

# 一种新颖的核学习算法用于小波特征的人脸识别<sup>\*</sup>)

周晓飞 姜文瀚 杨静宇

(南京理工大学计算机科学与技术学院 南京 210094)

**摘要** 文中将一种新颖的核学习算法—核最近邻凸包分类算法用于人脸的小波特征识别。该算法的设计受到支持向量机几何解释启发,利用核函数方法将数据映射到高维核空间,并在核空间构造以训练集凸包为扩展类集的最近邻分类器。文中采用的人脸图像的小波低频特征对人脸识别十分有效。人脸的小波低频特征不但保留了人脸的主要信息,而且具有较少的维度。在 ORL 人脸图像库上的“leave-one-out”测试方法的实验中,这种基于小波低频特征的核最近邻凸包分类算法取得了 99.25% 的识别率。

**关键词** 核学习,核最近邻凸包,小波变换,模式识别,人脸识别

## A New Kernel Learning Algorithm for Wavelet Features-based Face Recognition

ZHOU Xiao-Fei JIANG Wen-Han YANG Jing-Yu

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

**Abstract** A new kernel learning method called Kernel Nearest Neighbor Convex hull (KNNCH) algorithm is used for wavelet features based face recognition. Inspired by the intuitive geometric interpretation of SVM based on convex hulls, KNNCH maps the data in the original space to the kernel space with the kernel trick and constructs a nearest neighbor classifier in the kernel space, which takes the convex hulls of training sets as the extended class sets. The lower frequency features of face images extracted by 2D wavelet transform are efficient for face recognition. The features not only preserve the main information of face images, but also have the less dimensionality. KNNCH with wavelet features for face recognition shows very good performance, which can achieve 99.25% recognition rate with “leave-one-out” test method on ORL face database.

**Keywords** Kernel learning, Kernel nearest neighbor convex hull, Wavelet transform, Pattern recognition, Face recognition

## 1 引言

核机器学习方法是建立在统计学习理论和核函数方法基础上的一类以支持向量机为核心的统计机器学习方法,其基本思想是通过选择满足 Mercer 条件<sup>[1]</sup>的适当核函数替代特征空间中的样本内积,从而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。自支持向量机(Support Vector Machine, SVM)<sup>[2,3]</sup>成功引入核函数方法以来,许多基于核的学习算法相继被提出。核主分量分析(Kernel Principal Component Analysis, KPCA)<sup>[4]</sup>、核 Fisher 分类分析(Kernel Fisher Discriminant Analysis, KFDA)<sup>[5]</sup>以及核独立成分分析(Kernel Independent Component Analysis, KICA)<sup>[6]</sup>等都是核函数方法在机器学习中的推广。

本文的核学习算法—核最近邻凸包分类算法(Kernel Nearest Neighbor Convex Hull, KNNCH)受到支持向量机的几何解释<sup>[7-9]</sup>启发,采用核函数方法将以训练样本凸包为扩展训练集的最近邻分类器推广到核空间。该算法与二维小波变换(Wavelet Transform, WT)结合用于人脸图像识别,在 ORL 人脸库上取得了较好的实验效果。

本文接下来的第 2 部分简要介绍人脸图像的二维小波变

换,第 3 部分详细介绍核最近邻凸包分类算法,第 4 部分是实验分析,最后是总结。

## 2 人脸图像的二维小波变换

模式识别的重要环节就是特征提取,人脸识别也不例外。经过小波变换提取的人脸图像低频特征在一定程度上模糊了人脸的细节特征。这使得同类人脸间的相似性得以增强,同时也不失主要特征地保留了不同类人脸间的差异信息。此外,小波变换实现的特征维数压缩,客观上也极大地降低了识别处理的计算复杂度。以下是关于图像二维小波变换理论的简要介绍。

