

《人脸识别技术》专题文章导读

龚卫国

重庆大学 光电技术及系统教育部重点实验室

自 20 世纪 90 年代以来,人脸识别技术就逐渐成为模式识别领域的研究热点之一。尤其是近年来,随着反恐、国防安全和社会安全的需要,世界各国都加大了对人脸识别技术研究的投入。我国“863 计划”、国家科技支撑计划、自然科学基金等都拨出专款资助人脸识别的相关研究,国家“十一五”科技发展规划也将人脸识别技术的研究与发展列入其中。

人脸识别是一种非接触式的、友好的生物特征识别方法,在保密场所的用户认证,银行、住宅小区的安防监控,公安刑侦,人机交互等领域都有着极大的应用潜力。但正是由于非接触式这一特点,使人脸识别很容易受到光照、姿态、表情、年龄、遮挡等因素的影响,造成识别精度的大幅降低。目前,鲁棒的人脸识别方法是人脸识别技术研究的主要内容之一。

人脸识别是典型的高维、小样本模式识别问题。解决这类问题的关键是如何提取并选择对分类有意义的特征,以降低样本的维度,达到维数约简的目的。一个完整的人脸识别系统一般由图像预处理、特征提取和分类识别三个部分组成。本专题论文是长期从事人脸识别研究的重庆大学光电技术及系统教育部重点实验室科研人员在国家“863 计划”项目、国防“十一五”基础科研项目和省部委的 10 余项重点研究项目的支持下,在前期研究成果分别发表在专业刊物《PATTERN RECOGNITION》、《NEUROCOMPUTING》、《光学 精密工程》、《光学工程》、《仪器仪表学报》和国际会议 ICPR、CVPR、SPIE 等的基础上,进一步通过大量的理论分析和实验总结提出的一些提高人脸识别性能的方法:文一《人脸特征选择中的 SVM 泛化误差估计》将最小化 SVM 的 VC 误差界并将支持向量 span 误差界作为 Wrapper 模型的特征选择判据,通过递归特征排除搜索策略(RFE)选择优化的特征子集。文二《基于多级小波分解的人脸图像光照补偿方法》介绍了一种利用多级小波分解对人脸图像光照进行有效补偿的方法,在一定程度上解决了人脸识别的光照问题;文三《随机采样子空间保局投影人脸识别算法》融合了多个既有差异又互补的保局投影子空间上的分类识别结果,有效地增强了无监督保局投影算法的鉴别能力。文四《基于有监督核局部线性嵌入的面部表情识别》将核方法和有监督局部线性嵌入结合起来,在核空间中利用有监督局部线性嵌入降低样本的维度,在达到维数约简目的的同时保留了不同类别间的鉴别信息。

上述文章分别从人脸识别的图像预处理、特征提取等阶段提出了相应的解决方法,在一定程度上减轻了光照、表情等因素对人脸识别的影响,较为有效地解决了特征维数约简问题,取得了很好的效果。但同时我们也意识到,目前的大多数人脸识别方法都仅仅是从某个方面解决人脸识别中的难点问题,一种鲁棒的人脸识别方法仍将是当前和今后一段时间内研究的重点内容。个人认为,探索基于人类视觉感知机理的人脸识别新方法和新技术是今后该研究领域重要的研究方向之一。

最后,希望上述文章能对从事人脸识别研究领域的科研人员有所帮助,并请读者提出宝贵意见。

文章编号 1004-924X(2008)08-1452-07

人脸特征选择中的 SVM 泛化误差估计

李伟红,龚卫国,杨利平,辜小花

(重庆大学 光电技术及系统教育部重点实验室,四川 重庆 400044)

摘要:为了研究在人脸特征选择中用支持向量机(SVM)泛化误差界作特征选择判据的有效性,结合过滤(Filter)和封装(Wrapper)模型构造了人脸特征选择及识别的新框架,将小波变换(WT)和核主元分析(KPCA)作为 Filter 模型,最小化 SVM 的 VC 维(VC)留一法(LOO)误差界及支持向量 span 误差界作为 Wrapper 模型的特征选择判据;通过递归特

收稿日期:2007-12-21;修订日期:2008-02-01.

