

核最近邻凸包分类算法

周晓飞 杨静宇 姜文瀚

(南京理工大学计算机科学与技术学院, 南京 210094)

摘要 为了增强最近邻凸包分类器的非线性分类能力,提出了基于核函数方法的最近邻凸包分类算法。该算法首先利用核函数方法将输入空间映射到高维特征空间,然后在高维特征空间采用最近邻凸包分类器对样本进行分类。最近邻凸包分类器是一类以测试点到各类别凸包的距离为相似性度量,并按最近邻原则归类的分类算法。人脸识别实验结果证实,这种核函数方法与最近邻凸包分类算法的融合是可行的和有效的。

关键词 核最近邻凸包分类 最近邻凸包分类 模式识别 人脸识别

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2007)07-1209-05

Kernel Nearest Neighbor Convex Hull Classification Algorithm

ZHOU Xiao-fei, YANG Jing-yu, JIANG Wen-han

(Department of Computer Science and Technology, Nanjing University of Science and Technology of China, Nanjing 210094)

Abstract A novel pattern classification algorithm based on the kernel method named kernel nearest neighbor convex hull (KNNCH) algorithm is presented in this paper. First, the data from the input space are projected into a higher dimensional feature space by replacing the inner product with an appropriately chosen kernel function. Then, the nearest neighbor convex hull (NNCH) classifier is constructed for classification in the higher dimension feature space. In NNCH classifier, the distance between a test sample and a convex hull of training samples of a class is taken as the similarity measure for classification. According to the nearest neighbor rule, a test sample will be classified to the class of the nearest convex hull. The experiments on face recognition show good performance of the combination of kernel method and NNCH classifier.

Keywords kernel nearest neighbor convex hull (KNNCH) classifier, nearest neighbor convex hull (NNCH) classifier, pattern recognition, face recognition

1 引言

核机器学习方法是核函数方法在机器学习领域的应用。其基本思想是先通过选择满足 Mercer 条件的适当核函数来替代特征空间中的样本内积,以便隐式地将输入空间非线性映射到高维特征空间;然后在高维特征空间中利用传统方法进行学习处理。该思想在机器学习中的成功应用最早源于支持向量机 (support vector machine, SVM)^[1,2],而核主分量分析 (kernel principal component analysis, KPCA)^[3]、核 Fisher 分类分析 (kernel Fisher discriminant analysis, KFDA)^[4]以及核独立成分分析

(kernel independent component analysis, KICA)^[5]等也都是这一思想在机器学习领域的应用。

核学习机在用于处理模式分类问题时,它的设计除了要考虑核函数的作用外,还要设计合理有效的分类算法。例如支持向量机,一方面通过引入核函数方法提高了非线性问题的学习处理能力;另一方面以结构风险最小化为原则的分类器设计本身也具有优良的学习泛化性能。本文提出了一种新的基于核函数方法的模式分类算法-核最近邻凸包分类算法(kernel nearest neighbor convex hull, KNNCH)。该方法综合考虑了以上两方面因素,即首先通过核函数方法隐式地将样本映射到高维特征空间;然后在高维特征空间利用最近邻凸包分类器(nearest neighbor convex hull,

基金项目:国家自然科学基金资助项目(60472060)

收稿日期:2006-03-28;改回日期:2006-06-07

第一作者简介:周晓飞(1978~),女。现为南京理工大学计算机科学与技术学院博士研究生。主要研究方向为人工智能与模式识别。

E-mail: zhouxif@njust.edu.cn

NNCH)进行分类。NNCH的设计源于SVM的几何解释^[6,7]。SVM的目的是要求取一个最优分类超平面使其两侧训练集中的点到它的最小距离(间隔)最大。而实质上,该最优分类超平面是两类训练集凸包最近点对连线的中垂面,它与训练样本之间的最大间隔也正是它与凸包之间的最大间隔。因此,笔者认为SVM实际上已隐含地将训练集扩展为两类训练样本凸包,并构造了最大间隔分离两类凸包的分类超平面。受此启发,最近邻凸包分类器的设计思路就形成了:该分类器以各类训练样本生成的凸包作为扩展训练集,并由凸包来估计样本分布。但是与SVM不同,该分类器采用了最近邻原则进行分类,即将测试样本归为最近邻凸包所属的类别。

