

# 一种基于改进 BP 神经网络的 PCA 人脸识别算法

李康顺<sup>1 2</sup> 李 凯<sup>1</sup> 张文生<sup>3</sup>

<sup>1</sup>(江西理工大学理学院 江西 赣州 341000)

<sup>2</sup>(华南农业大学信息学院 广东 广州 510642)

<sup>3</sup>(中国科学院自动化研究所 北京 100190)

**摘 要** 人脸识别作为模式识别领域的热点研究问题受到了广泛的关注。传统 BP 算法虽然具有自学习、自适应以及强大的非线性映射能力并且在人脸图像识别准确率上占有很大的优势,但算法具有收敛缓慢、训练过程振荡、易陷入局部极小点等缺点。针对传统 BP 算法的不足提出一种基于改进 BP 神经网络的 PCA 人脸识别算法,该算法采用 PCA 算法提取图像的主要特征,并结合一种新的权值调整方法改进 BP 算法进行图像分类识别。仿真实验表明,通过使用该算法对 ORL 人脸数据库的图像进行识别,其结果比传统算法具有更快的收敛速度和更高的识别率。

**关键词** 人脸识别 主成分分析 BP 神经网络 附加动量 弹性梯度下降法

中图分类号 TP391

文献标识码 A

DOI: 10.3969/j.issn.1000-386x.2014.01.041

## PCA FACE RECOGNITION ALGORITHM BASED ON IMPROVED BP NEURAL NETWORK

Li Kangshun<sup>1 2</sup> Li Kai<sup>1</sup> Zhang Wensheng<sup>3</sup>

<sup>1</sup>(School of Science, Jiangxi University of Science & Technology, Ganzhou 341000, Jiangxi, China)

<sup>2</sup>(School of Information, South China Agricultural University, Guangzhou 510642, Guangdong, China)

<sup>3</sup>(Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract** Face recognition, as a focus of the research in pattern recognition field, has gained increasing attention. Traditional BP algorithm has a strong ability in self-learning, self-adaptivity and nonlinear mapping. Moreover, it has a significant predominance in human face recognition accuracy. However, the algorithm also has shortages including slow convergence, training process oscillation and easy to fall into local minima. In light of these deficiencies of traditional BP neural network, we propose a PCA face recognition algorithm which is based on improved BP neural network. The algorithm uses PCA algorithm to extract principal features of face image and uses a new weight adjustment method to improve the BP algorithm for image classification and recognition. Simulation experimental results show that faster convergence speed and higher recognition rate are achieved when using the improved algorithm to identify the images in ORL face database than the traditional algorithm.

**Keywords** Face recognition Principal component analysis BP neural network Additional momentum Elastic gradient descent method

## 0 引 言

人脸识别作为模式识别领域的热点研究问题受到了广泛的关注。研究人脸识别技术在众多领域的身份验证中有着广阔的应用前景<sup>[1-2]</sup>。PCA 方法是由 Turk 和 Pentlad 提出来的<sup>[3-4]</sup>。它以 K-L (Karhunen-Loeve 的缩写) 正交变换为基础,将数据从高维空间投影到低维空间中,从而达到降低数据维数的目的<sup>[5-6]</sup>。BP 神经网络具有自学习、自适应以及强大的非线性映射能力。通过权值的反向传播以及样本的不断学习,使网络具有很好的分类能力,适用于处理人脸识别问题。但是传统的 BP 神经网络算法具有收敛速度缓慢、训练过程易发生震荡、易陷入局部极小点等缺点<sup>[7-11]</sup>。本文针对传统 BP 神经网络算法的不足,提出一种基于改进 BP 神经网络的 PCA 人脸识别算法,该算法采

用 PCA 算法提取人脸图像中的主要特征,并结合一种新的权值调整方法改进 BP 算法进行人脸图像识别,最终达到快速准确的人脸图像分类和识别。

## 1 PCA 和 BP 算法

### 1.1 PCA 特征提取

主成分分析 PCA 是一种通过降低数据维数简化数据结构的方法。其本质就是将高维空间中的数据通过正交变换投影到低维空间中,从而实现了图像数据中主要特征的提取。运用主

收稿日期:2012-09-18。国家自然科学基金项目(70971043);江西省教育厅科学技术研究项目(GJJ112348)。李康顺,教授,主研领域:图像识别,演化计算。李凯,硕士生。张文生,研究员。