基金项目:国防“十一五”基础科研基金资助项目(No. C10020060355);国家 863 高技术研究发展计划资助项目(No. 2007AA01Z423);重庆市科技攻关研究项目(No. CSTC2007AC2018)

征排除法(RFE)在 UMIST 人脸图像库上进行人脸特征选择及识别实验。实验结果表明:判据为 VC 维的 LOO 误差界和支持向量 span 误差界时,特征维数可以分别降低到 80 和 70,而分类识别率仍然能达到 94% 以上,表明本文所提出的特征选择判据和特征搜索策略是解决人脸特征选择问题的一种有效方法。

关键词: SVM 泛化误差界;人脸特征选择;Filter 模型;Wrapper 模型;递归特征排除法

中图分类号: TP391.41 **文献标识码:** A

SVM generalization error estimation for facial feature selection

Li Wei-hong, Gong Wei-guo, Yang Li-ping, Gu Xiao-hua

(Key Laboratory of Opto-electronic Technology & System of the Ministry of Education,
Chongqing University, Chongqing 400044, China)

Abstract: To investigate the validity of Support Vector Machine(SVM) generalization error bounds as the feature selection criterion, a novel framework of facial feature selection based on Filter and Wrapper approaches was proposed. By taking a Wavelet Transformation(WT) and Kernel Principal Component Analysis(KPCA) as a Filter approach, the Vapnik-Chervonenkis(VC) Leave-one-Out (LOO) error bound was minimized. Then, the span bound of support vector was regarded as the feature selection criterion of Wrapper approach. Finally, Recursive Feature Elimination(RFE) search strategy was used for searching optimum facial subset. The experiments on UMIST face database were executed by the proposed method. The experimental results indicate that the facial feature dimensions can be reduced to 80 and 70, respectively, while both of the classification accuracies remain over 94%, so the proposed feature selection criterion and search strategy are effective for facial feature selection.

Key words: Support Vector Machine(SVM) generalization error; facial feature selection; Filter approach; Wrapper approach; Recursive Feature Elimination(RFE)

1 引言

特征选择(Feature Selection, FS)是人脸识别中的关键问题,也是模式识别及机器学习领域长期以来极其重要的研究方向。特征选择的任务是从大量基本特征中去掉不相关以及冗余的特征以获取对目标函数及分类有利的特征子集^[1]。好的特征选择方法应该使所选择的特征子集即使采用不同的预处理算法或最普通的分类器也能获得较好的分类效果。特征选择算法从模型上可以分为 Filter 和 Wrapper。Filter 模型将特征选择作为滤除不相关特征的预处理过程,独立于学习算法,直接利用所有训练数据的统计性能来评估特征。Wrapper 模型基于某种学习算法,将所选择的特征子集的性能作为特征评估判据反馈给学习算法并进行修正,不断循环此过程就可以获得最

优特征子集^[2]。

支持向量机(SVM)作为有效的分类工具已经广泛应用于机器学习及模式识别领域^[3-4],但 SVM 并不提供对相关特征及冗余特征进行自动分析的环节。对于高维小样本的人脸识别问题,将特征选择、SVM 模型优化及分类结合,理论上有利于 SVM 泛化能力的提高。根据统计学习理论^[5],SVM 是通过控制某个预测泛化误差界或其他相关性能来控制学习机复杂度的。近年有研究人员也开展了相关的工作,Weaton 等^[6]介绍了寻找具有最小化 VC 维(Vannik-cherouonenkis)留一法(Leave-One-Out, LOO)误差界的特征子集的特征选择方法,并通过梯度下降算法代替贪婪算法,分别对模拟数据和 DNA 分析的真实数据进行了实验。Grandvalet 等^[7]研究了可以自动判定输入特征相关性的算法。其基本原理是采用特征选择缩放因子法确定一个输入空间度量,并将

