

文章编号: 1000 - 582X(2006)01 - 0057 - 04

基于支持向量机的图像识别*

何江平¹,文俊浩²,邓恬洁³,王道乾²

(1. 重庆工学院 数理学院, 重庆 400050; 2 重庆大学 软件学院, 重庆 400030; 3 南京大学 软件学院, 江苏 南京 210093)

摘要:支持向量机是统计学习方法,正成为当今研究的热点.支持向量机在模式识别和文本分类等方面获得了极大的成功,分类的准确率很高,用支持向量机的方法处理一些二值图像和灰度图像,能获得较好的统计结果.从中摸索出了一种特征向量集的选取方法,讨论了判断结果优劣的标准,比较了支持向量机方法与其他方法得到的结果,得出了重要结论:用支持向量机识别图像的边缘具有非常优异的统计性能.

关键词:支持向量机;图像识别;边缘检测;统计学习;数字图像处理

中图分类号: TP391.41

文献标识码: A

V. Vapnik提出的支持向量机 (support vector machine,以下简称 SVM)的基本思想是:先在样本空间或特征空间,构造出最优超平面,使得超平面与不同类样本集之间的距离最大,从而达到最大的泛化能力.支持向量机结构简单,并且具有全局最优性和较好的泛化能力,自 20世纪 90年代中期提出以来得到了广泛的研究.支持向量机方法是求解模式识别^[1]和函数估计问题^[2]的有效工具.

SVM在数字图像处理方面的应用是:寻找图像像素之间的特征的差别,即从像素点本身的特征和周围的环境(临近的像素点)出发,寻找差异,然后将各类像素点区分出来.

对于较复杂的图形要区分边界像素^[3]和非边界像素是比较困难的,仅仅依靠像素本身的值无法区分.当加入灰度均值和灰度方差以后情况发生了变化,能够较好地区分边界像素和非边界像素,说明这些特征加在一起能够刻画边界和非边界.从直观上来讲,边界上的元素和周围的像素的方差较大,均值和原来像素的差也较大,因此,把灰度均值和灰度方差纳入了特征向量^[4].首先选取 2幅比较复杂的图形作为处理的对象,分别命名为 circle和 clip,见图 1.

把图像读入内存,转化为灰度图的格式.分别求得需要作为特征向量组成部分的梯度矩阵^[5]、均值矩阵、方差矩阵等.然后将每个矩阵降维成一维列向量,

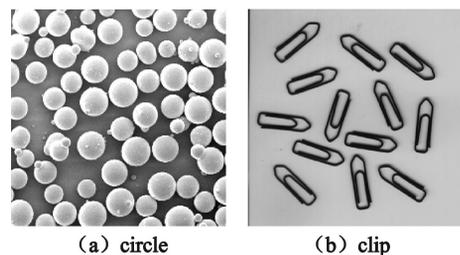


图 1 处理对象

将 5个一维列向量合并为一个多列向量.以这个多维列向量为特征向量,用 Sobel算子的方法求得图像边缘的矩阵,降维成一维列向量,以此作为目标集^[6-7].然后用 LS-SVM进行训练,分类,最后显示统计结果.

2幅图像较大, circle是 256 × 256像素, clip是 830 × 822像素.从运行时间上分析,98%以上的时间都在训练的时候用掉的,判别的时候花掉大约 1%,其他处理,比如获得特征向量、转化矩阵及降维用的时间非常少,总共不足 1%,见表 1.这主要源于以下 2点: 1) SVM的内在机制是通过求解二次规划问题来得到学习机,这是最重要的理论方面的原因; 2)工具包使用 MATLAB语言^[8],这个语言本身效率很低.

表 1 各项所用时间百分比

计算项	训练	判断	其他诸如准备、显示等
所用时间比例	98%	1%	1%

* 收稿日期: 2005 - 08 - 29

基金项目:重庆市自然科学基金资助项目 (2004BB2182)

作者简介:何江平 (1969 -),女,重庆人,重庆工学院讲师,硕士,主要研究方向:图形图像处理、基础数学等.

基于上面的原因,为了使计算量减少,只有将图像 clip 都用绘图工具缩放为 256 ×256 像素.

运算量主要在训练上,那么开始可试图用较小的训练集进行训练,如果结果很好,就不需要继续,如果结果欠佳,则需要加大训练集以达到较好的效果.

1 以像素作为训练集的图像识别

1.1 以 1/16 像素作为训练集的图像识别

先用图像的左上角的 1/16 像素作为训练集,然后识别整个图像,处理时间见表 2

表 2 1/16 作为训练集的处理时间

图像	circle	clip
处理时间 /s	445.64	465.32

第 1 种方法:只考虑边缘,也就是说计算正确率的分子分母都是边缘上的点的个数,按照这种算法,得到的正确率见表 3.