在二维多分辨率分析(Multi-Resolution Analysis, MRA)<sup>[10-13]</sup>中,存在一个二维尺度函数  $\Phi(x, y) = \varphi(x)\varphi(y)$ , 和三个二维小波函数  $\Psi_1(x, y) = \varphi(x)\psi(y)$ ,  $\Psi_2(x, y) = \psi(x)\varphi(y)$ ,  $\Psi_3(x, y) = \psi(x)\psi(y)$ 。其中  $\varphi(x)$  和  $\psi(x)$  分别是一维 MRA 中的尺度函数和小波函数。 $\varphi(t)$  和  $\psi(t)$  满足尺度方程:

$$\begin{cases} \varphi(t) = \sum_{k \in Z} h(k)\varphi(2t-k) \\ \psi(t) = \sum_{k \in Z} g(k)\varphi(2t-k) \end{cases}, g(k) = (-1)^k h(1-k) \quad (1)$$

$\{h(k)\}$  和  $\{g(k)\}$  分别是低通和高通数字滤波器。对于二维图像小波分解,就是竖直和水平的高通和低通算子  $\{h(k)\}$ 、

<sup>\*</sup> 基金项目:国家自然科学基金资助项目(No. 60472060)。周晓飞 博士研究生,主要研究方向为人工智能与模式识别;姜文瀚 博士研究生,主要研究方向为人工智能与模式识别;杨静宇 教授,博士生导师,主要研究方向为计算机视觉,机器学习,智能机器人与模式识别等。

$\{g(k)\}$ 先后作用于图像的行和列的过程。所以有四种滤波算子分别为：竖直低通水平低通滤波(LL)、竖直高通水平低通滤波(HL)、竖直低通水平高通滤波(LH)和竖直高通水平高通滤波(HH)。对给出  $M+1$  尺度上的大小为  $m \times n$  的二维图像函数(离散)  $f(m, n)$ , 计算可得二维 Mallat 小波分解如下:

$$\begin{cases} c_{m,n}^{M+1} = f(m, n) \\ \alpha_{m,n}^M = \sum_{k,l} h(k-2m)h(l-2n)c_{k,l}^{M+1} \\ \beta_{m,n}^M = \sum_{k,l} h(k-2m)g(l-2n)c_{k,l}^{M+1} \\ \gamma_{m,n}^M = \sum_{k,l} g(k-2m)h(l-2n)c_{k,l}^{M+1} \\ \delta_{m,n}^M = \sum_{k,l} g(k-2m)g(l-2n)c_{k,l}^{M+1} \end{cases} \quad (2)$$

小波系数  $\{c_{m,n}^M\}$  表现了图像的低频信息, 反映了图像平缓变化, 其视觉感觉明显。而系数  $\{\alpha_{m,n}^M\}$ ,  $\{\beta_{m,n}^M\}$  和  $\{\gamma_{m,n}^M\}$  的值往往都很小, 分别体现了竖直方向、水平方向和对角线方向的边界奇异性, 反映了图像的边缘和纹理等细节信息。小波分解后系数图像的结构如图 1 所示, 包括子图像 LL, HL, LH 和 HH。其中 LL 为低频系数  $\{c_{m,n}^M\}$ , HL, LH 和 HH 分别为系数  $\{\alpha_{m,n}^M\}$ ,  $\{\beta_{m,n}^M\}$  和  $\{\gamma_{m,n}^M\}$ 。



图 1 二维小波变换结构示意图

我们以人脸图像的小波变换实例来说明小波系数特点。图 2(a)和图 2(b)分别为同一人的两幅不同人脸图像的一级 Daubechies 1(db1)小波分解<sup>[14]</sup> 例图。



图 2 二维小波变换的一级分解实例图示意图

可以看到, 图 2(a)和图 2(b)中小波分解后的低频系数 LL 仍然与该类人脸图像近似, 并且脸部的细节特征也得到了-定程度的弱化, 突出表现在眼镜遮掩的滤除。这种保留主要信息, 弱化细节变化的处理, 对于人脸模式的分类是有利的。同时 LL 部分的维数仅是原图像维数的 1/4。维度的减少有效地降低了计算复杂度。