其中不相关的特征赋予零权重,然后基于标准 SVM 的经验风险最小化,再对输入的空间度量进行调节。Holger 等^[8]基于最小化正则风险结合一种改进型的递归特征排除法(Recursive Feature Elimination, RFE)进行特征选择。Rakotomamonjy^[9]报告了 SVM 中不同选择判据对特征选择的影响,并在模拟数据上进行了实验。Duan^[10]对于在 L1 和 L2 范式软间隔下的 k 折交叉验证误差, Xi-Alpha 界, GAVC (Generalized Approximate Cross-Validation), 半径/间隔误差界以及支持向量 span 误差界对 SVM 模型设计性能的影响进行了分析实验。作者的前期研究^[11]将权重矢量 $\|w\|^2$ 和半径/间隔 $R^2 \|w\|^2$ 作为 SVM RFE 的特征选择判据,采用缩放因子梯度搜索算法进行了人脸特征选择实验。本文继续研究了通过 RFE 搜索策略最小化 SVM VC 的 LOO 误差界和支持向量 span 误差界来获取优化的特征子集问题,并根据人脸特征高维、稀疏及小样本的特点建立有效的用于人脸识别的特征选择框架:将小波变换(Wavelet Transformation, WT)、核主元分析(Kernel Principal Component Analysis, KPCA)作为 Filter 模型部分,去除高维人脸特征的冗余和相关特征,再通过把最小化 SVM 的 VC 维的 LOO 误差界或支持向量 span 误差界作为 Wrapper 模型部分的特征选择判据来获取直接用于 SVM 分类的优化特征子集。在 UMIST 人脸图像库上进行了相关实验,结果显示所提出的特征选择判据、特征搜索策略及人脸特征选择及识别构架是有效的。

2 SVM

SVM^[3]通过非线性变换将输入空间(x)映射到一个高维空间($z = \phi(x)$),然后在这个新空间中通过 $w \cdot z - b = 0$ 构建一个最优线性超平面,最大化地将待分的两类分开。对于具有 L1 范式软间隔的 SVM,即为求解最优化问题($p=1$):

$$\begin{cases} \min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i^p \right) \\ \text{s. t. } y_i [(w \cdot x_i) + b] \geq 1 - \xi_i, \xi_i \geq 0 \end{cases}, \quad (1)$$

其中 x_i 是第 i 个样本, y_i 是相应的标签。当确定了核函数 $K(x_i, x_j)$ 后,问题就变成了求解二次规划对偶问题:

$$\begin{cases} \max W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \\ \text{s. t. } 0 \leq \alpha_i \leq C, i = 1, \dots, n, \sum_{i=1}^n y_i \alpha_i = 0 \end{cases}, \quad (2)$$

参数 C 起到了平衡分类偏差与泛化能力的作用,与核函数中的参数统称为超参数。

3 基于 SVM 的特征选择问题

在分类问题中,特征选择的目的是从原始特征中选出对分类有利的特征子集。根据 SVM 和泛化性理论,在有限样本下学习精度和泛化能力之间的矛盾是通过控制泛化误差或其他相关性能来进行惩罚的。

3.1 SVM 泛化误差界与特征选择

核函数的引入极大地提高了学习机的表达能力,并保持了内在的线性,从而使得学习容易得到控制,SVM 模型的假设是训练样本和测试样本根据未知但固定的分布 D 产生的。因此定义分类函数 h 在分布 D 上的误差 $\text{Err}_D(h)$ 为:

$$\text{Err}_D(h) = D\{(x_i, y_i) : h(x_i) \neq y_i\}. \quad (3)$$

根据泛化性理论,其误差一定存在与分布无关的误差界。控制核函数特征空间的适应性,其最好的描述是 VC 理论。SVM 的 VC 误差界可以近似描述为:

$$\text{Err}_{\text{VC}} = h \leq R^2 \|w\|^2 + 1, \quad (4)$$

其中, R 为 Hilbert 空间中包含 $\phi(x_1), \dots, \phi(x_n)$ 的最小超球的半径。

Vapnik^[10]提出硬间隔的 VC 维 LOO 误差界为:

$$\text{LOO Err} \leq \frac{1}{4n} R^2 \|w\|^2. \quad (5)$$

Chapelle^[10]提出 L1 范式的软间隔 VC 维 LOO 误差界为:

$$\text{LOO Err} \leq \frac{1}{n} \left[R^2 \sum_{i=1}^n \alpha_i + \sum_{i=1}^n \xi_i \right],$$

或

$$\text{LOO Err} \leq \frac{1}{n} \left[R^2 \|w\|^2 + (R^2 C + 1) \sum_{i=1}^n \xi_i \right], \quad (6)$$

式(5)、(6)也被称为半径/间隔误差界。

Vapnik^[13]还提出近似支持向量 span 方法,

它是基于映射函数 $\phi(\mathbf{w}_i)$ 和其他所有的支持向量间的距离 S_p 的一种近似 LOO 估计,其误差界为:

$$\text{Err}_s \leq \sum_p \alpha_p^* \mathbf{S}_p^2, \quad (7)$$

其中, \mathbf{S}_p^2 是支持向量间点积的矩阵,

$$\mathbf{S}_p^2 = 1/(\tilde{\mathbf{K}}_{\text{SV}}^{-1})_{pp}, \quad (8)$$

\mathbf{K} 是第一类别支持向量间的点积矩阵

$$\tilde{\mathbf{K}}_{\text{SV}} = \begin{pmatrix} \mathbf{K} & \mathbf{I} \\ \mathbf{1}^T & 0 \end{pmatrix}. \quad (9)$$

