

文章编号: 1001-0920(2010)01-0079-05

## 最大间隔最小体积球形支持向量机

文传军<sup>1,2</sup>, 詹永照<sup>1</sup>, 陈长军<sup>1</sup>

(1. 江苏大学 计算机科学与通信工程学院, 江苏 镇江 212013; 2. 常州工学院 理学院, 江苏 常州 213002)

**摘要:** 结合支持向量机(SVM)类间最大分类间隔和支持向量数据描述(SVDD)类内最小描述体积思想,提出一种新的学习机器模型——最大间隔最小体积球形支持向量机(MMH SVM). 模型建立两个大小不一的同心超球,将正负类样本分别映射到小超球内和大超球外,模型目标函数最大化两超球间隔,实现正负类类间间隔的最大化和各类类内体积的最小化,提高了模型的分类能力. 理论分析和实验结果表明该算法是有效的.

**关键词:** 支持向量机; 支持向量数据描述; 类间最大分类间隔; 类内最小描述体积; 球形支持向量机

中图分类号: TP391.41

文献标识码: A

### Maximal-margin minimal-volume hypersphere support vector machine

WEN Chuan-jun<sup>1,2</sup>, ZHAN Yong-zhao<sup>1</sup>, CHEN Chang-jun<sup>1</sup>

(1. School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China; 2. School of Science, Changzhou Institute of Technology, Changzhou 213002, China. Correspondent: WEN Chuan-jun, E-mail: wcjyhy@yahoo.com.cn)

**Abstract:** Inspired with the ideas of support vector machine (SVM) between-class maximal classification margin and support vector data description (SVDD) within-class minimal description volume, a novel learning machine model, maximal-margin minimal-volume hypersphere SVM(MMH SVM), is proposed in this paper. Two different concentric hyperspheres are built in the model, positive samples are packed in small hypersphere and negative samples are excluded outside large hypersphere. The between-class margin is maximized by model objective function, which realizes the maximization of between-class margin and the minimization within-class volume, and the model classification performance is improved. Theoretical analysis and experimental results show the effectiveness of the proposed method.

**Key words:** Support vector machine (SVM); Support vector data description (SVDD); Between-class maximal classification margin; Within-class minimal description volume; Hypersphere support vector machine (HSVM)

### 1 引言

类间间隔和类内聚类性是影响分类器分类性能的两重要因素,增大类间间隔和提高类内聚类性有利于分类能力的提高. 支持向量机(SVM)<sup>[1]</sup>是将若干重要技术有机集成的分类器,通过构造最优超平面寻找类间最大分类间隔以提高分类器的泛化能力. 支持向量机建立在结构风险最小化(SRM)基础上,具有完善的理论支撑和稳定优良的性能表现,在机器学习和模式识别领域得到了广泛的应用<sup>[2,3]</sup>. 支持向量数据描述(SVDD)<sup>[4]</sup>受到支持向量机的启发,对模式样本建立紧致包裹超球,实现对一类样本

的最小体积描述,即对包容一类样本的空间闭域体积最小化压缩——最小化类内体积,本质上等价于类内聚类性的最大化. SVDD能够解决一类样本的识别问题,主要应用于目标拒识、预警技术、变化检测等方面<sup>[5,6]</sup>. 这两种分类方法各自考虑了影响分类的两种因素:最大化类间间隔和最大化类内聚类性. 一个很自然的想法是希望将这两种方法结合起来,即在扩大分类间隔时,也提高类内聚类性,从而增强分类器的分类性能. 文献[7]提出分隔超球模型(SH),采用类似支持向量机的建模结构,力图通过一个超球将正负两类样本分隔. 支持向量机假设两

收稿日期: 2009-02-12; 修回日期: 2009-04-26.

基金项目: 国家自然科学基金项目(60673190).

作者简介: 文传军(1976—),男,重庆万州人,讲师,博士生,从事模式识别、表情识别等研究;詹永照(1962—),男,福建尤溪人,教授,博士生导师,从事计算机图形学、人机交互等研究.

类样本基于超平面对称,但两类样本基于分隔超球并不对称.文献[8]对 SH 进行改进,得到最大间隔球形支持向量机(MSSVM),能够实现正负类类间间隔的增大和正类类内体积的减小,但没有考虑负类体积的减小,即不能提高负类的聚类性,使之在对测试样本分类时还存在较大的混淆可能性.

