

基于 MEB-SVM 的静态手势识别研究

章丰明, 任 彧*

(杭州电子科技大学 智能与软件研究所, 浙江 杭州 310018)

摘要: 为了实现快速高准确率的静态手势识别, 提出了一种新的支持向量机分类方法。图像分割是图像识别的基本步骤, 采用彩色直方图为特征可以将手部区域从背景中分离出来; 然后将得到的手部图像去除手部以外的区域, 并二值化, 得到手的轮廓; 再用 8 邻域方法处理轮廓图, 得到一个表示轮廓的坐标序列, 再将复数序列采用傅里叶变换, 从而得到一个大小、平移、旋转不变的一维矢量。最后对矢量用 MEB-SVM 进行分类、识别。通过这种方法完成的手势识别正确率达到了 90% 以上, 基本实现了静态手势识别的目标。

关键词: 支持向量机; MEB-SVM; 识别; 分割

中图分类号: TH7; TP391.4

文献标识码: A

文章编号: 1001-4551(2010)06-0120-04

Static gesture recognition research with MEB-SVM

ZHANG Feng-ming, Ren Yu

(Institute of Software and Intelligence, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: Aiming at realizing high speed and high accuracy hand gesture recognition, a new support vector machine(abbreviated as SVM) classifier was proposed. Image segmentation is a first step of image recognition color-histogram method was applied to dig hand area from background; then binaries the hand image got from last step, and then, get the hand gesture's contour line; after that 8-connected neighborhood method was used to deal with the contour line to get a serial coordinates denote the main information of hand gesture, then use Fourier transform to get a suitable vector. Finally MEB-SVM was used to classify these vectors and get the meaning of hand gestures. The result indicates that the MEB-SVM method can get an accuracy of 90% in hand gesture recognition. And it can understand most hand gestures meaning.

Key words: support vector machine; MEB-SVM; recognition; segmentation

0 引 言

图像识别是人机交互方向的重要领域, 是让计算机理解人的行为的一种技术。手势识别在人机交互应用上能实现人类和计算机的自然交流。手势识别可分为静态手势识别和动态手势识别两类。静态手势如图 1 所示。

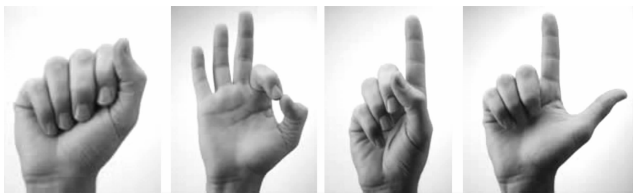


图 1 静态手势图

最简单的静态手势识别可以只是识别出目标手所表示的数字, 但对于复杂的手势则需要用到分类器来识别。Atid S. 等通过多分辨率和图像匹配实现了对戴手套的手的 26 个字母的识别^[1]。Jiang Qing Wu 等利用数据手套工具, 采用 DGMM 技术作为识别手段, 能够较快速而准确地实现手势词的识别^[2], 但缺乏交互的自然性。

本研究提出并实现了一个基于 2D 视觉、以 MEB-SVM(最小包围球的支持向量机)^[3]为分类方法的手势识别方法, 能有效解决以上问题。

1 Mean shift 图像分割

Mean shift 最初是以一个向量名词被提出的, 且提

出很久都没有得到重视,直到 Yizong Cheng 在 1995 年在他发表的论文中重新对这个问题进行了研究^[4]。Mean shift 是一个可以用作聚类、寻优等问题的解决方法。其基本的思想是一个逐渐寻优的过程。根据 Yizong Cheng 的论文,过程如下:

在一个 o 维的空间内 R^o , 存在 n 个向量 $x_1, x_2, x_3, \dots, x_n$, 空间内的某点 x 的向量可用如下的公式求得:

$$Mh(x) = \frac{\sum_{i=1}^n K(x_i - x)lw(x_i)(x_i - x)}{\sum_{i=1}^n K(x_i - x)lw(x_i)} \quad (1)$$

式中 h —点 x 的邻域样本空间; x_i —落入样本空间中的向量; $K(x)$ —一个单位核函数, 一般是一个高斯核函数; l —一个正定的对称 $o \times o$ 矩阵, 称为带宽矩阵; $w(x_i)$ — x_i 的权重 $w(x_i) \geq 0$ 。