3.2 泛化误差界梯度

设 SVM 泛化误差为 Err , D 为训练样本集, θ 为 SVM 超参数,则基于 SVM 的特征选择需要解决最优化问题:

$$\min_{\theta} \text{Err}(\theta). \quad (10)$$

通过控制 SVM 泛化误差界获取优化特征子集需求解公式(10)所示的最优化问题,为此需要计算出相应的梯度估计。

R_{θ}^2 梯度为式(11)目标函数的优化

$$\max_{\beta} \sum_{i=1}^n \beta_i K_{\theta}(\mathbf{x}_i, \mathbf{x}_i) - \sum_{i,j=1}^n \beta_i \beta_j K_{\theta}(\mathbf{x}_i, \mathbf{x}_j), \quad (11)$$

$$\text{s. t. } \begin{cases} \sum_{i=1}^n \beta_i = 1 \\ \beta_i \geq 0, i = 1, \dots, n \end{cases}$$

而 \mathbf{w}_{θ}^2 由式(12)得到:

$$\mathbf{w}_{\theta}^2 = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K_{\theta}(\mathbf{x}_i, \mathbf{x}_j), \quad (12)$$

$$\text{s. t. } \sum_{i=1}^n \alpha_i y_i = 0, i = 1, \dots, n$$

则 Err_{VC} 的梯度估计为

$$\frac{\partial \text{Err}_{\text{VC}}}{\partial \theta} = \left\| \mathbf{w}_{\theta} \right\|^2 \sum_{i,j} (\beta_i \beta_j - \beta_i \delta_{ij}) \frac{\partial K_{\theta}(\mathbf{x}_i, \mathbf{x}_j)}{\partial \theta} + R_{\theta}^2 \sum_{i,j} \alpha_i^* \alpha_j^* y_i y_j \frac{\partial K_{\theta}(\mathbf{x}_i, \mathbf{x}_j)}{\partial \theta} \Big| \quad (13)$$

支持向量 span 误差界的梯度估计为

$$\frac{\partial \text{Err}_s}{\partial \theta} = \sum_p \alpha_p^* \frac{\partial \mathbf{S}_{\theta_p}^2}{\partial \theta} = \sum_p 2 \left(-\mathbf{H}^{-1} \frac{\partial \mathbf{H}}{\partial \theta} \alpha^* \right)_{pp} \mathbf{S}_p^2 + \alpha_p^* \mathbf{S}_p^4 \left(\tilde{\mathbf{K}}_{\text{SV}}^{-1} \frac{\partial \tilde{\mathbf{K}}_{\text{SV}}^{-1}}{\partial \theta} \tilde{\mathbf{K}}_{\text{SV}}^{-1} \right)_{pp}, \quad (14)$$

其中, $\mathbf{H} = \begin{pmatrix} \mathbf{K}^T & \mathbf{Y} \\ \mathbf{Y}^T & 0 \end{pmatrix}$, $K_{ij}^r = y_i y_j K_{\theta}(\mathbf{x}_i, \mathbf{x}_j)$ 。

4 人脸特征选择及识别构架

人脸特征具有高维、稀疏及小样本的特点。为了满足人脸识别的有效性及时性要求,本论文提出了如图 1 所示的基于 Filter 和 Wrapper 特征选择模型的新框架。Filter 模型的作用是有效降低人脸原始特征的维数并去掉不相关以及冗余的特征;Wrapper 模型的作用是通过将最小化 SVM