本文建立一种新的模型,称之为最大间隔最小体积球形支持向量机(MMHSVM).运用 SVM 基本理论和 SVDD 超球结构,构造两个大小不同的同心超球,小超球将正类样本包裹其中,大超球将负类样本排斥在外.模型目标函数优化两超球间隔,同时实现类间间隔的增大和正负两类类内体积的缩小,更有利于分类器性能的提高.

## 2 最大间隔最小体积球形支持向量机

### 2.1 MSSVM

MSSVM 如图 1 所示,正类样本用 + 号表示,负类样本用 x 号表示.设两类样本为超球可分,存在两个同心超球  $S_1$  和  $S_2$ ,球心为  $c$ .  $S_1$  为小超球,  $S_2$  为大超球,  $S_1$  将正类样本包裹其中,  $S_2$  将负类样本排除其外.在超球  $S_2$  上任选一点  $A$  对超球  $S_1$  做切平面(在图 1 中对应为切线),切点为  $D$ ,连接  $AD$ ,  $Dc$  和  $Ac$ . 设  $Dc = R$ ,  $AD = d$ , 则  $Ac = \sqrt{R^2 + d^2}$ . 超球  $S_1$  的半径为  $R_1 = R$ , 超球  $S_2$  的半径为  $R_2 = \sqrt{R^2 + d^2}$ , 则正负类类间间隔为  $\Delta = R_2 - R_1 = \sqrt{R^2 + d^2} - R$ . MSSVM 的目标为最大化类间间隔  $\Delta$ , 等价于最大化  $d^2$  和最小化  $R^2$ , 并建立相应的数学模型和定义类判决规则.

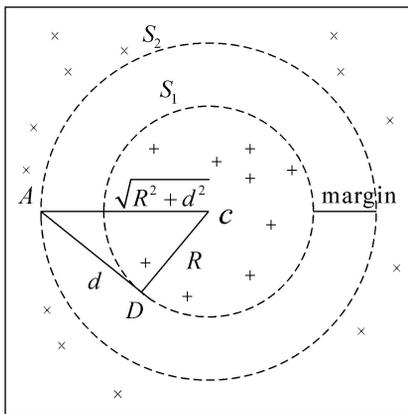


图 1 MSSVM

MSSVM 通过最大化  $d^2$  和最小化  $R^2$  实现分类间隔  $\Delta$  的最大化, 同时也可最小化正类类内体积. 但最大化  $d^2$  和最小化  $R^2$  并不能得到  $R_2 = \sqrt{R^2 + d^2}$  的最大化, 无法保证负类类内体积最小化, 不能确定模型对负类类内聚类性的贡献.

### 2.2 MMHSVM

MMHSVM 如图 2 所示, 设正负类样本  $\{x_i, y_i\}, i = 1, 2, \dots, n, x_i \in R^d, y_i \in \{-1, 1\}$ , 记正类训练样本为  $x_i^+$ , 正类训练样本数为  $n^+$ , 负类训练样本为  $x_j^-$ , 负类训练样本数为  $n^-$ . 模式样本在输入空间中为超球可分, 即存在两个同心超球  $S_1$  和  $S_2$ , 球心为  $c$ .  $S_1$  为小超球, 半径为  $R_1$ ,  $S_2$  为大超球, 半径为  $R_2, R_2 > R_1$ .  $S_1$  将正类样本包裹其中, 正类样本在超球体  $V_1$  中,  $S_1$  为  $V_1$  的边界;  $S_2$  将负类样本排除其外, 设  $S_2$  所对应的超球体为  $V_2$ , 则负类样本位于  $V_3 = \mathbb{R}^d \setminus V_2$  中,  $\mathbb{R}^d$  为输入空间.