因此, 通过式(1) 得到的向量 x 是 R^o 内在邻域 h 内的向量在核函数作用下的平均偏移向量。

在图像分割中, 一张图像, 即为 R^o , 每个像素是一个 3 维向量 (R, G, B) , h 是像素的邻域。

Mean shift 对图像的处理过程就是逐步地将各个像素向量朝邻域内最优向量接近, 最后以此最优向量为模式归一化此邻域, 完成图像分割。

实验处理结果如图 2 所示。

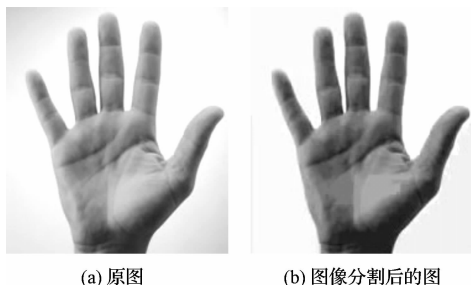


图 2 图像分割

2 傅里叶特征提取

2.1 轮廓提取

特征提取前用高斯模式训练好一个用于手部皮肤的颜色向量 V 。将分割后的图像用向量 V 去除背景, 得到仅剩手部的图像, 最后二值化图像。接下来用傅里叶描述子提取其边界特征。

傅里叶边界特征提取算法的特点有: 对轮廓大小, 平移, 旋转不变性, 且可根据精度和复杂度的平衡情况改变系数个数。

为保证得到的轮廓的封闭性, 可以用八邻域搜索算法提取得到的二值图的轮廓。八邻域搜索算法记录

一个一维序列, 序列中存放边界像素的位置。图中背景像素为零值, 目标像素为非零值。从开始坐标出发, 在当前坐标的八邻域内搜索非零像素, 将搜到的第一个坐标记为序列的下一个坐标, 并将当前坐标存入序列中; 直到当前坐标和开始坐标相等, 则完成搜索。得到的一维序列即为二值图的轮廓, 如图 3 所示。

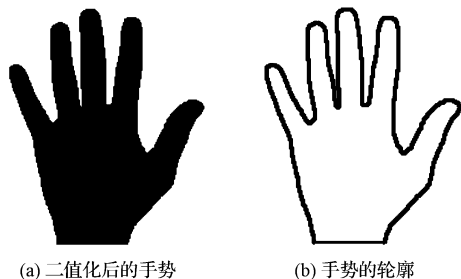


图 3 轮廓提取

2.2 轮廓的傅里叶变换

图像轮廓的傅里叶描述子 (Fourier descriptors), 是对边界曲线的频域变换的结果, 能够得到不受边界旋转、平移和缩放的影响的结果。

傅里叶描述子的作用是将边界曲线上的点都表示成复数形式。 $\{x(i), y(i) | i = 1, 2, 3, \dots, n\}$ 可表示为 $d(m) = x(m) + y(m)i$ ($m = 1, 2, 3, \dots, l-1$)。那么这个一维序列的离散傅里叶系数定义即为:

$$a_n = \frac{1}{l-1} \sum_{m=1}^{l-1} x(m) e^{-i[2\pi/(l-1)]nm} \quad (1)$$

$$b_n = \frac{1}{l-1} \sum_{m=1}^{l-1} y(m) e^{-i[2\pi/(l-1)]nm} \quad (2)$$

但此系数还不是理想的特征向量, 它还不具有旋转、平移和缩放不变性, 需要对此系数进行归一化变换:

$$r_n = (|a_n|^2 + |b_n|^2)^{1/2} \quad (3)$$

r_n 是平移和旋转不变量, 为了具有缩放不变性, 可将每项都除以其幅值 r_1 , 得 w_n :

$$w_n = r_n / r_1 \quad (4)$$

傅里叶描述子的个数的多少表明对轮廓的描述的详细程度, 但若个数太多, 计算复杂度无疑会增加。根据手势识别的需要, 一般通过 12 个描述子就能够区别不同的手势, 且复杂度也不会太高。

3 MEB-SVM

3.1 MEB-SVM 相对于其他 SVM 的优点

SVM (支持向量机)^[5-6] 最初是由 Vapnik 提出来的, 是一种有效的模式分类和识别方法。SVM 的原理是: 对向量空间中的分属不同种类的点, 用一个最优超平面将它们分开, 使得属于同一种类的点被尽可能地

分往同一侧,同时,这两个类之间的间距达到最大。

SVM 基本的理论是针对二分类问题的,对于多个类的分类问题,SVM 包括一对多方法 (OneAgainst-Two),一对一方法 (OneAgainstOne),树分类方法。这些方法在分类过程中普遍存在计算量大、需要多次二分类、算法复杂等缺点。