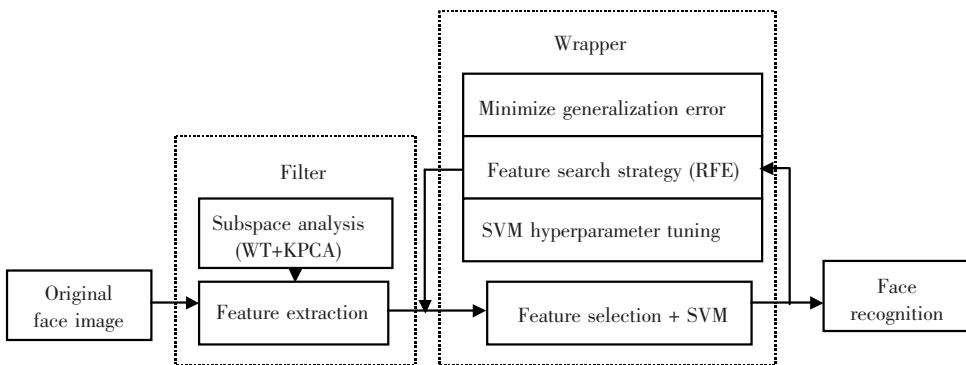


图 1 基于 Filter 和 Wrapper 模型的人脸特征选择及识别框架

Fig. 1 Outline of facial feature selection and recognition based on Filter and Wrapper

泛化误差界作为特征选择判据,选择对分类直接有效的特征。理论上讲,由于提出的框架以 SVM 及泛化理论为基础,它具有 SVM 本质上的优点,可以克服“过拟合”,具有很好的泛化能力。

4.1 Filter 特征选择模型

Filter 特征选择模型部分由 WT 和 KPCA 组成。在图像分析中小波实质上是对信号用一种不同尺度的带通滤波器进行滤波,将信号分解到不

同的频带,高频部分描述的是图像的细节信息,而人脸的光照、少许遮罩、旋转扭曲和面部表情等只影响图像中高频部分。为此,利用 WT 所获得的人脸低频信息可以较好地描述对分类有用的人脸特征。

KPCA^[14]是将主元分析(Principal Component Analysis, PCA)推广到非线性领域,其基本思想是通过一个非线性变换 Φ 把输入数据 $x, x \in \mathbf{R}^N$ 映射到一个高维的特征空间 $F = \{\Phi(x) : x \in \mathbf{R}^N\}$ 。则可以把在输入空间无法线性分类的数据变换到特征空间,再利用 PCA 进行特征提取,即将在输入空间线性不可分问题转换到高维特征空间进行线性分类。

4.2 Wrapper 特征选择模型

Wrapper 特征选择模型通过控制 SVM 某个预测泛化误差界或其他相关性能选择对分类直接有效的特征。本文利用 SVM 的 VC 维 LOO 误差界和支持向量 span 误差界,通过 RFE 特征搜索策略进行特征选择。RFE 是由 Guyon 等^[12]提出的在一定的特征排列判据下逐个排除特征以获得优化的特征子集的方法。在进行特征选择时,首先训练分类器,然后根据最小化特征选择判据 $\min_{\theta} \text{Err}(\theta)$ 计算特征的排列,移去排列最后的特征,不断循环直到判据没有明显减小或者已经得到需要的特征数为止。

5 实验结果及分析

为了验证所提出的特征选择判据、特征搜索策略及人脸特征选择及识别框架的有效性,在 UMIST 人脸图像库^[16]上进行了相关实验。UMIST 人脸图像库包括 20 个人的 559 幅图像

(112×92),每人的图像为 19~42 幅,主要具有不同视点、姿态。原始图像实例如图 2 所示,实验中每人随机取 6 幅作为训练样本,13 幅作为测试样本。



图 2 UMIST 人脸图像库的原始图像

Fig. 2 Image examples from UMIST image database

实验主要比较特征选择前后特征维数、识别率及识别时间的变化以及不同 SVM 泛化误差界对特征选择的影响。通过实验确定核函数为多项式核 $K(x_i, x_j) = (x_i \cdot x_j)^d$ ^[16]。SVM 超参数调节实验用交叉验证法,多项式核参数选择 $d=0.1$ 。

表 1 为在 UMIST 人脸图像库中人脸原始图像和人脸原始图像通过 WT 和 KPCA 后, SVM 的分类结果,并与本文提出的人脸特征选择及识别方法进行比较。Wrapper 模型中分别选用了最小化 $\|w\|^2$ 、VC 维 LOO 误差界及支持向量 span 误差界。

从表 1 可以看到直接采用原始数据进行分类,不仅分类识别率低,同时识别时间长。经过 WT 或 KPCA 处理,在特征维数及分类识别率上都有不同程度的改善。而采用本文提出的方法可以保证在分类识别率不变的情况下特征维数继续降低。可以看出这是一种解决高维数据特征选择及分类的有效方法。当最小化泛化误差界不同时,实验结果略有不同。