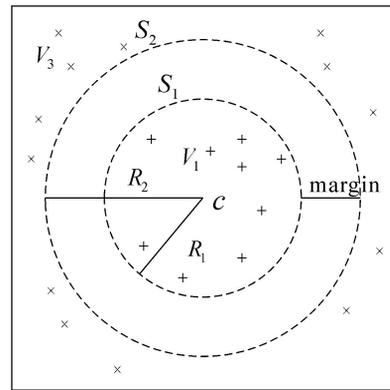


图 2 MMHSVM

正负类类间间隔  $\Delta = R_2 - R_1$ , 最大化类间间隔 等价于最大化  $R_2$  和最小化  $R_1$ , 亦即等价于最大化  $R_2^2$  和最小化  $R_1^2$ , 建立如下对应的数学模型:

$$\begin{aligned} \max_{R_1^2, R_2^2, c, \dagger, \ddagger, \ddot{}} & R_2^2 - MR_1^2 - C_1 \sum_i \dagger_i - C_2 \sum_j \ddot{j}; \quad (1) \\ \text{s.t.} & x_i^+ - c \quad^2 = (x_i^+ - c)^T (x_i^+ - c) \quad R_1^2 + \dagger_i, \\ & x_j^- - c \quad^2 = (x_j^- - c)^T (x_j^- - c) \quad R_2^2 - \ddot{j}, \\ & R_2^2 - R_1^2 \geq 0, \quad \dagger_i \geq 0, \quad \ddot{j} \geq 0, \quad \forall i, j. \end{aligned}$$

其中:  $\dagger_i$  和  $\ddot{j}$  为松弛因子, 分别用于约束正负类奇异点;  $C_1$  和  $C_2$  为惩罚因子, 用于控制模型训练精度和泛化性能的平衡; 参数  $M (M > 0, M \leq 1)$  用于  $R_2^2$  和  $R_1^2$  的折衷.

由式(1) 可得模型对应的 Lagrange 辅助函数

$$\begin{aligned} L(R_1^2, R_2^2, c, \dagger, \ddot{j}) = & R_2^2 - MR_1^2 - C_1 \sum_i \dagger_i - C_2 \sum_j \ddot{j} + \\ & \sum_i \dagger_i [R_1^2 + \dagger_i - (x_i^+ - c)^T (x_i^+ - c)] + \\ & \sum_j \ddot{j} \ddot{j} [R_2^2 - \ddot{j} - (x_j^- - c)^T (x_j^- - c) - \\ & R_2^2 + \ddot{j}] + \sum_j \ddot{j} \ddot{j} + (R_2^2 - R_1^2). \quad (2) \end{aligned}$$

其中:  $\dagger_i, \dagger_i, \ddot{j}, \ddot{j}$  为对应的非负 Lagrange 乘子. 令式(2) 分别对  $R_1^2, R_2^2, c, \dagger_i, \ddot{j}$  求偏导, 并令其

等于 0, 得到

$$\sum_i \alpha_i^+ = M + C_1, \quad (3)$$

$$\sum_j \alpha_j^- = 1 + C_2, \quad (4)$$

$$C_1 = \sum_i \alpha_i^+ + \sum_i \beta_i^+, \quad (5)$$

$$C_2 = \sum_j \alpha_j^- + \sum_j \beta_j^-, \quad (6)$$

$$\frac{\partial L}{\partial c} = 2 \sum_i \alpha_i^+ x_i^+ - 2c \sum_i \alpha_i^+ - 2 \sum_j \alpha_j^- x_j^- + 2c \sum_j \alpha_j^- = 0. \quad (7)$$

由式(3), (4) 和(7) 可得球心  $c$  为

$$c = \frac{1}{M - 1} \left( \sum_i \alpha_i^+ x_i^+ - \sum_j \alpha_j^- x_j^- \right). \quad (8)$$

由式(3) ~ (6) 可知

$$M - \sum_i \alpha_i^+ = C_1, \quad 1 - \sum_j \alpha_j^- = C_2. \quad (9)$$

将式(3) ~ (6) 和(8) 带入(2), 得到模型的 Wolf 对偶二次规划问题

$$\begin{aligned} \min - & \sum_i \alpha_i^+ (x_i^+ \cdot x_i^+) + \sum_j \alpha_j^- (x_j^- \cdot x_j^-) + \\ & \frac{1}{M - 1} \sum_{i,l} \alpha_i^+ \alpha_l^+ (x_i^+ \cdot x_l^+) - \\ & \frac{2}{M - 1} \sum_{i,j} \alpha_i^+ \alpha_j^- (x_i^+ \cdot x_j^-) + \\ & \frac{1}{M - 1} \sum_{j,k} \alpha_j^- \alpha_k^- (x_j^- \cdot x_k^-). \end{aligned} \quad (10)$$

$$\begin{aligned} \text{s. t. } & \sum_i \alpha_i^+ = M, 0 \leq \alpha_i^+ \leq C_1, \quad \forall i; \\ & \sum_j \alpha_j^- = 1, 0 \leq \alpha_j^- \leq C_2, \quad \forall j. \end{aligned}$$