本文通过人脸识别实验来对文中的核最近邻凸包分类算法进行测试,取得了较好的实验效果。结果证实了这种将核技术与最近邻凸包分类器有效结合的方法是可行的和有效的。

2 最近邻凸包分类算法

2.1 基本概念

定义 1(凸集定义)^[7]: 设集合 $S \subset \mathbf{R}^n$, 如果对任意 $x_1, x_2 \in S$ 和任意的 $\lambda \in [0, 1]$, 都有 $\lambda x_1 + (1 - \lambda)x_2 \in S$, 则称 S 是凸集。

定义 2(凸组合定义)^[7]: 设 $x_1, x_2, \dots, x_k \in \mathbf{R}^n$, 如果存在满足 $\sum_{i=1}^k \lambda_i = 1$ 的非负的 $\lambda_1, \lambda_2, \dots, \lambda_k$, 使得 $x = \sum_{i=1}^k \lambda_i x_i$, 则称 x 是 x_1, x_2, \dots, x_k 的一个凸组合。

定义 3(凸包定义)^[7]: 设集合 $S \subset \mathbf{R}^n$ 是由 \mathbf{R}^n 中的 K 个点组成的集合, 即 $S = \{x_1, \dots, x_k\}$, 则定义 S 的凸包为 $co(S) = \left\{ \sum_{i=1}^k \lambda_i x_i \mid \sum_{i=1}^k \lambda_i = 1, \lambda_i \geq 0 \right\}$ 。

定理 1(闭凸集投影定理)^[8]: 设 H 是 Hilbert 空间, K 是 H 中的一个闭凸集, $x \in H$ 是给定的向量, 则必存在唯一向量 $y \in K$, 使得对于 $\forall \eta \in K$ 有 $\|x - y\| \leq \|x - \eta\|$ 。此外, y 是 K 中唯一最小化向量的充分必要条件是, 对于 $\forall \eta \in K$, 都有 $(x - y) \cdot (\eta - y) \leq 0$ 。

本文定义点 x 到凸集 K 的距离为 $d(x, K) = \min_{\eta \in K} \|x - \eta\|$ 。

2.2 最近邻凸包分类算法

对一类训练样本集合 $S \subset \mathbf{R}^n, S = \{x_1, \dots, x_k\}$, S

的凸包为: $co(S) = \left\{ \sum_{i=1}^k \alpha_i x_i \mid \sum_{i=1}^k \alpha_i = 1, 0 \leq \alpha_i \leq 1 \right\}$ 。

对 \mathbf{R}^n 中的任意点 $x \in \mathbf{R}^n$ 到 $co(S)$ 的距离 $d(x, co(S))$ 为

$$d(x, co(S)) = \min_{\eta \in co(S)} \|x - \eta\| \quad (1)$$

由闭凸集投影定理知, 在 l_2 范数意义下, 点 x 在凸包上的投影点 y , 即式(1)的最优解, 且唯一存在。 y 是凸包 $co(S)$ 内与 x 最相似的点, 也是凸包对 x 的最佳逼近点。本文通过 x 与 y 之间的距离来衡量 x 与此类凸包所表征类别的相似程度, 距离越小越相似。

求式(1)优化问题等价于求 $d^2(x, co(S))$ 。因式(1)中 η 是凸包内的点, 其可由 S 中样本的凸组合表示, 即

$$\eta = \sum_{i=1}^k \alpha_i x_i, \sum_{i=1}^k \alpha_i = 1, 0 \leq \alpha_i \leq 1 \quad (2)$$