图 3 为原始图像经过 KPCA 特征提取后,在采用最大化 margin(最小化 $\|w\|^2$)、最小化 VC LOO 误差及最小化支持向量 span 误差界所得到

表 1 实验结果

Tab. 1 Experimental results

UMIST Facial database	Feature number	Recognition rate/%	Recognition time/s
Original +SVM	10 304	86.92	15.893 0
WT+SVM	1 326	92.69	1.041 0
KPCA+SVM	119	94.62($d=0.1$)	0.110 0
WT+KPCA+SVM RFE($\ w\ ^2$)	100	94.62($d=0.1$)	0.080 0
WT+KPCA+SVM RFE(LOO)	80	94.62($d=0.1$)	0.081
WT+KPCA+SVM RFE(span)	70	94.23($d=3$)	0.032

的结果,从图中可以得到,基于支持向量 span 误差界的方法在维数为 70 时,可以获得 94% 的分类识别率,在 3 种方法中最好,即识别率最高时维度最低。

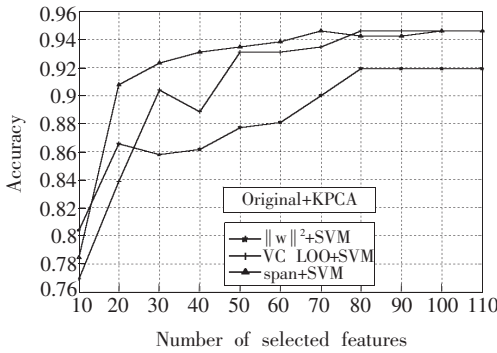


图 3 泛化误差界对人脸特征选择的影响(KPCA)

Fig. 3 Effect of different generalization error bounds on feature selection (KPCA)

图 4 为原始图像经过 WT 和 KPCA 特征提取后,基于最大化 margin(最小化)、最小化 VC 维的 LOO 误差界及最小化支持向量 span 误差界所得到的结果。从图中可以得到,仍然是基于支持向量 span 误差界的方法最好-维数为 70 时,94% 的分类识别率。但此方法三种泛化误差界所获得的结果比较近似。根据特征选择的原则,好的特征选择方法应该是在经过同预处理或前期处理后均能获得较好的结果。可以认为经过 WT+KPCA 后,基于最小化支持向量 span 误差界是一种较好的特征选择方法。

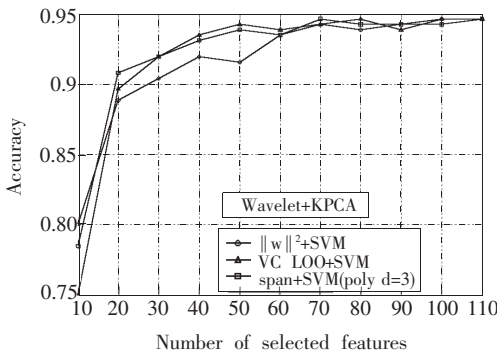


图 4 泛化误差界对人脸特征选择的影响(WT+KPCA)

Fig. 4 Effect of different generalization error bounds on feature selection (WT+KPCA)

图 4 结果表明在 UMIST 人脸图像数据库中,本论文所提出的方法是一种有效的特征选择方法。但由于核参数的不同意味着原始数据投影到的 Hilbert 空间不同,即从分类意义上得到的最优特征空间将不同。为此,本文采用 grid search 法搜索采用最小化支持向量 span 误差界作为特征选择判据时的多项式最佳核参数。图 5 表明多项式核参数 $d=1$ (可以认为是线性核) 时实验结果较差,其余结果较类似。

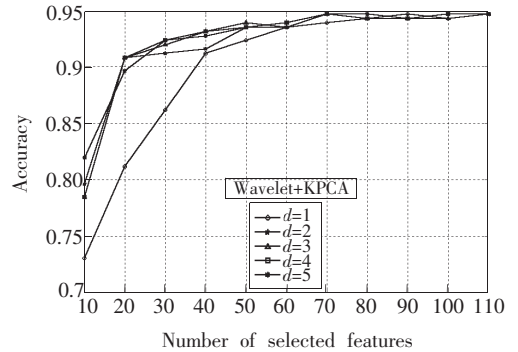


图 5 在不同核参数下最小化支持向量 span 误差界对人脸特征选择的影响

Fig. 5 Effect of support vector span error bound on feature selection with different kernel parameters