成分分析的方法求解一个特征子空间,将用于训练和测试图像向该特征子空间里做投影,可以得到用于区分不同人脸的主要特征(即投影系数)作为识别的特征向量。提取的主要特征就可以用分类器进行识别与分类了。

## 1.2 BP 算法的基本原理

BP 神经网络是一种采用误差反向传播的多层前馈神经网络,Robert Hecht-Nielsen 证明了具有一个隐含层的 3 层 BP 神经网络可以有效的逼近任意连续函数<sup>[12]</sup>。其中 3 层的 BP 网络包括输入层、隐含层、输出层。其基本结构如图 1 所示。

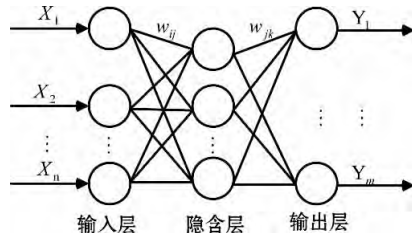


图1 三层BP网络结构图

BP 神经网络的基本思想是采用信号的正向传播和误差的反向传播。在正向传播中,从输入层传入的数据,依次在各隐含层处理,最终到达输出层。如果得到的不是期望输出,则将输出层的误差作为调整信号进行反向传播,根据预测误差不断调整权值和阈值,最终得到网络可以接受精度的输出。BP 神经网络具有自学习、自适应以及强大的非线性映射能力。

## 2 基于改进 BP 网络的 PCA 人脸识别算法

### 2.1 改进 BP 网络

传统的 BP 网络采用梯度下降算法在学习过程中存在收敛缓慢、训练过程震荡、易陷入局部极小点的问题。针对上述问题,本文提出一种基于附加动量的弹性梯度下降法的权值调整方式。利用这种权值调整方法改进 BP 神经网络,以加快网络的收敛速度和学习精度。

#### 1) 附加动量法

传统的 BP 神经网络在权值的调整过程中,没有考虑前一次的权值调整,从而常使学习过程容易发生震荡,不容易取得全局最优解。为了减小训练过程中的震荡问题,可以通过引入一个动量项来反映权值以往经验的积累<sup>[13]</sup>。BP 网络的权值调整量不仅和误差项对权值的梯度有关而且和前一次权值调整量有关。此时的权值调整公式为:

$$\Delta W(t+1) = (1 - mc) \eta \delta x + mc \Delta W(t) \quad (1)$$

其中  $\Delta W(t)$  为权值的变化量,  $t$  为训练次数,动量系数用  $mc$  表示,  $\eta$  为学习速率,  $\delta x$  为误差项,  $x$  为输入量;  $\delta x$  实质上反映了误差项对权值的梯度,可以看出附加动量法是考虑了以前经验的积累对权值调整的影响,其中动量系数  $mc \in (0, 1)$ 。通过引入一个动量项,可以将前一次的调整经验加入到权值的调整过程中来,从而降低了网络对局部极小点的敏感性,当误差曲面出现骤然变化时,更容易跳过局部极小点。

#### 2) 弹性梯度下降法

弹性梯度下降法是由德国的 Martin Riedmiller 和 Heinrich Braun 在 1993 年提出来的<sup>[14,15]</sup>。权值的调整量跟更新值  $\Delta t$  有关,梯度只影响权值的调整方向。这样可以有效地避免梯度大小造成网络收敛缓慢的问题。其中权值的调整量的大小与更新

值的关系为:

$$\Delta W(t) = \begin{cases} -\Delta t & \frac{\partial E(t)}{\partial W} > 0 \\ +\Delta t & \frac{\partial E(t)}{\partial W} < 0 \\ 0 & \frac{\partial E(t)}{\partial W} = 0 \end{cases} \quad (2)$$

则相应的  $t+1$  时刻权值调整公式为:

$$W(t+1) = W(t) + \Delta W(t) \quad (3)$$

更新值  $\Delta t$  的调整规则为:

$$\Delta t = \begin{cases} \alpha \times \Delta(t-1) & \frac{\partial E(t)}{\partial W} \times \frac{\partial E(t-1)}{\partial W} > 0 \\ \beta \times \Delta(t-1) & \frac{\partial E(t)}{\partial W} \times \frac{\partial E(t-1)}{\partial W} < 0 \\ \Delta(t-1) & \frac{\partial E(t)}{\partial W} \times \frac{\partial E(t-1)}{\partial W} = 0 \end{cases} \quad (4)$$

