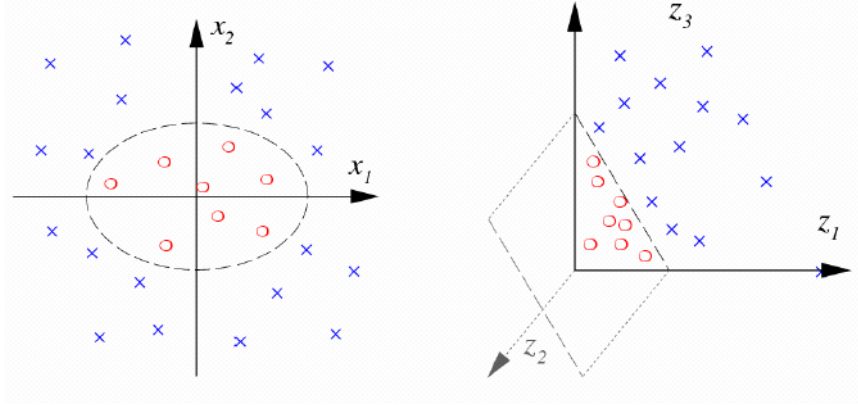
因此，最优分类面问题可以表示成如下的约束优化问题，即在式 (4) 的约束下，求函数

$$\Phi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w^T w) \quad (6)$$

的最小值。解这个最优化问题便得到最优分类函数：

$$f(x) = \text{sgn}((w^*)^T x + b^*) = \text{sgn}\left(\sum_{i=1}^n a_i^* y_i x_i^* x + b^*\right) \quad (7)$$

3.2 非线性可分情况



图（1）通过非线性变换将输入空间变换到一个高维空间

当在原始空间中的简单超平面不能得到满意的分类效果时，必须以复杂超曲面作为分界面。SVM通过非线性变换 Φ 将输入空间变换到一个高维空间（图1），然后在这个新空间中求取最优线性分类面，而这种非线性变换是通过定义适当的核函数（内积函数）实现的，用核函数 $K(x_i, x_j)$ 代替最优分类平面中的点积 $x_i^T x_j$ ，就相当于把原特征空间变换到了某一新的特征空间，从而使学习样本成为线性可分的。此时优化函数变为：

$$f(x) = \text{sgn}[(w^*)^T \phi(x) + b^*] = \text{sgn}\left(\sum_{i=1}^n a_i^* y_i K(x_i, x) + b^*\right) \quad (8)$$

常用的核函数有：多项式形式的核函数，即 $K(x, x_i) = [(x^T x_i) + 1]^q$ ，对应SVM是一个q阶多项式分类器；径向基形式的核函数，即 $K(x, x_i) = \exp(-\sigma \|x - x_i\|^2)$ ，对应SVM是一种径向基函数分类器；s形核函数，如 $K(x, x_i) = \tanh(v(x^T x_i) + c)$ 。

4 实验

实验用到的图像来源于图像检索专用数据库。将第2节得到的特征向量作为SVM的输入。其中图像划分以及特征提取、聚类均利用matlab6.5完成。利用LIBSVM[10]训练图像的语义分类。我们选取湖泊，雪山，红花，古建筑四类图像。每类图像80幅，按5: 3的比例分成训练集和测试集，采用“一对其它”的方式，即将这四类分类分解为4个二类分类，正例50幅，反例随机选自其它三类图像，亦取50幅。训练时采用第3节介绍的三种核函数分别实验（实验1），只列出效果最好的径向基形式的核函数，其中参数 $\sigma=2$ ，惩罚系数 $c=100$ 。进一步实验证明，当训练集包含所有图像集合时，分类准确率有一定提高（实验2）。实验结果见表（1）。

	湖泊	雪山	红花	古建筑
准确率(实验1)	83.3%	80.0%	73.3%	63.3%
准确率(实验2)	90.0%	80.0%	76.6%	70.0%

表1 实验结果

5 结论

实验证明基于块划分及聚类提取特征向量作为机器学习的训练样本的方法是可行的,做到了图像语义自动分类,为下一步基于语义检索提供了前提。此方法受到训练图像包含样本多少的影响,当训练图像包含样本较少时,准确率有所下降。另一个缺陷是图像特征提取比较繁琐,聚类时间较长,当图像库中图像数量增多时,预处理工作时间增长。进一步研究的内容包括更有效的特征提取方法,以及下一步图像检索时语义规则的建立。

参考文献

- [1] Y. Rui, T.S. Huang, S.F. Chang, "Image Retrieval: Past, Present, And Future"[A], Proceedings International Symposium on Multimedia Information Processing[C], 1997.
- [2] Colombo C, Bimbo AD, Pala P. Semantics in visual information retrieval[J]. IEEE Multimedia, 1999,6(3):38~53.
- [3] 刘金梅, 王国宇. 一种基于块划分颜色特征的图像检索方法[J], 中国海洋大学学报, 2005,35(2):330~334.
- [4] Ying Wu. Color, Edge and Texture[Z].ECE510-Computer Vision Notes Series 3
- [5] A.K. JAIN, M.N. MURTY and P.J. FLYNN, Data Clustering: A Review [J], ACM Computing Surveys, Vol. 31, No. 3, September 1999
- [6] Cortes C, Vapnik V. Support Vector Networks[J]. Machine Learning, 1995, 20: 273~297.
- [7] Osuna E, Freund R, Girosi F. Training Support Vector Machines: An Application to Face Detection[A]. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition[C], New York: IEEE, 1997, 130~136.
- [8] Joachims T. Text Categorization with Support Vector Machines: Learning with Many Relevant Features[A]. In: Proceedings of the 10th European Conference on Machine Learning[C], 1998.
- [9] Dumais S, Platt J, Heckerman D, Sahami M. Inductive Learning Algorithms and Representations for Text Categorization[A]. In: Proceedings of the 7th International Conference on Information and Knowledge Management[C], 1998.
- [10] C.-W. Hsu, C.-C. Chang, C.-J. Lin. A Practical Guide to Support Vector Classification.[Z]. Department of Computer Science and Information Engineering. National Taiwan University.

作者简介:

邢慧强 (1980-), 男 (汉族), 山东青岛人, 硕士, 主要研究方向: 图像处理、图像检索。
 王国宇 (1962-), 男 (汉族), 山东青岛人, 博士, 教授, 主要研究方向: 图像处理、模式识别。

联系人: 邢慧强

通信地址: 中国海洋大学信息学院电子系 03 级研究生, 邮编, 266071

e_mail: qdxinghq@163.com