表 3 第 1 种方法的正确率 (以 1/16 作为训练集,识别整个图像)

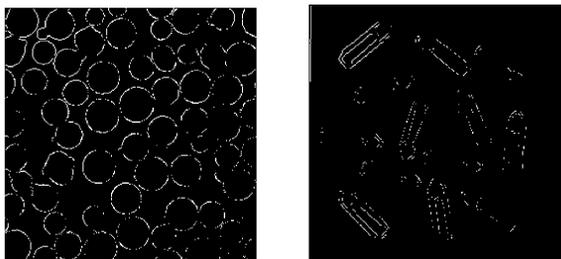
图像	Circle	clip
正确率 /%	47.32	15.07

第 2 种方法:只要识别正确,不管是不是边缘,这种算法下,得到的正确率见表 4.

表 4 第 2 种方法的正确率 (以 1/16 作为训练集,识别整个图像)

图像	Circle	clip
正确率 /%	95.77	95.43

因为只关注边缘,得到 47.32% 和 15.07% 的正确率 (见表 3). 从视觉上看 (见图 2), 这个结果也不能接受.



(a) 训练 1/16 circle 识别整个图像 (b) 训练 1/16 clip 识别整个图像

图 2 训练 1/16, 识别整个图像的结果

猜想: circle 较 clip 要简单、规则一些,对 circle 的处理效果会好一些. 从上面 2 幅图中看到:无论是从统计数字,还是视觉上看,对 circle 的结果明显较好. 这主要是 circle 这幅图较 clip 简单、规则,在以后的处理结果中,也期望这个结论是正确的.

1.2 以左上角 1/16 作为训练集,识别左上 1/4 图像

先用图像左上角的 1/16 的像素作为训练集,然后识别左上 1/4 的图像,处理时间见表 5.

表 5 训练 1/16, 识别 1/4 的处理时间

图像	circle	clip
处理时间 /s	419.54	427.36

第 1 种方法:只考虑边缘,也就是说计算正确率的分子分母都是边缘上的点的个数,按照这种算法,得到的正确率见表 6.

表 6 第 1 种方法的正确率 (以 1/16 作为训练集,识别 1/4)

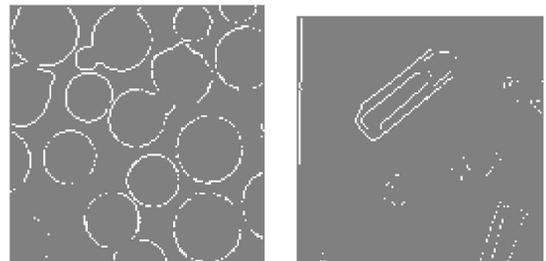
图像	circle	clip
正确率 /%	62.97	12.25

第 2 种方法:只要识别正确,不管是不是边缘,按照这种算法,得到的正确率见表 7.

表 7 第 2 种方法的正确率 (以 1/16 作为训练集,识别 1/4)

图像	circle	Clip
正确率 /%	96.79	94.85

因为只关注边缘,得到的 62.97% 和 12.25% 的正确率 (见表 6). 从视觉上看 (见图 3), 12.25% 这个结果也不能接受.



(a) circle 训练 1/16 识别 1/4 (b) clip 训练 1/16 识别 1/4

图 3 训练 1/16, 识别 1/4 图像的结果

训练 1/16, 识别整个图像和识别 1/4 图像的结果证实了猜想是正确的. 也说明使用 SVM 对数字图像的处理虽然在 SVM 处理的过程中没有应用环境的因素,仅仅从每个像素本身的特征向量出发,但是在预先的处理过程中,考虑了环境因素,这些环境因素在数学上的表达就是周围像素的值,这里用到了梯度、均值、方差等,这些都是与环境有关的参量. 因为 circle 规则,所以环境很容易描绘,所以处理的时候,用相同的参数,结果就好些.

1.3 用半幅图像作为训练集,识别整个图像

鉴于前面的处理效果都不是很好,至少 clip 相关的统计数据以及视觉效果是不能让人接受的,所以用半幅图像作为训练数据,不再考虑计算机处理时间,但把计算机处理时间记录下来作为比较的数据.

先用图像上面一半的像素作为训练集,然后识别整个图像,计算机的处理时间见表 8.