**定理 1** 当  $M > 1$  时, 二次规划(10) 为凸二次规划, 即对偶二次规划问题的解为全局最优解.

求解对偶二次规划问题, 得到 Lagrange 乘子的解  $\alpha_i^+$  和  $\alpha_j^-$ , 由式(8) 可求得超球球心  $c$ .

任一样本  $x$  与球心  $c$  的距离平方公式

$$\begin{aligned} \|x - c\|^2 = & (x, x) - \frac{2}{M - 1} \left( \sum_i \alpha_i^+ (x \cdot x_i^+) - \sum_j \alpha_j^- (x \cdot x_j^-) \right) + \\ & \frac{1}{(M - 1)^2} \sum_{i,l} \alpha_i^+ \alpha_l^+ (x_i^+ \cdot x_l^+) - \frac{2}{(M - 1)^2} \sum_{i,j} \alpha_i^+ \alpha_j^- (x_i^+ \cdot x_j^-) + \\ & \frac{1}{(M - 1)^2} \sum_{j,k} \alpha_j^- \alpha_k^- (x_j^- \cdot x_k^-). \end{aligned} \quad (11)$$

由 KKT 条件和 MMHSVM 模型可知:

当  $0 < \alpha_i^+ < C_1$  时,  $\beta_i^+ = 0, R_1^2 = \|x_i^+ - c\|^2$ ,  $\alpha_i^+$  所对应的  $x_i^+$  为超球  $S_1$  的支持向量;

当  $0 < \alpha_j^- < C_2$  时,  $\beta_j^- = 0, R_2^2 = \|x_j^- - c\|^2$ ,  $\alpha_j^-$  所对应的  $x_j^-$  为超球  $S_2$  的支持向量;

当正负两类样本在输入空间中不为超球可分时, 可引入非线性映射, 将输入空间的模式样本映

射到高维甚至无限维特征空间  $H$  中, 实现超球可分. 而在特征空间  $H$  中, 两个向量  $(x_i), (x_j)$  的内积可利用 Mercer 核函数  $K(x_i, x_j)$  映射来表示和计算, 即有

$$K(x_i, x_j) = (x_i) \cdot (x_j).$$

分类判决规则:

1) 对于正负类二类分类问题, 定义分类半径  $R = R_1 + (R_2 - R_1)/2 = (R_2 + R_1)/2$ . 测试样本为  $x$ , 当  $\|x - c\| < R$  时, 判决  $x$  为正类, 反之  $x$  为负类.

2) 对于多类分类问题, 类似 SVM 的处理方法, 可采用一对多、一对一、决策树等方式将多类分类问题转换为多个二类分类问题进行类别判决.

### 3 模型比较分析

#### 3.1 分类性能比较分析

SVM 目标函数为  $\min(1/2) \|w\|^2$ , 类间间隔为  $\gamma = 2/\|w\|$ , 通过最大化类间间隔以提高分类器分类性能. SVDD 目标函数为  $\min R^2$ , 追求模式样本类内体积最小描述, 即最大化类内聚类性. MSSVM 目标函数为  $\min R^2 - Md^2$ , 等价于  $\min R^2$  和  $\max d^2$ , 则有  $\max d = \sqrt{R^2 + d^2} - R$ , 即最大化分类间隔; 当  $\min R^2$  时, 也可最小化正类类内体积; 但 MSSVM 模型不能对负类边界  $R_2 = \sqrt{R^2 + d^2}$  最大化, 无法提高负类类内聚类性.