### 3 核最近邻凸包分类算法

支持向量机在处理线性不可分问题时, 通过巧妙运用核函数方法将输入空间中的样本非线性的升维映射到高维空间, 使原本线性不可分的问题转化为高维空间(可能)线性可分的情况(升维合理性有 cover 定理<sup>[15]</sup>), 然后再在高维空间运用线性手段构造最优分类超平面。从几何角度解释<sup>[7-9]</sup>, 高维空间中的最优分类超平面实质就是两类训练特征样本凸包最近点对线段的中垂面。所以, 支持向量机所表达的“最大间隔”也正是两个凸包间的距离(最近点对距离)。因此, 我们可以这样理解: SVM 算法已隐含地将训练特征集扩展为凸

包, 并构造最大间隔分离凸包的分类超平面。受此启发, 我们也由凸包描述样本分布, 从分离凸包的角度设计分类器。因此, 我们提出核最近邻凸包分类算法, 该算法在核空间以凸包为扩展训练集, 将测试特征样本归属于离其最近的凸包类。以下详细介绍该算法。

假设有  $l$  个类别  $\omega_1, \omega_2, \dots, \omega_l$  的模式识别问题, 各类训练集  $S_1, S_2, \dots, S_l \subset R^n$  中分别包含有  $m_i$  个训练样本:  $S_i = \{x_i^1, x_i^2, \dots, x_i^{m_i}\}, i=1, 2, \dots, l$ 。训练样本  $x_i^j$  的上角标  $i$  表示类别  $\omega_i$ , 下角标  $j$  表示  $S_i$  中的第  $j$  个样本。如果存在特征映射  $\Phi: R^n \rightarrow H$ , 那么各训练集在  $H$  空间的映射集为:  $\tilde{S}_i = \{\Phi(x_i^j) | x_i^j \in S_i, S_i \subset R^n, j=1, 2, \dots, m_i, \Phi: R^n \rightarrow H\}, i=1, 2, \dots, l$ 。我们称  $\tilde{S}_i (i=1, 2, \dots, l)$  为  $\omega_i$  类的训练特征集, 则训练特征集  $\tilde{S}_i (i=1, 2, \dots, l)$  的凸包  $co(\tilde{S}_i)$  为:

$$co(\tilde{S}_i) = \left\{ \sum_{j=1}^{m_i} \alpha_j \Phi(x_i^j) \mid \sum_{j=1}^{m_i} \alpha_j = 1, 0 \leq \alpha_j \leq 1, i=1, 2, \dots, l \right\} \quad (3)$$

若训练集  $S_1, S_2, \dots, S_l$  通过高维映射  $\Phi$  投影到  $H$  空间后的训练特征集  $\tilde{S}_1, \tilde{S}_2, \dots, \tilde{S}_l$  彼此线性可分, 那么, 凸包  $co(\tilde{S}_1), co(\tilde{S}_2), \dots, co(\tilde{S}_l)$  彼此必不相交(分离)<sup>[9]</sup>。在高维  $H$  空间, 我们认为各类特征样本分布在各自训练特征集凸包周围, 以这些凸包为扩展类集, 比较测试点到各凸包的-距离, 按最近邻原则将测试点归为最近邻的凸包类。

对于给定  $R^n$  中待鉴别样本  $x, x$  在高维空间  $H$  的投影为  $\Phi(x)$ , 计算  $\Phi(x)$  到各类训练特征集凸包的-距离  $d(\Phi(x), co(\tilde{S}_1)), \dots, d(\Phi(x), co(\tilde{S}_2)), \dots, d(\Phi(x), co(\tilde{S}_l))$  为(求点到凸集的最小距离<sup>[16]</sup>):

$$d(\Phi(x), co(\tilde{S}_i)) = \min_{\eta \in co(\tilde{S}_i)} \|\Phi(x) - \eta\|, i=1, 2, \dots, l. \quad (4)$$