表 8 1/2 图像作为训练集的处理时间

图像	circle	clip
处理时间 /s	5280	3300

从表 8 中可以看出,花费时间很长,不过结果还是让人倍受鼓舞的. 在视觉效果来看,图 4 的结果不错.

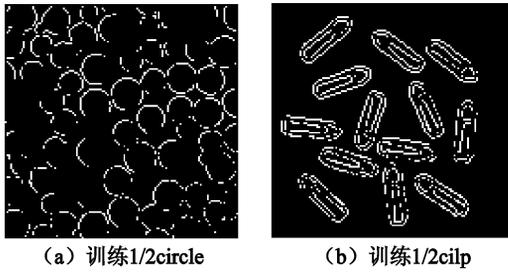


图 4 训练 1/2,识别整个图像的结果

第 1 种方法:只考虑边缘,也就是说计算正确率的分子分母都是边缘上的点的个数,按照这种算法,得到的正确率见表 9.

表 9 第 1 种方法的正确率 (以 1/2 作为训练集,识别整个图像)

图像	circle	clip
正确率 / %	68.94	68.72

第 2 种方法:只要识别正确,不管是不是边缘,按照这种算法,得到的正确率见表 10.

表 10 第 2 种方法的正确率 (以 1/2 作为训练集,识别整个图像)

图像	circle	clip
正确率 / %	96.88	96.07

比较图 2 和图 4 发现:前者明显比后者在观感上的效果要好,但是后者在统计数据上却高出了很多.为什么达到了那么高的正确率,视觉效果还是不好?

首先因为主要是从单一的像素入手,没有考虑像素间的关系,其他的方法是从视觉的方面入手,考虑了周边像素的关系,比如细化的算法,总是考虑连通性,甚至牺牲正确性,这样得到的视觉效果当然好.出发点不同,得到的结果也会不同.但是可以考虑用连接相邻的边缘像素的算法,对结果进行处理,使之视觉效果增强.在本文中笔者要做的是对 SVM 学习机的学习能力、识别能力的检验与实现,采用的是用数据说话,统计性能为标准,而不是要得到“好看”的图像.

其次是所采用的标准的原因.采用其他算法所获得的边缘作为判断的标准,这个所谓的边缘是不是真正的边缘?还要用实验结果去和它吻合.

最后进行横向比较,在统计数据上,circle 对应的数据要比 clip 对应的数据好,印证了前面的猜想:circle 较 clip 要简单、规则一些,对 circle 的处理效果会更好一些.

2 目标集的指定方法和对结果的判断标准

在前面的讨论中,所采用的训练集的标准是直接使用了 sobel 算子的方法求得图像边缘的矩阵,并且用这个矩阵的相应部分作为训练集的目标集,后来又用它作为训练结果统计数据的标准.但是必须注意到:这个标准不是绝对可靠的.因为究竟哪个像素才是边缘

无法判断.边缘本身是一个模糊的概念.但是必须找到一种可靠的标准.如把一幅只由线条组成的图像叠加到一幅没有明显边缘的平滑的图像上,那么原来的线条就是后来生成图像的边缘.这个方法简单易行,可以有效验证 SVM 的性能.在实际问题中,可以逐步积累“经验”,让 SVM 学习,但是 SVM 不支持增量学习.

3 特征向量集的选取方法

选取特征向量一方面是要使特征足以表达所要识别的不同种类的对象特征,这是“完备性”的要求.另一方面,又要减少“冗余性”,要求各个向量的相关性要小,能够达到正交,不过实际应用中一般是达不到这样的条件.只要能够达到相关性小,不是线性相关就可以了.对于线性相关可以应用线性代数以及线性回归的相关知识获得判断的依据.所以首要的任务是找到尽量多的特征向量,使得它们联合起来能够刻画不同种类的对象特征.然后在此基础上,应用独立的向量一定不相关,找出在最后特征向量集中的那些向量.这样得到一组无关的向量组.之后,逐个判断其他向量是否与它们线性相关,相关系数有多大,决定一个阈值,小于这个阈值的就将它加入(阈值越大,最后选入的向量个数可能就多,最后的完备程度就可能越高,运算量当然也大,要在运算量和效果之间权衡).最后再判断其他向量和已经得到的那个组内的向量的相关性,一直到所有向量都判断完成;或者中途试验一下,看效果是否已经令人满意,如果已经达到了很好的效果,则不需要继续判断剩下的向量.总之,要在付出的代价(运算量)和获得的效果两方面进行权衡.