MMHSVM 的目标函数  $\max R_2^2 - MR_1^2$ , 等价于  $\max R_2^2$  和  $\min R_1^2$ , 由图 2 可知,  $\gamma = R_2 - R_1$ , 即 MMHSVM 模型最大化类间间隔  $\gamma = R_2 - R_1$ ;  $\min R_1^2 \Leftrightarrow \min R_1 \Leftrightarrow \min V_1$ , 提高了正类类内聚类性;  $\min R_2^2 \Leftrightarrow \min R_2$ , 从直观上可知, 模型可以  $\min V_3$ , 但在绝对量度上, 无论  $R_2$  如何变化, 只要  $R_2$  为有限值, 则  $V_3$  的体积都为  $\pi^3/6$ , 无法类似  $(R_1, V_1)$  来说明  $(R_2, V_3)$  的变化关系.

为了严格说明  $(R_2, V_3)$  的对应变化关系, 通过同构的方式, 可将排斥于超球  $S_2$  外的映射空间  $V_3$ , 同构于一个超球体  $V_2$ , 其对应的超球面为  $S_2$ , 半径为  $R_2$ . 其中  $V_3 = \{z \mid d(z, c) > R_2, z \in R^n\}$ ,  $S_2 = \{z \mid d(z, c) = R_2, z \in R^n\}$ ,  $R_2 > 0$ ,  $d(z, c)$  表示点  $z$  到球心  $c$  的距离.

设  $z$  为超球  $S_2$  外任意一点, 即  $z \in V_3$ , 连接球心  $c$  与点  $z$ , 并向外延伸为射线  $l$ , 在射线  $l$  上选取一点  $z'$ , 使得  $d(z', c) = 1/d(z, c)$ . 点  $z$  与  $z'$  是一一对应的, 即存在双射映射  $f: z \leftrightarrow z'$ , 则  $V_3$  与  $V_2$  同构.  $V_2 = \{z \mid d(z, c) = 1/R_2 = R_2, z \in R^n\}$ , 即  $V_2$  为一超球体, 对应的超球面为

$$S_2 = \{z \mid d(z, c) = 1/R_2 = R_2, z \in R^n\}, \quad x_j^- \in V_2 \Leftrightarrow x_j^- \in V_2.$$

$\max R_2^2 \Leftrightarrow \min (R_2)^2, \max R_2 \Leftrightarrow \min R_2$ , 则  $\min V_2$ , 又  $V_3$  与  $V_2$  同构, 于是有  $\min V_3$ . 因此, MMHSVM 同时实现类间间隔的增大和正负类类内体积的减小, 等价于类间间隔的增大和类内聚类性的提高.

### 3.2 算法复杂度比较分析

SVM, SVDD, MSSVM, MMHSVM 的求解都归结于具有线性约束条件的二次凸规划问题, 线性约束二次凸规划问题的算法复杂度主要取决于规划中变量的个数和线性约束方程的个数. 采用符号  $Q(d, s)$  表示一个线性约束二次凸规划问题,  $C_Q(d, s)$  表示对应的复杂度, 其中  $d$  为变量个数,  $s$  为线性约束方程的个数. 设训练样本数为  $n$ , 则 SVM, SVDD, MSSVM, MMHSVM 算法的复杂度分别表示为  $C_Q(n, 2n+1)$ ,  $C_Q(n, 2n+1)$ ,  $C_Q(n, 2n+2)$ ,  $C_Q(n, 2n+2)$ . 由文献[9]知, SVM 的求解在时间和空间上的复杂度为  $O(n^2)$ , 则

$$C_Q(n, 2n+1) = O(n^2). \quad (12)$$

令式(12)中的  $n$  取值  $n+1$ , 则有

$$C_Q(n+1, 2n+3) = O((n+1)^2) = O(n^2), \quad (13)$$

显然有下式成立:

$$\begin{aligned} C_Q(n, 2n+1) &= C_Q(n, 2n+2) \\ C_Q(n+1, 2n+3) & \end{aligned} \quad (14)$$

由式(12)~(14)可知,  $C_Q(n, 2n+2) = O(n^2)$ , 各算法复杂度同级.

## 4 仿真实验

为了验证 MMHSVM 的有效性, 分别在日本女性表情数据库(JAFFE 静态图像库)和 UCI 机器学习数据库上进行了一系列的仿真实验, 并将 MMHSVM 与 SVM, MSSVM 算法进行测试比较.