由闭凸集投影定理<sup>[9,16,17]</sup> 可知,  $\Phi(x)$  在凸集  $co(\tilde{S}_i)$  内必存在唯一的向量  $\eta^*$  使得  $\|\Phi(x) - \eta^*\| \leq \|\Phi(x) - \eta\|, \eta \in co(\tilde{S}_i)$ 。  $\eta^*$  是凸包  $co(\tilde{S}_i)$  内与  $\Phi(x)$  最相似的点, 是凸包  $co(\tilde{S}_i)$  对  $\Phi(x)$  的最佳逼近元。我们通过  $\Phi(x)$  与  $\eta^* (i=1, 2, \dots, l)$  之间的距离  $d(\Phi(x), co(\tilde{S}_i)) = \|\Phi(x) - \eta^*\|$  来衡量  $\Phi(x)$  与  $\omega_i (i=1, 2, \dots, l)$  类凸包的相似程度, 距离越小越相似。

在  $l_2$  范数意义下, 由  $d(\Phi(x), co(\tilde{S}_i))$  作为相似性度量与  $d^2(\Phi(x), co(\tilde{S}_i)) = \min_{\eta \in co(\tilde{S}_i)} \|\Phi(x) - \eta\|^2 (i=1, 2, \dots, l)$  等价。又因  $\eta$  是凸包  $co(\tilde{S}_i)$  内的点,  $\eta$  可由  $\tilde{S}_i$  中样本的凸组合表示:  $\eta = \sum_{j=1}^{m_i} \alpha_j \Phi(x_i^j), \sum_{j=1}^{m_i} \alpha_j = 1, 0 \leq \alpha_j \leq 1$ , 因此求凸包最佳逼近元  $\eta^*$  的问题就是求凸组合权系数  $\alpha_1, \alpha_2, \dots, \alpha_{m_i}$  的最佳值问题:

$$\begin{cases} \text{令 } \alpha = (\alpha_1, \alpha_2, \dots, \alpha_{m_i}), \\ d^2(\Phi(x), co(\tilde{S}_i)) = \min_{\eta \in co(\tilde{S}_i)} \|\Phi(x) - \eta\|^2 \\ = \min_{\alpha} \left\| \Phi(x) - \sum_{j=1}^{m_i} \alpha_j \Phi(x_i^j) \right\|^2 \\ = \min_{\alpha} \left[ \Phi(x) \cdot \Phi(x) - 2 \sum_{j=1}^{m_i} \alpha_j \Phi(x) \cdot \Phi(x_i^j) \right. \\ \left. + \sum_{j=1}^{m_i} \sum_{k=1}^{m_i} \alpha_j \alpha_k \Phi(x_i^j) \cdot \Phi(x_i^k) \right] \\ \text{s. t. } \sum_{j=1}^{m_i} \alpha_j = 1, 0 \leq \alpha_j \leq 1, j=1, 2, \dots, m_i \\ i=1, 2, \dots, l. \end{cases} \quad (5)$$

我们看到求测试样本点到凸包距离平方的优化方程(5)

完全可以写成已知样本(包括测试样本和训练样本)在  $H$  空间的内积形式。所以可以利用核函数方法,根据 Mercer 条件<sup>[1]</sup>选择适当的核函数  $k(x, y)$  替代  $H$  空间的内积运算  $\Phi(x) \cdot \Phi(y)$ 。则(5)式可表示为如下只与核函数有关的形式:

$$d^2(\Phi(x), co(\tilde{S}_i)) = \min_a k(x, x) - 2 \sum_{j=1}^{m_i} a_j k(x, x_j) + \sum_{j=1}^{m_i} \sum_{k=1}^{m_i} a_j a_k k(x_j, x_k)$$

$$s. t. \sum_{j=1}^{m_i} a_j = 1, 0 \leq a_j \leq 1, j=1, 2, \dots, m_i$$

$$i=1, 2, \dots, l. \quad (6)$$

(6)式是凸二次规划优化问题,我们采用 Matlab 优化工具箱求解。求出  $d^2(\Phi(x), co(\tilde{S}_i)) (i=1, 2, \dots, l)$  后,我们依最近邻原则判别分类:

$$x \in \omega_j, j = \arg \min_i d^2(\Phi(x), co(\tilde{S}_i)), i=1, 2, \dots, l. \quad (7)$$