鉴于以上分类方法的缺点,借用几何学的方法提出了最小包围球的 SVM 方法 (MEB-SVM)。MEB-SVM 是最新的学术研究成果,并成功地应用于大规模数据集分类上。实验表明这个方法相比其他分类方法在解决大规模数据集上有很大的优势,不仅在精确度上,而且在计算时间上也能获得很大的提高。

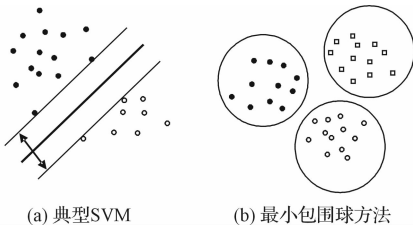


图 4 两种 SVM 分类方法示意图

如图 4 所示,最小包围球方法在多分类问题上存在明显的优势,不用进行多次的二分法,只需要对各个类分别计算其“球心”和“半径”,即可完成一类的判别。

3.2 最小包围球的核方法

在给定在样本空间 $S = \{x_1, x_2, x_3, \dots, x_n\}, x_i \in R^d$,则 S 的最小包围球 $MEB(S)$ 是包含 S 内所有样本的最小的球; $B_R(c), c$ 表示球心, R 表示半径。

和传统 SVM 一样,给定的样本空间并不一定会对各类形成自然归类的分布,所以需要将样本空间映射到特征空间 φ 内。根据泛函理论,只要核函数满足 Mercer 条件,它就对应某一空间中的内积^[7]。本研究中所用的核函数为高斯核函数,也作径向基函数 (RBF): $k(x, x_i) = \exp\{- (\|x - x_i\|^2)/2\sigma^2\}$ 。

因此求最小包围球的问题可表示为:

$$\min_{R, c} R^2: \|c - \varphi(x_i)\|^2 \leq R^2, i = 1, 2, 3, \dots, m \quad (1)$$

其对偶问题为:

$$\text{Subject to: } \alpha_i \geq 0, i = 1, 2, 3, \dots, m$$

$$\sum_{i=1}^m \alpha_i = 1$$

式中 α_i —拉格朗日乘子。

3.3 MEB-SVM 算法

将给定的手势训练样本集 S 经图像分割、二值化等步骤得到一组特征向量;再用基函数对这些向量进

行空间映射。为了能快速得到样本集的最小包围球,按 Ivor W. Tsang 等人的方法可采用近似算法^[8]。其主要思想是通过递增迭代的方式逐步扩展包围球的半径。扩展过程中将离当前球心最远的样本加入包围球中。记第 t 次迭代时的包围球包含的样本为 S_t ,球心和半径为 C_t 和 R_t ,包围球为 B 。则对于给定的 $\varepsilon > 0$,算法可描述为以下过程:

- (1) 对于一类样本,任选一个样本点作为 S_0 ,记球心 C_0 为该点向量,半径 R_0 为 0;
- (2) 结束训练条件:不存在样本点 z 在 $(1 + \varepsilon)$ 包围球 $B(C_t, (1 + \varepsilon)R_t)$ 的外面;
- (3) 找到特征空间内离当前包围球心 C_t 最远的样本点,并添加到当前包围球: $S_{t+1} = S_t \cup \{z\}$;
- (4) 更新包围球的球心和半径 C_{t+1}, R_{t+1} ;
- (5) $t = t + 1$, 返回到 (2);

算法流程图如图 5 所示。

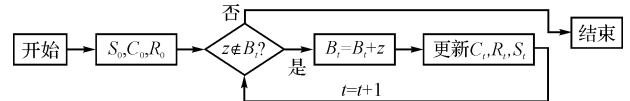


图 5 MEB-SVM 流程图

经过以上步骤后即得到本类的最小包围球。对于各训练样本类都执行以上步骤则得到全样本空间的所有最小包围球。对于测试样本 T 的分类,只需计算 T 与各个包围球的球心 C_i 的欧式距离,并加以比较,则与 T 距离最近的球心所在的包围球的类即为 T 所属的类。

4 实 验

本研究选取了手语中 1~10 的 10 个手势作为测试对象,并分别对每个手势进行 10 次实验。

实验的测试平台为 Intel Core2 Duo 3.00 Hz, 1 G 内存的电脑,程序所使用的编程环境为 Visual Studio 2008 C++ 和 OpenCV 开源图像处理库。图像采集采用具有 130 万像素的 10moons 天弓摄像头。