**实验 1** 数据集为 JAFFE 静态表情图像库, 图像库总计 213 张图片, 7 种表情(愤怒、厌恶、恐惧、高兴、悲伤、惊讶和中性表情), 通过 10 个对象采集而成, 每人每种表情采集了 3 幅图像. 首先对原始图像进行人脸图像子区域分割, 截取出纯人脸表情图像, 并通过双线性插值算法和灰度直方图修正对人脸表情图像进行尺度和灰度归一化处理, 经过处理后的人脸图像为  $50 \times 50$  像素大小. 利用 Gabor 小波变换进行特征提取, 其中 Gabor 小波变换中的网格大小为  $5 \times 5^{[10]}$ . 分类器所选核函数均为高斯径向基核函数, 即

$$K(x, y) = \exp(- (r^{-1} \|x - y\|)^2), \quad r = 10.$$

SVM 的惩罚因子  $C = 10$ , MSSVM 的类决规则与 MMHSVM 类似, 取  $C_1 = C_2 = 10$ , 采用 one against rest 方式构造多分类器. 这里固定参数的目的在于只考虑算法的相对性能, 而不考虑其绝对性能. 采用

3 折交叉验证测试分类器的分类性能, 平均测试结果如表 1 所示, 其中 MS, MMS 分别为 MSSVM, MMHSVM 的简写. 从表 1 可知, 本文所提出的方法能够有效地提高分类器的分类能力.

表 1 各分类器在 JAFFE 数据集上的平均识别率 %

SVM	MS					MMS			
	M					M			
	1	1.5	2	2.5	3	1.5	2	2.5	3
生气	89.7	90.7	90.7	90.7	90.7	95	95	95	95
厌恶	84.7	79	77.3	75.7	75.7	83.3	83.3	83.3	81.7
害怕	79.6	87.3	87.3	87.3	85.7	84	88.3	88.3	88.3
高兴	91.3	89	89	89	84	84	95	95	95
中性	81.3	85.7	85.7	85.7	84	84	83.6	83.6	83.3
悲伤	84.6	89	89	89	89	89	93.3	93.3	93.3
惊讶	78	89	89	90.7	90.7	90	90	91.7	90
平均	84.1	87.1	86.9	86.9	85.7	85.4	89.8	89.8	90

表 1 中 MSSVM 的  $M$  为控制模型中  $d^2$  和  $R^2$  的折衷参数, 从平均识别率可知, 随着  $M$  的增大, MSSVM 的平均识别率有降低的趋势; MSSVM 的模型目标函数为  $\min R^2 - Md^2$ , 随着  $M$  的增大,  $R^2$  在目标函数中的变化率大于  $d^2$ , 模型类间间隔

$$= \sqrt{R^2 + d^2} - R = (d^2/2R)(1 + o(d/R)),$$

$R$  在类间间隔最大化中仅起到一次方的作用, 没有  $d$  在最大化过程的作用明显, 影响了的最大化.

表 1 中 MMHSVM 的  $M$  为控制模型中  $R_1^2$  和  $R_2^2$  的折衷参数, 模型的目标函数为  $\max R_2^2 - MR_1^2$ . 而模型所追求类间间隔  $= R_2 - R_1$  中,  $R_1$  和  $R_2$  是同次的,  $M$  的改变对类间间隔的最大化影响较小, 表现为实验中的平均识别率相对稳定.

**实验 2** 数据集来自 UCI 机器学习数据库, 它们是 iris, wine, parkinsons, 表 2 给出所选数据集特征. 各分类器核函数都选择高斯核函数, 采用 one against rest 方式构造分类器. 每个分类问题使用不同的核参数  $= [2^{-3}, 2^{-2.5}, \dots, 2^4]$  和折衷参数  $C = C_1 = C_2 = [2^1, 2^2, \dots, 2^{10}]$  估计推广能力; 使用 10 折交叉验证选择最优参数. 规范化训练数据为均值 0, 方差 1, 使用相同线性变化标度测试数据.