综上所述,核最近邻凸包分类算法不需要训练过程,也不必直接求出决策界,但是该算法在核空间所隐含的决策界是将凸包非线性分离的。如图 3,我们以二维空间为例,比较支持向量机和核最近邻凸包两种分类器的决策差异。

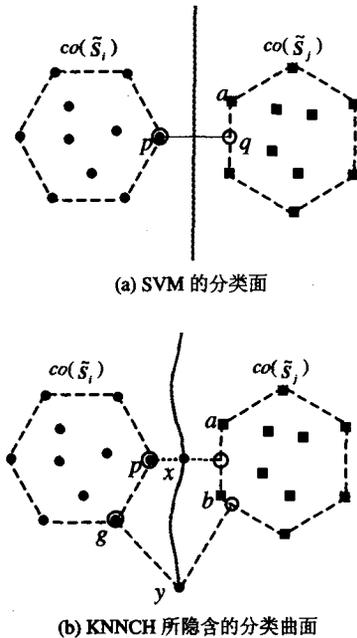


图 3 核空间中 SVM 和 KNNCH 的决策界示意图

由图 3 显而易见,在核空间,支持向量机的决策界是线性的,而 KNNCH 所隐含的决策界是非线性的。从几何解释,支持向量机在核空间的线性决策界由凸包最近点对  $p, q$  决定,即分类面可直接通过垂直平分最近点对来构造。而最近点对又分别是各类支持向量(Support Vectors, SV)<sup>[3]</sup>的凸组合,如图 3(a)中,  $p$  也是支持向量  $p = \lambda p, \lambda=1, q$  落在支持向量  $a, b$  之间  $q = \lambda_1 a + \lambda_2 b, \lambda_1 \lambda_2 \in [0, 1], \lambda_1 + \lambda_2 = 1$ 。所以,最终 SVM 分类只与支持向量有关。而核最近邻凸包分类器所隐含的决策界不仅只与支持向量有关,还与凸包上的其他边缘训练样本有关。如图 3(b)的决策界上的点  $x$  和  $y$  到两侧凸包距离相等,  $x$  在两侧凸包上的投影点分别落在支持向量  $p$  上和  $a, b$  间,而  $y$  的投影点却分别落在边缘样本  $g$  上和  $b, c$  之间。

另外,核最近邻凸包分类器不同于支持向量机仅针对两类问题,可直接用于多类别问题。

## 4 实验和分析

我们选择在 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>)上将本文提出的小波变换与核最近邻凸包分类相结合的算法(WT+KNNCH)同最近邻分类器(1NN)、核支持向量机(径向基核)以及核最近邻凸包分类器(KNNCH)进行比较。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,分辨率是  $92 \times 112$ 。该库中的人脸图像摄于不同时期,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  $20^\circ$ )和尺度等的变化。图 4 为 ORL 标准人脸库中的人脸例图。实验中,将所有图像先预处理转化为 JPEG 格式,双三次插值缩小为  $16 \times 16$ 。

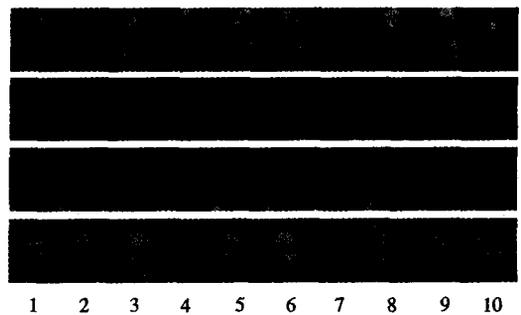


图 4 ORL 人脸库图像

本文算法首先采用 db1 小波对图像进行 1 级二维小波分解,将低频图像作为人脸分类特征,然后再用 KNNCH 分类器分类。实验中,各种分类器都是将特征图像按列展成一维向量作为特征样本。我们采用“leave-one-out”方法进行实验测试。该方法每次留出一个样本测试,其余的 399 个样本用来训练,所有样本都分别测试一次,共测试 400 次。实验分别统计各分类算法的正确识别率(正确识别次数/400)。SVM 和 KNNCH 分类器的优化采用 Matlab 优化工具箱的 quadprog 实现。对于 SVM, KNNCH 以及 WT+KNNCH 的核函数,我们都选择径向基核。核参数根据经验设置,选择使各算法识别率最佳的参数。SVM 和 KNNCH 的参数为 2, WT+KNNCH 的核参数为 1。实验结果见表 1。