所以, 求 $d^2(x, co(S))$ 可写成以下优化问题:

$$\begin{aligned} d^2(x, co(S)) &= \min_{\eta \in co(S)} \|x - \eta\|^2 = \min_{\alpha_1, \dots, \alpha_k} \left\| x - \sum_{i=1}^k \alpha_i x_i \right\|^2 \\ &= \min_{\alpha_1, \dots, \alpha_k} \left[(x \cdot x) - 2 \sum_{i=1}^k \alpha_i (x \cdot x_i) + \sum_{i=1}^k \sum_{j=1}^k \alpha_i \alpha_j (x_i \cdot x_j) \right] \quad (3) \end{aligned}$$

$$\text{s. t. } \sum_{i=1}^k \alpha_i = 1, 0 \leq \alpha_i \leq 1, i = 1, 2, \dots, K$$

显然式(3)是一个凸二次规划问题。式(3)的最优解 $\hat{\alpha}_1, \dots, \hat{\alpha}_k$ 是 S 对 x 最优逼近的凸组合系数, 由于由该系数可求出 x 在 $co(S)$ 上的投影 $y = \sum_{i=1}^k \hat{\alpha}_i x_i$, 从而可求出 $d^2(x, co(S)) = \|x - y\|^2$ 。

对于 \mathbf{R}^n 中的 l 类问题 $\omega_1, \omega_2, \dots, \omega_l$, 有 l 类训练集 $S_1, S_2, \dots, S_l \subset \mathbf{R}^n, S_i$ 是第 i 类训练集。每类训练集分别生成各自类凸包 $co(S_1), co(S_2), \dots, co(S_l)$ 。对待鉴别样本 x 的判别分类问题, 就是先计算 x 到各凸包的距离平方 $d^2(x, co(S_1)), \dots, d^2(x, co(S_l))$ 的值, 然后依以下最近邻原则进行判别分类:

$$x \in \omega_j, j = \arg \min_i d^2(x, co(S_i)), i = 1, 2, \dots, l$$

当两类样本凸包交叠时(线性不可分), 由于交叠区域中的点到两类凸包的距离都为零, 所以 NNCH 对该区域中测试样本的分类是随机的或指定的。本文以图 1(a)和图 1(b)所示的 2 维空间中两种线性不可分情况为例来进行说明。图 1(a)和图 1(b)中, 凸包分别由虚线标出; NNCH 划分的不同类别区域分别由不同颜色表示(NNCH 对区域的划分并非直接可求, 而是隐含的结果)。

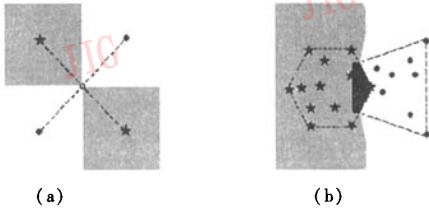


图 1 凸包交叠情况 NNCH 算法分类示意图

Fig. 1 NNCH classification in intersectional convex hull case

从图 1 可见,图 1(a)中,凸包交叠区域仅为一点,即图中两条虚线段的交点,该点为随机分类点;图 1(b)中,凸包交叠区域较大,即图中的深色阴影区,该区域的样本归类则存在随机性。

NNCH 与最近邻分类器(1-nearest neighbor, 1-NN)^[9]相比,都是通过比较距离,然后依最近邻原则进行分类的。但是,1-NN 比较的是测试点与训练点间的距离,而 NNCH 的比较对象则是测试点与凸包间的距离。

NNCH 与线性 SVM 相比,在处理线性可分问题时,虽然都是通过训练样本的凸包来构造分类器,但不同的是:线性 SVM 是线性的,而 NNCH 是非线性的。

另外,NNCH 与文献[10]的 k 局部凸距离最近邻分类算法(k -local convex distance nearest neighbor, CKNN)在相似性度量上都是以点到凸包的距离作为依据。但是,CKNN 是从 k 近邻角度出发的,它更强调样本流形的局部线性结构。与之相比,NNCH 则是从全局考虑,用一类训练样本凸包来估计其所属类别的样本分布。