其中  $0 < \beta < 1 < \alpha$ 。

更新值的调整量与连续两次梯度的调整方向有关。如果连续两次的调整方向相同则增加更新值,如果连续二次的调整方向相反则减小更新值,其它情况更新值保持不变。通过引入弹性梯度下降法克服了梯度大小对网络造成的不利影响,使网络能够快速收敛。

#### 3) 一种新的权值调整方法

通过进一步的理论和实验分析得出,如果只引入附加动量项来改进 BP 神经网络,则能够使算法找到全局最优解,但是训练的时间过长;如果只采用弹性梯度下降法,虽然可以加快网络的收敛速度,缩短网络的训练时间,但是很容易陷入局部最优解。本文将两种方法有机地结合起来,提出了一种新的 BP 神经网络权值调整方式。其中  $t+1$  时刻的权值的调整公式为:

$$W(t+1) = \begin{cases} W(t) - \text{sign}\left(\frac{\partial E(t)}{\partial W}\right) \Delta t & \frac{\partial E(t)}{\partial W} \neq 0 \\ 0 & \frac{\partial E(t)}{\partial W} = 0 \end{cases} \quad (5)$$

当连续二次迭代的梯度方向相同,则权值的更新值增加,并且权值的调整方向不变。更新值  $\Delta t$  的调整公式为:

$$\alpha \times (1 - mc) \times \Delta(t-1) + mc \times \Delta(t-1) \quad (6)$$

其中  $mc \in (0, 1)$ 。

当连续二次迭代的梯度方向相反,则更新值将减小,权值以一定的概率  $P$  保持调整方向不变,本文中概率  $P$  定义为一个随着时间的增加而变小的数。其中  $P = \beta / ((1 + \beta)(e^{0.2t} + 1))$ ,  $P \in (0, 1)$ 。通过权值以一定的概率  $P$  保持调整方向不变,可以使网络有一定的几率跳出局部最优解。此时更新值  $\Delta t$  的调整公式为:

$$\beta \times (1 - a) \times \Delta(t-1) \times \text{sign}\left(\frac{\partial E(t)}{\partial W}\right) + a \times \Delta(t-1) \quad (7)$$

其中  $0 < \beta < 1$ ,  $a \in \left(0, \frac{\beta}{1 + \beta}\right)$ 。

另外,权值以  $1 - P$  的概率改变调整方向,其中更新值调整公式为:

$$\beta \times (1 - mc) \times \Delta(t-1) + mc \times \Delta(t-1) \quad mc \in (0, 1) \quad (8)$$

其他情况下,权值的更新值调整公式保持不变。这样,克服梯度的大小对网络造成收敛缓慢的不利影响,当网络的误差曲面出现骤然变化时,可以有效地减小网络陷入局部最优解,从而提高了网络的识别速度和识别率。

## 2.2 改进 BP 神经网络的 PCA 人脸识别算法流程

首先,使用主成分分析的方法将高维空间中图像数据映射到低维空间,提取出人脸库图像中的主要特征;然后运用本文提出的附加动量和弹性梯度下降法相结合的权值调整方式改进 BP 网络,对人脸库图像的主要特征进行训练;最后用测试图像对训练好的网络测试,获得相应的识别结果。模型的具体步骤如下:

**Step1** 读取人脸库中的图像,将  $M$  幅图像用于训练和  $M'$  幅图像用于测试。将分辨率为  $m \times n$  的图像按照每列互相连接,构成一个  $N = m \times n$  维的列向量。则可将整个人脸库中用于训练的图像存储在一个矩阵  $R^{N \times M}$ 。其中第  $i$  ( $i = 1, 2, \dots, M$ ) 幅图像  $x_i$  存储在矩阵  $R^{N \times M}$  的第  $i$  列。同理可将整个人脸库中用于测试的图像存储在一个矩阵  $S^{N \times M'}$  中。其中第  $j$  ( $j = 1, 2, \dots, M'$ ) 幅图像  $x'_j$  存储在矩阵  $S^{N \times M'}$  的第  $j$  列。则能求出用于训练图像的平均脸  $\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i$ 。

**Step2** 根据用于训练的图像与平均脸的差值构造一个协方差矩阵

$$C = \frac{1}{M} \sum_{i=1}^M (x_i - \bar{x})(x_i - \bar{x})^T A A^T \quad (9)$$