表 1 ORL 人脸库的实验结果

NN	SVM	KNNCH	WT+KNNCH
98.25%	98.75%	98.75%	99.25%

实验结果可见,WT+KNNCH 取得了最高的识别率 99.25%,较 1NN, SVM 和未经过小波处理的 KNNCH 具有明显的优势。SVM 和 KNNCH 结果相同为 98.75%,都要低于 WT+KNNCH,而高于 1NN 的 98.25%。由此可见,本文提出的基于小波特征的核最近邻凸包分类算法用于人脸识别是可行有效的。

**总结** 本文提出的基于小波特征的核最近邻凸包人脸识别算法是将二维小波变换与核最近邻凸包分类器有效融合的非线性模式分类方法。ORL 人脸数据库上的实验表明了这一算法的良好分类能力,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

- 1 边肇祺,张学工,等编著. 模式识别[M](第二版). 清华大学出版社,1999
- 2 Boser B Z, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers[A]. In: Proceedings of the 5th Annual ACM. Workshop on Computational Learning Theory[C]. Pittsburgh, PA, July ACM Press, 1992. 144~152
- 3 Vapnik V N. The Nature of Statistical Learning Theory[M]. New York: Springer-Verlag, 1995
- 4 Schölkopf B, Smola A, Muller K R. Kernel principal component analysis[A]. In: W. Gerstner, ed. Artificial Neural Networks - ICANN'97[C], Berlin, 1997. 583~588
- 5 Mika S, Ratsch G, Weston J, Scholkopf B, Muller K R. Fisher discriminant analysis with kernels[A]. In: IEEE Neural Networks for Signal Processing Workshop[C], 1999. 41~48
- 6 Bach F, Jordan M L. Kernel independent component analysis[A]. Technical Report CSD-01-1166[R]. Computer Science Division, University of California, Berkeley, 2001
- 7 Bennett K P, Bredensteiner E J. Duality and geometry in SVM classifiers[A]. In: P. Langley, ed. Proceedings of the 17th International Conference on Machine Learning[C], San Francisco, California, Morgan Kaufmann,2000. 57~64

- 8 Keerthi S S, Shevade S K, Bhattacharyya C, Murthy K R K. A fast iterative nearest point algorithm for support vector machine classifier design. Neural Networks[J]. IEEE Transactions on, 2000, 11(1):124~136
- 9 邓乃扬,田英杰著. 数据挖掘中的新方法—支持向量机[M]. 科学出版社,2004
- 10 Mallat S. Multiresolution Approximation and Wavelet Orthonormal Bases of L2(R)[J]. Trans. Amer. Math. Soc., 1989,315: 69~87
- 11 Mallat S. A Theory for Multiresolution Signal Decomposition; the Wavelet Representation[J]. IEEE Trans. PAMI, 1989, 11 (5):674~693
- 12 Meyer Y. Ondelettes et Opérateurs[M]. Paris: Hermann Press, 1990
- 13 徐长发,李国宽著. 实用小波方法[M]. 华中科技大学出版社, 2004
- 14 Daubechies I. Orthonormal Bases of Compactly Supported Wavelets[J], Comm. Pure and Appl. Math, 1988,41: 909~996
- 15 海金著,叶世伟,史忠植译. 神经网络原理(原书第2版)[M]. 机械工业出版社,2004
- 16 王日爽编著. 泛函分析与优化理论[M]. 北京航空航天大学出版社,2003
- 17 Kinderlehrer D. 变分不等方程及其应用[M]. 郭友中等译,科学出版社,1991

(上接第 217 页)

由表 2 看到 HMM/SVM 模型的提高的识别效果是显著的。但有的序列,例如 Nematode promoters 经过二次识别后虽然比第一次大幅度提高了识别效果,但仍然只有 88.8%。