最近邻凸包分类器有以下几方面特点:

(1) 非线性性 最近邻凸包分类器是以测试点到凸包的距离为相似性度量的分类器,其继承了最近邻分类器非线性的特点;

(2) 无参数性 和最近邻分类器一样,该分类器不需要预先设定任何经验参数即可实现模式的非线性分类;

(3) 样本虚拟性 测试点在各类别凸包上的投影点不一定是已知训练样本点,而各类别训练集则相当于被拓展为包括各自训练样本集的最小凸集,并由有限集变为无限集;

(4) 直接测试性 最近邻凸包分类器无需训练过程,就可直接进行实时测试;

(5) 多类适用性 算法本身对适用类别数没有特殊限制,完全适用于多类别问题的分类处理。

3 核最近邻凸包分类算法

最近邻凸包分类器用于处理如图 1(a)所示的一类交叠区域较小的线性不可分问题时,其决策的不确定性影响较小,因为严格地讲,各种分类器决策界上的点的分类也都是随机的,但当交叠区域过大(如图 1(b)所示)时,不确定区域问题就变得不可忽视。一个可采用的处理方法就是引入当前流行的核技术。该方法先通过核函数方法进行升维映射,以求将原空间线性不可分问题转化为高维空间(可能)线性可分(凸包分离)的问题;然后再用最近邻凸包分类器进行分类。核函数方法的升维映射是通过以满足 Mercer 条件的核函数来替代内积运算隐式地实现的。这种升维映射的合理性来自模式可分的 Cover 定理^[11,12],该定理可定性表述为“将复杂的模式分类问题非线性地投影到高维空间将比投影到低维空间更可能是线性可分的”。另外,由于该升维映射仅改变了内积运算,从而使得算法的复杂性并没有随维数的增加而增加。

3.1 核函数

统计学习理论指出,只要一种运算满足 Mercer 条件,那么它就可以作为 Hilbert 空间的内积使用^[13],即对于任意的对称函数 $k(x, y)$,它作为某个特征空间的内积运算的充分必要条件是,对 $\forall \varphi(x) \neq 0$ 且 $\int \varphi^2(x) dx < \infty$,有 $\int k(x, y) \varphi(x) \varphi(y) dx dy \geq 0$ 。

如果函数 $k(x, y)$ 满足 Mercer 条件,其中 $x, y \in F$,那么就意味着存在从空间 F 到某一 Hilbert 空间 H 的映射 $\Phi: F \rightarrow H, x \mapsto \Phi(x)$ 使得 $k(x, y) = (\Phi(x) \cdot \Phi(y))$,其中 (\cdot) 表示 H 中的内积。根据核函数的定义^[7], $k(x, y)$ 就是核函数。

尽管 Mercer 条件为核函数的判定提供了理论依据,然而,对于确定一个给定函数是否符合 Mercer 条件仍然是一件十分困难的事^[12]。关于核函数的选择、构造以及参数的设定等问题,目前都还处在研究阶段,在实际应用中,一般是通过合理利用先验知识来决定核函数(参数)的采用。以下是最常用的几种核函数:

线性核: $k(x, y) = x \cdot y$

径向基核: $k(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$ (σ 为尺度变量)

多项式核: $k(x, y) = (1 + (x \cdot y))^d$ ($d \geq 2$)

Sigmoid 核: $k(x, y) = \tanh(b(x \cdot y) - c)$ (b, c 为常数)。

3.2 核最近邻凸包分类算法

一个分类算法能否被推广到核空间,取决于它的训练过程与决策准则是否都能被表示为样本内积的形式。笔者看到,求点到凸包距离的优化方程式(式(3))完全可以写成样本(包括测试样本和训练样本)的内积形式,这就为核函数方法的应用提供了可行的途径。以下详细介绍该算法。