**Step3** 求解图像投影的一个特征子空间。协方差矩阵  $C$  是一个  $N \times N$  的矩阵,由于  $N$  是一个很大的数,直接求解将协方差矩阵  $C$  的特征值与特征向量很困难,为了简化运算量,运用奇异值分解的方法即通过求解矩阵  $A^T A$  的特征值和特征向量来间接求解  $A A^T$ 。将  $A^T$  的特征值  $\lambda$  按照从大到小的顺序排序。定义前  $r$  个特征值与总的特征值的比值为特征值贡献率  $\Phi$ ,即  $\Phi = \sum_{i=1}^r \lambda_i / \sum_{i=1}^M \lambda_i$ 。选取特征值贡献率  $\Phi \geq P\%$  的前  $r$  个特征值以及对应的特征向量构建一个特征子空间。令  $\lambda_i$  ( $i = 1, 2, \dots, r$ ) 是矩阵  $A^T$  的前  $r$  个特征值,  $v_i$  ( $i = 1, 2, \dots, r$ ) 为前  $r$  个特征值对应的特征向量,则由奇异值分解定理知  $A A_i^T$  为:

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i \quad i = 1, 2, \dots, r \quad (10)$$

由此求的一个特征子空间  $U = (u_1, u_2, \dots, u_r)$ 。

**Step4** 将用于训练的图像与平均脸的差值向  $U^T$  中投影,得到相应的投影系数:

$$Y = U^T (x_i - \bar{x}) \quad i = 1, 2, \dots, M \quad (11)$$

**Step5** 将用于测试的图像与平均脸的差值向  $U^T$  中投影,得到相应的投影系数:

$$U^T Y' = U^T (x'_j - \bar{x}) \quad j = 1, 2, \dots, M' \quad (12)$$

**Step6** 将投影系数  $Y$  和  $Y'$  进行归一化,并且用归一化后的  $Y$  作为网络的输入训练 BP 神经网络。其中 BP 神经网络输入层神经元的数目由投影系数  $Y$  的维数确定。隐含层的神经元数目需要根据经验和多次测试来确定,输出神经元根据人脸数据库中人脸类别数目来决定。网络的权值则采用文中所提的权值调整方式。

**Step7** 根据训练好的网络对测试集人脸图像在特征子空间的投影系数进行分类得到正确识别的训练集图像个数  $N$  和测试集图像个数  $N'$ 。

**Step8** 最后通过计算用于训练图像正确识别个数与训练图像总数的比值和用于测试图像正确识别个数与测试图像总数的比值得到用于训练的图像识别率  $(N/M) \times 100\%$  和用于测试的图像识别率  $(N'/M') \times 100\%$ 。

## 3 实验结果

本文选择 ORL 人脸数据库进行人脸识别实验。该数据库共有 40 个人的人脸图像,每人有 10 幅  $112 \times 92$  图像,共 400 幅图像。人脸的面部表情和细节均有一定的变化,如笑或不笑,眼睛睁开或闭合、戴眼镜或不戴眼镜等;另外人脸的姿态也有很大的变化,其深度旋转和平面旋转可达 20 度,人脸尺寸也有最多的 10% 的变化。ORL 人脸数据库部分图像如图 2 所示。



图2 ORL 人脸数据库中部分人脸图像

由图像分辨率为  $112 \times 92$  可知,  $m = 112, n = 92, N = 10304$ 。选择库中 40 个人的图像。在每个人的人脸图像中随机选择 8 幅用于训练,另外 2 幅用于测试。则  $M = 320$ 。根据  $\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i$  可求出用于训练图像的平均脸。库中的图像运用主成分分析的方法提取主要特征,找到了一组特征脸子空间。将特征值按从大到小的顺序排序,则训练集图像的平均脸和第 1、2、20、60、150、200、240、280、310 个图像的特征值对应的特征脸图像如图 3 所示。