由于一开始选取向量的时候就考虑了相关性,在使用的 5 个向量 $T = \{\text{像素值}, DX, DY, \text{灰度均值}, \text{灰度方差}\}$ 中,毋庸置疑,前面 4 个向量是独立的,一定不相关,肯定应该选入特征向量集.至于灰度方差和 $\{\text{像素值}, DX, DY, \text{灰度均值}\}$ 虽然不是线性相关,但是却是二次相关,权衡识别效果,也将它加入,到这里正如笔者所看到的一样,效果已经很好,不必再加入其他的向量在特征向量集中了.总之,选取特征向量集要考虑“完备性”和“相关性”,在效果和效率之间权衡.

4 SVM 和其他方法获得的边缘效果比较

在数字图像处理中,当前用得比较广泛的有 Robert 算子、Sobel 算子、Pewitt 算子、LOG 算子、Canny 算子等.下面将以 circle 为例对它们的效果进行比较,阐述其中的差异,分析原因.比较图 circle 对应的各种方法的效果,见图 5.

图 5 中的 6 幅子图依次是使用 Robert 算子、Sobel

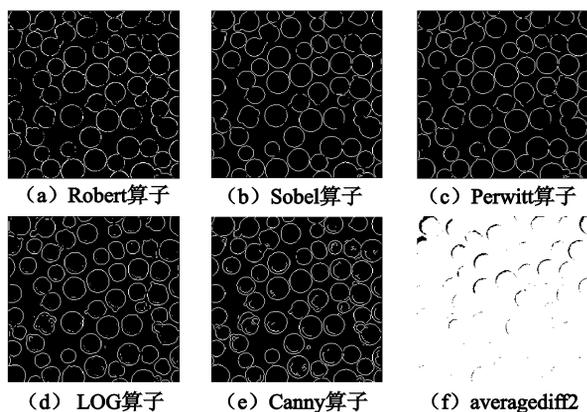


图 5 各种算子和 averagediff2 得到的结果

算子、Prewitt算子、LOG算子和 Canny算子求图像的边缘以及 averagediff2函数所得到的图像. 它们和图 6 (用 LS - SVM方法得到的边缘)相比较,除了 averagediff2以外,其余图像在视觉上都更加“好看”。

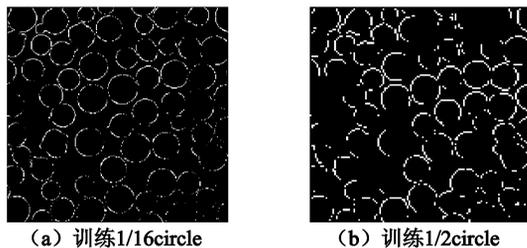


图 6 训练 1/16circle 和 1/2circle 处理结果

但是,很容易看出它们彼此差异非常大. 如果都以 Sobel方法作为参考,对吻合程度作了如下统计:

定义吻合的程度由前面提到的第 1 种方法来表示 (用第 2 种标准大家的吻合度都很高,分辨率比较低)。

只考虑边缘,以此作为标准,表 11 - 12 列出了它们和 Sobel 相比较的吻合的数据 (称为吻合度)。

表 11 各种方法的吻合度 (由 circle 试验获得的数据)

算子	Sobel	Roberts	Prewitt	Log	Canny
Sobel	1.000 0	0.444 1	0.980 1	0.491 4	0.614 1
Roberts	0.434 9	1.000 0	0.442 0	0.311 3	0.388 8
Prewitt	0.956 3	0.440 5	1.000 0	0.479 0	0.601 4
Log	0.540 7	0.349 8	0.540 1	1.000 0	0.471 5
Canny	0.763 5	0.493 6	0.766 2	0.532 8	1.000 0

表 12 各种方法的吻合度 (由 clip 试验获得的数据)

算子	Sobel	Roberts	Prewitt	Log	Canny
Sobel	1.000 0	0.493 9	0.971 6	0.208 1	0.664 4
Roberts	0.559 2	1.000 0	0.559 2	0.255 2	0.531 6
Prewitt	0.971 2	0.493 8	1.000 0	0.207 3	0.668 4
Log	0.224 1	0.242 7	0.223 3	1.000 0	0.282 5
Canny	0.750 5	0.530 4	0.755 3	0.296 4	1.000 0

从表 11 - 12 中的数据可以看出,由几种算子得到的边缘虽然看起来都非常漂亮,但是它们都不是一致的. 因为除了 Prewitt 和 Sobel 以外,其他的吻合度都不高,这和它们的算子形式有关系. Prewitt 和 Sobel 算子

只有 2 个值大小不一致,而 Robert 算子是 2×2 的矩阵,LOG 算子和其他的 3×3 的算子的非 0 元素的分布以及正负都不一致,这也使得 LOG 算子所得的结果和其他算子所得的结果差异最大 (表格中的数字小说明差异大,数字大说明差异小,等于 1 表示完全一样)。