假设通过特征映射 $\Phi: \mathbf{R}^n \rightarrow H, \mathbf{x} \mapsto \Phi(\mathbf{x})$, 即可将训练集 $S \subset \mathbf{R}^n, S = \{x_1, \dots, x_k\}$ 所在的输入空间 \mathbf{R}^n 映射到特征空间 H , 那么, S 在空间 H 的映射集为 $\tilde{S} = \{\Phi(x_i) \mid \Phi: \mathbf{R}^n \mapsto H, x_i \in S, S \subset \mathbf{R}^n, i = 1, 2, \dots, K\}$, \tilde{S} 的凸包为

$$co(\tilde{S}) = \left\{ \sum_{i=1}^K \alpha_i \Phi(x_i) \mid \sum_{i=1}^K \alpha_i = 1, 0 \leq \alpha_i \leq 1 \right\} \quad (4)$$

对于 $\forall x \in \mathbf{R}^n$, 该点在空间 H 中的映射点为 $\Phi(x)$, $\Phi(x)$ 到 $co(\tilde{S})$ 的距离平方为

$$d^2(\Phi(x), co(\tilde{S})) = \min_{\eta \in co(\tilde{S})} \|\Phi(x) - \eta\|^2 \quad (5)$$

同样, η 可由 \tilde{S} 中样本的凸组合表示, 即

$$\eta = \sum_{i=1}^K \alpha_i \Phi(x_i), \sum_{i=1}^K \alpha_i = 1, 0 \leq \alpha_i \leq 1 \quad (6)$$

因此, 求 $\Phi(x)$ 到 $co(\tilde{S})$ 的距离平方可写成以下优化问题:

$$\begin{aligned} d^2(\Phi(x), co(\tilde{S})) &= \min_{\eta \in co(\tilde{S})} \|\Phi(x) - \eta\|^2 \\ &= \min_{\alpha_1, \dots, \alpha_K} \left\| \Phi(x) - \sum_{i=1}^K \alpha_i \Phi(x_i) \right\|^2 \\ &= \min_{\alpha_1, \dots, \alpha_K} \left[(\Phi(x) \cdot \Phi(x)) - 2 \sum_{i=1}^K \alpha_i (\Phi(x) \cdot \Phi(x_i)) + \sum_{i=1}^K \sum_{j=1}^K \alpha_i \alpha_j (\Phi(x_i) \cdot \Phi(x_j)) \right] \\ \text{s. t. } &\sum_{i=1}^K \alpha_i = 1, 0 \leq \alpha_i \leq 1, i = 1, 2, \dots, K \quad (7) \end{aligned}$$

若选择满足 Mercer 条件的适当的核函数 $k(x, y)$ 替换空间 H 中的内积 $(\Phi(x) \cdot \Phi(y))$, 则式(7)可表示为如下与核函数有关的形式:

$$\begin{aligned} d^2(\Phi(x), co(\tilde{S})) &= \min_{\alpha_1, \dots, \alpha_K} \left[k(x, x) - 2 \sum_{i=1}^K \alpha_i k(x, x_i) + \sum_{i=1}^K \sum_{j=1}^K \alpha_i \alpha_j k(x_i, x_j) \right] \\ \text{s. t. } &\sum_{i=1}^K \alpha_i = 1, 0 \leq \alpha_i \leq 1, i = 1, 2, \dots, K \quad (8) \end{aligned}$$

式(8)是一个凸二次规划优化问题, 其与式(3)

原理相同, 只是内积运算不同。

同样, 对于 \mathbf{R}^n 中的 l 类问题, 在特征空间 H 中相应就有 l 个类别的凸包生成。而判断一个测试样本的类别, 就是先通过优化式(8)分别求得 l 个距离值, 然后再依据最近邻原则归类。