6 结 论

本文研究了通过 RFE 搜索策略、最小化 SVM VC 维的 LOO 误差界和支持向量 span 误差界来获得优化的特征子集的方法。根据人脸特征高维、稀疏及小样本的特点建立了 Filter(WT+KPC)和 Wrapper(最小化 SVM 泛化误差界)特征选择构架,并在 UMIST 人脸图像库上进行了相应的实验。结果显示:判据为最小化 VC 的 LOO 误差界和支持向量 span 误差界时,人脸特征维数可以分别降低到 80 和 70,而分类识别率仍然能达到 94% 以上。这表明本文所提出的特征选择判据和特征搜索策略以及构建的人脸特征选择框架是合理、有效的,同时也表明通过最小化 SVM 泛化误差界进行特征选择是一种有效的方法,可以有效地控制 SVM 学习机的复杂性,提高其泛化能力。

参考文献:

- [1] DASH M, LIN H. Feature selection for classification [J]. *Intelligent Data Analysis*, 1997 (1):131-156.
- [2] JOHN G, KOHAVII R, PFLEGER K. Irrelevant features and the subset selection problem[C]. *The Eleventh International Conference on Machine Learning*, 1994:121-129.
- [3] BURGESS J C. A tutorial on support vector machines for pattern recognition [J]. *Data Mining Knowledge Discovery*, 1998, 2(2):121-167.
- [4] 赵吉文,刘永斌.基于 SVM 和遗传算法的新型直线电机结构参数优化[J]. *光学 精密工程*,2006,14(5):870-875.
ZHAO J W, LIU Y B. Parameter optimization of novel cylinder type linear motor based on SVM and genetic algorithm [J]. *Opt. Precision Eng.*, 2006,14(5):870-875. (in Chinese)
- [5] VAPNIK V. *Statistical Learning Theory* [M]. New York:John Wiley and Sons Inc.,1998.
- [6] WESTON J, MUKHERJEE S, CHAPPELLE O. *et al.*. Feature selection for SVM [J]. *In Advances in Neural Information Processing Systems*, 2000(13):668-674.
- [7] GRANDVALET Y, CANU S. Adaptive scaling for feature selection in SVM[J]. *Advances in Neural Information Processing Systems*, 2003, 15:553-560.
- [8] HOLGER F, ANDREAS Z. Feature subset selection for support vector machines by incremental regularized risk minimization [J]. *Proc. Int. Joint Conf. on Neural Networks*, 2004(3):2041-2046.
- [9] RAKOTOMAMONJY A. Variable selection using SVM based criteria [J]. *Machine Learning Research*,2003,3:1357-1370.
- [10] DUAN K, KEETHI S S. Evaluation of simple performance measures for tuning SVM hyper parameters [J]. *Neurcomputing*,2003, 51:41-59.
- [11] 李伟红,龚卫国,陈伟民,等. RFE 的人脸特征选择方法研究[J]. *光电工程*,2006,33(5):113-117.
LI W H, GONG W G, CHEN W M, *et al.*. Facial feature selection method based on SVM RFE [J]. *Opto-Electronic Engineering*, 2006, 33(5):113-117. (in Chinese)
- [12] GUYON I, WESTON J. Gene selection for cancer classification using support vector machines[J]. *Machine Learning*, 2002,46:389-422.
- [13] VAPNIK V, CHAPPELLE V. Bounds on Error Expectation for Support Vector Machines [J]. *Neural Computation*,2000,12(9):2013-2036.
- [14] CAO L J, CHUA K S, CHONG W K, *et al.*. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine [J]. *Neurocomputing*, 2003(55):321-336.
- [15] GRAHAM D B, ALLINSON N M. Characterizing virtual eigensignatures for general purpose face recognition [J]. *In:Wechsler H, Phillips P J, Bruce V, Soulie F F, Huang T S. (eds.);Face Recognition:From Theory to Application. NATO ASI Series F, Computer and Systems Sciences*,1998,163:446-456.
- [16] 李粉兰,唐文彦.分数次幂多项式核函数在核直接判别式分析中的应用[J]. *光学 精密工程*,2007,15(9):1410-1414.
LI F L,TANG W Y. Application of fractional power polynomial kernel function of kernel direct discriminant analysis[J]. *Opt. Precision Eng.*,2007,15(9):1410-1414. (in Chinese)

作者简介:李伟红(1962—),女,博士,副教授,主要研究方向为模式识别及图像处理。E-mail:weihongli@cqu.edu.cn