笔者认为,有三种可能的原因,一是所训练的序列太少,导致 L 值范围的偏离,影响了识别效果;另一个是本身序列数据的原因,也许是提交的数据库序列存在测序上的误差;三是我们的模型还需进一步改进。

表 2 双层识别模型进行 DNA 多家簇分类

DNA 类	L 值范围	测试数	第一次识别(传统)(HMM)		第二次识别(新型)(HMM+SVM)	
			正确识别数	识别率	正确识别数	识别率
Human intron	-402.3~-426.8	200	180	90%	186	93%
Man promoters	-433.3~-480.8	100	87	87%	91	91%
Plant promoters	-285.6~-305.9	100	88	88%	91	91%
Echinoderm promoters	-420.2~-425.6	20	0(无法识别)	0%	17	85%
Nematode promoters	-448.1~-453.7	18	15	83.3%	16	88.8%

近年来,SVM 和 HMMS 都是模型识别的主流分支,SVM 有它独有的小样本处理能力及基于核函数特征映射使之得到迅速发展;而 HMMS 由于能对基因序列符号进行编解码,其插入、删除、匹配正好模拟了基因的进化过程,使它在生命科学领域备受欢迎。国际上已经建立的著名的基于隐马尔可夫模型蛋白质家族识别数据库(基于多重序列比对) Pfam [http://pfam.wustl.edu/] 就是一例。

**讨论** 本文利用隐马尔可夫模型训练中不同结构的序列其 L 值分布有不同范围的特点,对传统分类“投票模型”进行改进,提出一种 HMM/SVM 混合识别模型。先用 HMM 进行识别,同时,对于 L 值区域有重叠造成识别率较低的部分,我们再用支持向量机进行第二次过滤识别。该方法克服了传统单一 HMM 分类方法的不足,提高了识别效果。虽然隐马尔可夫模型也有许多不足,例如 HMM 只是线性模型、独立性假设(HMM 假设观察符号之间是相互独立的,但它们之间可能相互依赖)、分类域值选取问题(过紧和过松的域值选取,会导致样本的少分或错分)、小样本(有的数据库仅有少数几条序列)训练带来的 L 值误差,这些缺点结合 SVM 可以得到弥补,因为 SVM 有极强的小样本处理能力,SVM 欠缺的多分类能力又恰恰是 HMM 所具有的,所以 HMM 和 SVM 结合过滤,达到了优势互补。

征以及如何进一步优化模型仍是将来要继续研究的工作。

参考文献

- 1 Vaseghi S V, Milner B P. Noise compensation methods for hidden Markov model speech recognition in adverse environments. Speech and Audio Processing, IEEE Transactions on, 1997, 5(1): 11~21
- 2 Gordan M, Kotropoulos C, Pitas I. Application of support vector machines classifiers to visual speech recognition. Image Processing. In: 2002. Proceedings. 2002 International Conference on, June 2002, 3: 129~132
- 3 Lefevre S, Bouton E, Brouard T, Vincent N. A new way to use hidden Markov models for object tracking in video sequences. Image Processing, 2003. ICIP 2003. In: Proceedings. 2003 International Conference on Volume 3, Sept. 2003, 2(III-1):17~20
- 4 Fu Y, Shen R, Lu H. Watermarking scheme based on support vector machine for colour images. Electronics Letters, 2004, 40 (16): 986~987
- 5 Baum L E, Sell G R. Growth functions for transformations on manifolds. Pac. J. Math., 1968, 27(2):211~227
- 6 Baum L E, Petrie T. Statistical inference for probabilistic functions of finite state Markov chains. Annmath. Stat., 1996, 37: 1554~1563
- 7 Kundsens S. Promoter 2.0: for the Recognition of Poll Promoter Sequences. Bioinformatics, 1999, 15: 356~361
- 8 Kasabov N, Pang S. TRANSDUCTIVE SUPPORT VECTOR MACHINES AND APPLICATIONS IN BIOINFORMATICS FOR PROMOTER RECOGNITION. In: IEEE int. Conf. Neural Networks & Signal Processing. Nanjing, China, December, 2003. 14~17