4 实验和分析

为了验证 KNNCH 分类算法的有效性, 本文选择 AT&T 剑桥实验室 (AT&T laboratories cambridge) 的 ORL (olivetti research lab) 标准人脸图像库 (<http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html>) 中的全部人脸图像对该算法与 1-NN 和 SVM 进行实验比较。ORL 标准人脸库由 40 个人的 400 幅灰度图像构成。其中, 每人 10 幅图像, 分别用自然数 1, 2, ..., 10 标识。图像格式为 PGM, 大小为 92×112 。该库中的人脸图像摄于不同时期, 并具有光照、表情 (如睁眼, 闭眼, 微笑, 不笑)、细节 (如戴眼镜, 不戴眼镜)、姿态 (深度旋转和平面旋转可达 20°) 和尺度等的变化。图 2 为 ORL 标准人脸库中的一类人脸例图。

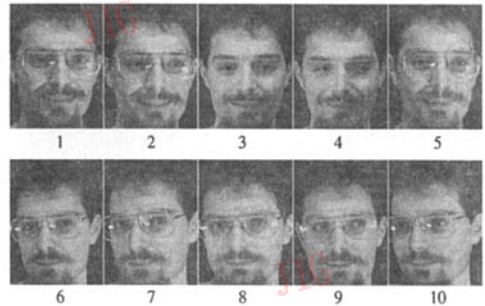


图 2 ORL 人脸库图像

Fig. 2 Face images from ORL face database

本文的实验目的是要在同等实验条件下, 通过比较 3 种分类算法的识别率来证实 KNNCH 分类算法的有效性和可行性。因此实验没有对图像进行特殊的特征提取处理, 只将每幅图像转化为 JPEG 格式, 通过双三次插值缩小为 16×16 大小, 并按列展成 1 维样本向量。实验分别取所有人的前 6 幅、前 7 幅、后 4 幅、后 3 幅图像以及相应的镜像图像用于训练, 其余的图像 (分别为后 4 幅、后 3 幅、前 6 幅、前 7 幅图像) 作为测试集, 共进行了 4 组比较实验。实验中, SVM 和 KNNCH 的优化都采用 Matlab 优化工具箱 quadprog 实现。对于核函数的选择, 本实验

采用线性核 (Linear), 径向基核 (Rbf) 和多项式核 (Poly) 3 种核函数。其中, 径向基核参数 $\sigma = 6$, 多项式核参数 $d = 2$ 。对于支持向量机的惩罚参数, 本实验取 $C = \infty$ 。实验结果见表 1。

表 1 ORL 人脸库的不同方法的正确识别率

Tab. 1 Results of experiments on ORL face database

训练集 $i \sim j$	测试集 $i \sim j$	正确识别率 (%)						
		1-NN	SVM ($C = \infty$)			KNNCH		
			Linear	Rbf(6)	Poly(2)	Linear	Rbf(6)	Poly(2)
1 ~ 6	7 ~ 10	96.25	96.25	96.25	96.25	96.25	96.25	97.50
1 ~ 7	8 ~ 10	97.50	97.50	96.67	97.50	97.50	97.50	97.50
7 ~ 10	1 ~ 6	91.25	92.08	93.33	92.92	93.75	93.75	94.17
8 ~ 10	1 ~ 7	88.93	91.43	91.43	92.14	91.43	91.43	92.50

注: $i \sim j$ 代表从第 i 个到第 j 个样本 (如 1 ~ 6, 前 6 个样本), Linear 代表线性核, Rbf(6) 代表选用径向基核 ($\sigma = 6$), Poly 代表选用多项式核 ($d = 2$)。黑体表示各组实验的最高识别率。

从表 1 实验结果可以看到, 多项式核的 KNNCH 的识别效果明显好于 1-NN 和 SVM。4 组实验中, 多项式核的 KNNCH 除第 2 组实验的正确识别率 (97.50%) 与 1-NN 和 (Linear, Poly) SVM 相同外, 其它 3 组实验的正确识别率 (97.50%, 94.17% 和 92.50%) 均高于 1-NN 和 SVM 的正确识别率。