表 3 给出了 SVM, MSSVM, MMHSVM 在这 3

表 2 所选 UCI 数据集的数据特征

数据集	总样本数/个	类数	特征数
Iris	150	3	4
Wine	178	3	13
Parkinsons	195	2	22

表 3 各分类器在所选 UCI 数据集上的平均识别率

	%						
	SVM	MS			MMS		
	$r_{rate}$ ( $C, r$ )	$M = 1.5$ ( $C, r$ )	$M = 2$ ( $C, r$ )	$M = 2.5$ ( $C, r$ )	$M = 1.5$ ( $C, r$ )	$M = 2$ ( $C, r$ )	$M = 2.5$ ( $C, r$ )
Iris	94.7 ( $2^1, 2^{-3}$ )	95.7 ( $2^1, 2^{-1}$ )	94.3 ( $2^2, 2^{-1.5}$ )	93.7 ( $2^1, 2^{-1.5}$ )	94.9 ( $2^7, 2^{-1.5}$ )	95.3 ( $2^4, 2^{-1.5}$ )	95.3 ( $2^4, 2^{-1.5}$ )
Wine	94.4 ( $2^1, 2^{-3}$ )	95.0 ( $2^1, 2^1$ )	93.3 ( $2^1, 2^1$ )	93.3 ( $2^1, 2^1$ )	94.3 ( $2^1, 2^1$ )	94.1 ( $2^4, 2^1$ )	95.7 ( $2^1, 2^1$ )
Parkin- sons	87.5 ( $2^1, 2^{-1}$ )	86.5 ( $2^2, 2^{1.5}$ )	86.5 ( $2^2, 2^{1.5}$ )	86.5 ( $2^1, 2^{1.5}$ )	87.5 ( $2^1, 2^1$ )	87.5 ( $2^1, 2^1$ )	87.5 ( $2^1, 2^1$ )

个基准问题上的实验结果. 由表 3 可知, MMHSVM 表现出良好的推广能力.  $r_{rate}$  表示测试样本的正确率, ( $C, r$ ) 为对应核参数,  $M$  为模型参数.

### 5 结 论

增大类间间隔和提高类内聚类性是提高分类器分类性能的两种不同方式, SVM 关注于类间间隔的最大化; SVDD 追求一类样本的最小体积约束, 即类内聚类性最大化; MSSVM 在对类间间隔最大化时也可最小化正类类内体积, 但不能使得负类类内体积最小化. 本文提出的 MMHSVM 模型, 在对模型中两超球间隔即类间间隔进行最大化的同时, 实现了正负两类样本体积的最小化. 将提高分类性能的因素: 增大类间间隔和提高类内聚类性统一于同一模型中, 实现了模型分类能力的提高, 且模型的算法复杂度与其他算法相比是同级的. 公共数据库上的实验验证了算法的有效性.

### 参考文献 (References)

[1] Cortes C, Vapnik V N. Support vector networks[J]. Machine Learning, 1995, 20(3): 273-297.  
 [2] Guo B F, Gunn S R, Dampier R I, et al. Customizing kernel functions for SVM-based hyperspectral image classification[J]. IEEE Trans on Image Processing, 2008, 17(4): 622-629.  
 [3] 张钦礼, 王士同, 郭琦. 使用超椭球参数化坐标的支持向量机[J]. 控制与决策, 2008, 23(6): 626-630.  
 (Zhang Q L, Wang S T, Guo Q. Hyper-ellipsoid

support vector machine classifiers [J]. Control and Decision, 2008, 23(6): 626-630.)  
 [4] Tax D M J, Duin R P W. Support vector data description[J]. Machine Learning, 2004, 54(1): 45-66.  
 [5] Lin J, Peng M J. Interactive early warning technique based on SVDD [J]. J of Systems Engineering and Electronics, 2007, 18(3): 527-533.  
 [6] Camps-Valls G, Gomez-Choval L, Munoz-Mari J, et al. Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection [J]. IEEE Trans on Geoscience and Remote Sensing, 2008, 46(6): 1822-1835.  
 [7] Wang J G, Neskovic P, Cooper L N. Pattern classification via single spheres [J]. Lecture Notes in Computer Science, Discovery Science, 2005, 3735: 241-252.  
 [8] Hao P Y, Chiang J H, Lin Y H. A new maximal-margin spherical structured multi-class support vector machine [J]. Applied Intelligence, 2009, 30(2): 98-111.  
 [9] Collobert R, Bengio S. SVM Torch: Support vector machine for large-scale regression problems [J]. J of Machine Learning Research, 2001, 1(2): 143-160.  
 [10] Zhan Y Z, Ye J F, Niu D J, et al. Facial expression recognition based on gabor wavelet transformation and elastic templates matching [J]. Int J of Image and Graphics, 2006, 6(1): 125-138.