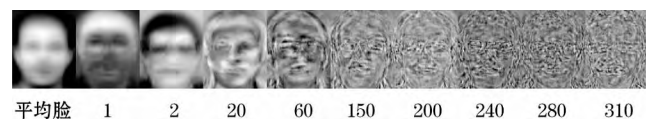


图3 训练图像平均脸及部分图像的特征脸

从上图可以看出较大特征值反映出图像的大体轮廓而较小的特征值,更多的反映出图像的细节部分。

利用前  $r$  个特征值对应的特征向量构造一个特征子空间,然后将用于训练的图像和测试的图像分别向该空间投影,得到相应的投影系数。将训练图像的投影归一化到  $[-1, 1]$  作为改进 BP 神经网络的输入,进行神经网络的训练,再对测试集上的人脸图像做相同的归一化处理,然后进行识别。本文中取隐含层神经元个数为 100,输出层神经元个数为 40,学习率为 0.03,动量系数  $mc = 0.5$ ,  $\alpha$  是满足  $\alpha \in (1/3, 1)$  的随机数,训练 1500 次,系数  $\alpha = 1.2$ ,系数  $\beta = 0.5$ 。图 4 显示了用传统 BP 神经网络做分类器和用改进 BP 神经网络做分类器时选取的特征向量个数  $r$  与测试集图像识别率(注:本文所提到的识别率都是将程序运行 100 次所求的平均识别率)的关系图。

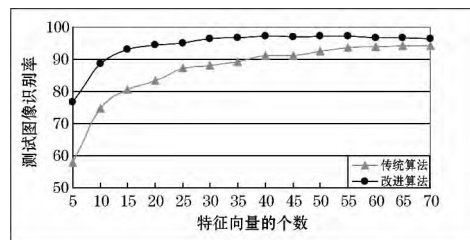


图4 测试图像的识别率与特征向量个数的关系

从图 4 中可以看出,运用本文所提出的改进 BP 网络的算法比传统算法具有更高的识别率,且图像的识别率随着特征向量个数的增加而增加。当特征向量个数为 40 的时候,改进算法

的识别率高达 97.15%。但在此后,图像的识别率随着特征向量个数的增加出现了小范围的上下波动。这说明了当特征向量个数增加到一定程度后,引入的一些特征可能并不是能够区分图像的主要特征,同时也说明图像的识别率并不与特征向量个数成正比,通过反复的实验可以找到使识别率最高的特征向量个数。另外在程序运行过程中,传统算法的识别部分平均运行时间为 23.4 s,利用本文所提的改进算法的识别部分平均运行时间为 1.3 s。运行结果显示本文所提的改进算法比传统算法具有更快的收敛速度和更高的识别率。

另外,随机选择人脸库中每个人的  $k(k=1, 2, \dots, 9)$  幅图像用于训练,另外  $10-k$  幅图像用于测试,选择特征值贡献率大于 85% 的特征值以及对应的特征向量。得到本文所提到的改进方法与传统方法做分类器时每类中用于训练图像的个数与测试图像上识别率的关系,如图 5 所示。

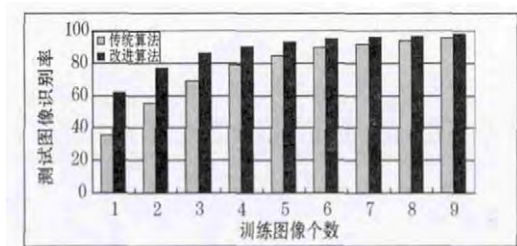


图5 训练图像个数与测试图像上识别率的关系

由图 5 可以看出,本文所提出的改进算法比传统算法具有更高的识别率,且用于训练图像个数对测试图像上识别率有很大的影响。随着每类中训练图像个数的增加,测试图像上的识别率不断增高。当人脸库中每类用于训练的图像个数为 1,而用于测试图像个数为 9 时,本文所提出的改进算法在测试图像上的识别率为 60.61%。网络的分类能力随着训练图像个数的增加不断增强。特别是当人脸库中每类用于训练的图像个数为 9,而用于测试图像个数为 1 时,测试图像上的识别率达到 97.7%。

## 4 结 语

通过主成分分析的方法提取人脸库中图像的主要特征数据,然后根据提取的图像数据,确定 BP 网络的参数,最后利用本文提出的权值学习方法改进的 BP 神经网络作为分类器对图像数据进行分类识别。仿真实验表明,通过使用本文提出的改进算法对 ORL 人脸库的图像进行识别,其实验结果比传统方法具有更快的收敛速度和更高的识别率。

## 参 考 文 献

- [1] Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: A survey[J]. //Proceedings of the IEEE, 1995, 83(5): 705-740.
- [2] Zhao W, Chellappa R, Rosenfeld A, et al. Face recognition: A literature survey[J]. ACM Computation Survey, 2003, 35(4): 399-458.
- [3] Turk M, Pentland A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.
- [4] Kirby M, Sirovich L. Application of the Karhunen-Loeve procedure for the characterization of human faces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(1): 103-108.
- [5] 邓楠. 基于主成分分析的人脸识别研究[D]. 西北大学, 2006.
- [6] 许丽艳. 人脸识别技术研究[D]. 天津大学, 2010.
- [7] 邵长斌. 基于神经网络的人脸识别技术研究[D]. 江南大学, 2009.