具体过程如图 6 所示。

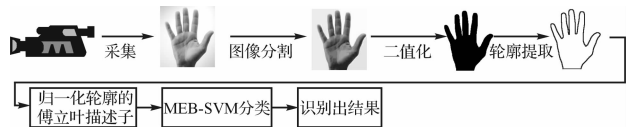


图 6 手势识别整体流程图

在用 Mean shift 对图像分割时,经过试验,笔者以 5×5 为邻域像素,核函数采用高斯核 $N(x) = \exp(-\|x\|^2)$,带宽矩阵简化为 $I = h^2 I$,即为一个单位数量阵。

权重系数按与 x 点的距离赋值。

经二值化和八邻域搜索得到轮廓后,对轮廓即可作傅里叶变换。为使傅里叶描述子的误差尽量减少,作傅里叶变换的轮廓图像应该大小一致。同时计算时需要兼顾计算的复杂度和精确度,经过经验对比可知,对一个轮廓采用 12 个描述子是较为合适的。

MEB-SVM 算法中的 ε 是决定算法精确度与复杂度的关键因素。实验中分别取不同的 ε 进行训练,得到的结果如表 1 所示。

表 1 实验结果		
ε	训练时间/s	识别准确度/%
0.000 5	3.07	80.2
0.000 8	4.11	86.5
0.001	5.29	91.4
0.001 4	7.18	92.2
0.002	10.26	93.6
0.002 5	15.37	94.5

实验结果中, ε 越小,算法收敛时间越短,但由于包含在包围球中的向量太少,对算法的准确度会有一些的影响,当 ε 大于 0.001 以上时,算法训练时间增加明显,同时算法准确度也有所增加。

5 结束语

本研究介绍了一种基于 MEB-SVM 的手势识别算法,MEB-SVM 算法是在具有强分类能力的 SVM 的基础上提出来的,且此方法的计算复杂度与精确度又超过了传统的 SVM 方法,具有非常广泛的应用前景。

参考文献(References):

[1] ATID S, WU H, ALISTAIR S. Hand gesture recognition for human computer interaction[N]. ERCIM NEWS, European Research Consortium for Informatics and MAtematics,2001 (46).

[2] WU J Q, GAO W, PANG B, et al. A fast sign word recognition technique for Chinese sign language[J]. **High Technology Letters**,2001,11(6):23-27.

[3] CERVANTES J, LI Xiao-ou, YU Wen, et al. Support vector machine classification for large data sets via minimum enclosing ball clustering[J]. **Neurocomputing**,2008,71(4-6):611-619.

[4] CHENG Yi-zong. Mean shift, mode seeking and clustering [J]. **IEEE Trans on Pattern Analysis and Machine Intelligence**,1995,17(8):790-799.

[5] CORTES C, VAPNIK V N. Support-vector networks[J]. **Machine Learning**,1995,20:273-297.

[6] BOSER B, GUYON I, VAPNIK V N. A training algorithm or optimal margin classifiers[C]//Proc. Fifth Annual Workshop on Computational Learning Theory. New York: [s. n.],1992:144-152.

[7] 段一洪,陈一民,林 锋. 基于 LSSVM 的静态手势识别[J]. 计算机工程与设计,2004,12(25):2352-2368.

[8] TSANG I W, KWOK J T, CHEUNG P M. Core vector machines: fast SVM training on very large data sets[J]. **Journal of Machine Learning Research**,2005,6:363-392.

[编辑:李 辉]

(上接第 111 页)

参考文献(References):

[1] QUINLAN J R. C4.5: programs for machine learning[M]. San Mateo: Morgan Kaufmann,1993.

[2] LIU Bing, HSU W, MA Yi-ming. Integrating Classification and Association Rule Mining[C]//Proc. of the International Conf. on Knowledge Discovery and Data Mining. New York: [s. n.],1998: [s. n.].

[3] LU J, XU B, KANG D. Classification methods of association rules with linguistic terms[J]. **Journal of Southeast University**,2004,20(1):21-25.

[4] LU J, XU B, JIANG J. A Prediction Method of Fuzzy Association Rules[C]//Proc. of the International Conf. on Information

Reuse and Integration. Nevada: [s. n.],2003: [s. n.].

[5] FRIGUI H, KRISHNAPURAM R. Clustering by competitive agglomeration[J]. **Pattern Recognition**, 1998, 30(7): 1109-1119.

[6] AGRAWAL R, SKIKANT R. Fast Algorithms for Mining Association Rules[C]//Proc. of 20th International Conf. Very on Very Large Databases. Morgan Kaufmann: [s. n.], 1994: [s. n.].

[7] 王立新,王迎军. 模糊系统与模糊控制教程[M]. 北京:清华大学出版社,2003.

[8] SMITH S F. A learning system based on genetic algorithms [D]. Pittsburgh:University of Pittsburgh,1980.

[编辑:李 辉]