此外, 通过 3 种不同核函数的 KNNCH 分类器结果的比较可看到, 多项式核的分类结果明显要好于线性核和径向基核, 而线性核方法就是原来的 NNCH。可见, 由于多项式核函数的引入, 使得核最近邻凸包分类器比最近邻凸包分类器具有更好的非线性分类能力。

5 结 论

本文提出的核最近邻凸包分类算法是将核函数方法与最近邻凸包分类器有效融合的非线性模式分类方法。该算法兼有核函数方法和最近邻凸包分类器的诸多长处。ORL 人脸数据库上的实验结果表明, 这一算法具有良好的分类能力, 也证实了这种融合策略对处理模式分类问题是可行的和有效的。

参考文献 (References)

- 1 Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers [A]. In: Proceedings of the 5th Annual Workshop on Computational Learning Theory [C], Pittsburgh, PA, USA, 1992: 144 ~ 152.
- 2 Vapnik V N. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1995.
- 3 Schölkopf B, Smola A, Müller K R. Kernel principal component analysis [A]. In: Proceedings of Artificial Neural Networks-ICANN'97 [C], Lausanne, Switzerland, 1997: 583 ~ 588.
- 4 Mika S, Rätsch G, Weston J, et al. Fisher discriminant analysis with kernels [A]. In: Proceedings of IEEE Neural Networks for Signal Processing Workshop [C], Madison, Wisconsin, USA, 1999: 41 ~ 48.
- 5 Bach F, Jordan M I. Kernel independent component analysis [R]. Technical Report CSD-01-1166, Computer Science Division, University of California, Berkeley, CA, USA, 2001.
- 6 Bennett K P, Bredensteiner E J. Duality and geometry in SVM classifiers [A]. In: Proceedings of Seventeenth International Conference on Machine Learning [C], Stanford University, CA, USA, 2000: 57 ~ 64.
- 7 Deng Nai-yang, Tian Ying-jie. A New Approach in Data Mine—Support Vector Machine [M]. Beijing: Science Press, 2004. [邓乃扬, 田英杰著. 数据挖掘中的新方法——支持向量机 [M]. 北京: 科学出版社, 2004.]
- 8 Wang Ri-shuang. Functional Analysis and Optimization Theory [M]. Beijing: Bei Hang University Press, 2003. [王日爽编著. 泛函分析与优化理论 [M]. 北京: 北京航空航天大学出版社, 2003.]
- 9 Cover T M, Hart P E. Nearest neighbor pattern classification [J]. IEEE Transactions on Information Theory, 1967, 13(1): 21 ~ 27.
- 10 Vincent P, Bengio Y. K-local hyperplane and convex distance nearest neighbor algorithms [A]. In: Dietterich T G, Becker S, Ghahramani Z, editors. Advances in Neural Information Processing Systems 14 [C], Cambridge, MA, USA: MIT Press, 2002: 985 ~ 992.
- 11 Cover T M. Geometrical and statistical properties of systems of linear inequalities with application in pattern recognition [J]. IEEE Transactions on Electronic Computers, 1965, EC-14: 326 ~ 334.
- 12 Haykin S. Neural Networks: A Comprehensive Foundation, 2nd Edition (in Chinese) [M]. Beijing: China Machine Press, 2004. [Haykin S 著, 叶世伟, 史忠植译. 神经网络原理 (原书第 2 版) [M]. 北京: 机械工业出版社, 2004.]
- 13 Bian Zhao-qi, Zhang Xue-gong. Pattern Recognition (2nd edition) [M]. Beijing: Tsinghua University Press, 1999. [边肇祺, 张学工等编著. 模式识别 (第二版) [M]. 北京: 清华大学出版社, 1999.]