- [8] 罗炳海, 黄永武, 刘远飞. 基于 PCA-BP 神经网络的磨粒自动识别[J]. 润滑与密封, 2010, 35(6): 117-120.
- [9] 俞利强, 马道钧. 基于 PCA 技术的神经网络说话人识别研究[J]. 计算机工程与应用, 2010, 46(19): 211-213.
- [10] 陈先锋, 舒志兵, 赵英凯. “弹性”BP 神经网络在识别带有噪声字母中的应用[J]. 计算机仿真, 2005, 22(9): 153-155.
- [11] 李志清, 傅秀. 基于 PCA 的 3 种改进 BP 算法性能研究[J]. 计算机工程, 2011, 37(21): 108-110.
- [12] Robert Hecht-Nielsen. Theory of the Back Propagation Neural Network[C]. //Proceedings of the International Joint Conference on Neural Networks, 1989: 593-605.
- [13] 许宜申, 顾济华, 陶智, 等. 基于改进 BP 神经网络的手写字符识别[J]. 通信技术, 2011, 5(44): 106-109.
- [14] 袁曾任. 人工神经网络及其应用[M]. 北京: 清华大学出版社, 1999.
- [15] Riedmiller M, Braun H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm[C]. //Proceedings of the IEEE International Conference on Neural Networks. Piscataway (NJ): IEEE, 1993: 586-591.
- [16] 周德龙, 高文, 赵德斌. 基于奇异值分解和判别式 KL 投影的人脸识别[J]. 软件学报, 2003, 14(4): 783-788.

(上接第 157 页)

## 参 考 文 献

- [1] Briggs F, Raich R, Fern X Z. Audio Classification of Bird Species: A Statistical Manifold Approach[C]. //Proc. of the 9th IEEE International Conference on Data Mining. Miami, Florida, USA: [s. n.], 2009.
- [2] Vilches E, Escobar I A, Vallejo E E, et al. Data mining applied to acoustic bird species recognition[C]. //The 18th International Conference on Pattern Recognition. Hong Kong, China: IEEE Computer Society, 2006: 400-403.
- [3] Behnaz Ghoraani J, Sridhar Krishnan. Time-Frequency Matrix Feature Extraction and Classification of Environmental Audio Signals[J]. IEEE Transactions on Audio, Speech and Language Processing, 2011, 19(7): 2197-2209.
- [4] Lawrence Neal, Forrest Briggs, Raviv Raich, et al. Time Frequency Segmentation of Bird Song in Noisy Acoustic Environments[C]. //ICASSP 2011: 2012-2015.
- [5] 毕福昆, 边明明. 多分类器融合在语音情感识别中的应用[J]. 计算机工程与应用, 2010, 46(28): 205-208.
- [6] Sundararajan Rangachari, Philipos C Loizou, et al. A noise-estimation algorithm for highly non-stationary environments[J]. Speech Communication, 2006(48): 220-231.
- [7] 商琳, 杨育彬, 王亮, 等. 基于颜色共生矩阵的纹理检索算法 MCM[J]. 南京大学学报: 自然科学版, 2004, 40(5): 540-547.
- [8] 贾富仓, 李华. 基于随机森林的多谱磁共振图像分割[J]. 计算机工程, 2005, 31(10): 20-23.
- [9] Universitat Pompeu Fabra. Repository of sound under the creative commons license[DB/OL]. [2012-05-14]. <http://www.freesound.org/>.
- [10] 韩林峰, 赵晖. 基于支持向量机的联机手写维吾尔字符识别[J]. 计算机应用与软件, 2012, 29(3): 179-182.
- [11] 王爱平, 万国伟, 程志全, 等. 支持在线学习的增量式极端随机森林分类器[J]. 软件学报, 2011, 22(9): 2059-2074.