不同的方法差异很大,也就是说,其中一种算法认为的边缘,另外一种算法不一定认为是边缘. 然而, SVM 方法,试图达到这样的效果:只要前面训练集中选用的方法 (主要用 Sobel 算子,也可以使用其他算子)认为是边缘,以后在识别的时候也要认为那是边缘. 在这样的假设下,使用机器学习的方法,吻合度能够达到 68.94% 和 68.72% (比 Prewitt 低,因为 Sobel 算子和 Prewitt 算子相近)。

5 结束语

SVM 的统计性能非常优异,在准确分类上, SVM 具有很高的性能,和使用 Robert 算子、Sobel 算子、Prewitt 算子、LOG 算子和 Canny 算子求图像边缘的机制不同,不能只从视觉上看效果——使用 SVM 的初衷是获得较好的统计性能,这是机器学习的特点,因为机器可没有人的肉眼那样感性。

参考文献:

- [1] DANIEL SAGE, FRANCK R NEUMANN, FLORENCE HEDIGER, et al. Automatic Tracking of Individual Fluorescence Particles: Application to the Study of Chromosome Dynamics[J]. IEEE Trans Image Processing, 2005, 14(9): 1372 - 1383.
- [2] YIKEN CHEN, JAMES ZWANG, ROBERT KROVETZ. CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning[J]. IEEE Trans Image Processing, 2005, 14(8): 1187 - 1199.
- [3] YONGJIAN YU, SCOTT TACTON. Edge Detection in Ultrasound Imagery Using the Instantaneous Coefficient of Variation[J]. IEEE Trans Image Processing, 2004, 13(12): 1640 - 1655.
- [4] RISHIR RAKESH, PROBAL CHAUDHURI, MURTHY C A. Thresholding in Edge Detection: A Statistical Approach[J]. IEEE Trans Image Processing, 2004, 13(7): 927 - 936.
- [5] BARNARD K, DUYGULU P, FORSYTH D, et al. Matching Words and Pictures[J]. J Mach Learn Res, 2003, (3): 1107 - 1135.
- [6] CARSON C, BELONGIE S, GREENSPAN H, et al. Blobworld: Image Segmentation Using Expectation Maximization and Its Application to Image Querying[J]. IEEE Trans Pattern Anal Mach Intell, 2002, 24(8): 1026 - 1038.
- [7] 王新成. 高级图像处理技术 [D]. 北京: 中国科学技术出版社, 2001.
- [8] 孙兆林. MATLAB 6 X 图像处理 [D]. 北京: 清华大学出版社, 2002.

(下转第 65 页)

Vehicle Type Recognition by Using Support Vector Machine SVM

XIAO Han-guang^{1,2}, CAI Cong-zhong^{1,2}, WANG Wan-lu¹

(1. Department of Applied Physics, Chongqing University, Chongqing 400030, China;

2. Department of Computational Science, National University of Singapore, Singapore 117543)

Abstract: The Support Vector Machine (SVM) has shown excellent learning and generalization ability in the practice problems of binary classification, and has been widely employed in multi-class classification. Based on the framework features of the vehicles, the SVM is used to classify 4 types of vehicles. The results of the SVM are compared with that of different classifiers. The testing accuracy to this vehicle dataset reaches 85.59% by means of 9-fold cross-validation which demonstrates that the classification performance of SVM is superior to those of other classifiers.

Key words: support vector machine (SVM); vehicle recognition; framework feature

(编辑 陈移峰)

(上接第 60 页)

Image Recognising Based on Support Vector Machine

HE Jiang-ping¹, WEN Jun-hao², DENG Tian-jie³, WANG Dao-qian²

(1. School of Mathematical Sciences, Chongqing Institute of Technology, Chongqing 400050, China;

2. School of Software Engineer, Chongqing University, Chongqing 400030, China;

3. Software Institute, Nanjing University, Nanjing Jiangsu 210093, China)

Abstract: Support vector machine is the uniform method of statistical learning method, and has become more and more popular in research field. Support vector machine has achieved excellence in pattern recognition and text classification for its high performance in veracity. Support vector machine method was used to process several Binary image and Gray scale image and got a good statistical result. A strategy to select feature vectors has been found out, and the criterion to judge the results are discussed. Compared with other methods, the important conclusion was draw: it performs perfectly in statistic when using support vector machine to detect the image edge.

Key words: support vector machine; image recognising; edge detection; statistical learning; digital image processing

(编辑 张 苹)