

文章编号: 1000-582X(2012)12-133-06

## 局部匹配的人脸识别方法

何光辉<sup>a</sup>, 张太平<sup>b</sup>

(重庆大学 a. 数学与统计学院; b. 计算机学院, 重庆 400044)

**摘要:** 从人类认知方式出发, 提出了一种基于统计学习的局部匹配人脸识别方法。该方法将人脸图像划分成若干小块, 各个子块中包含不同的人脸形状特征, 而不同的子块则描述了人脸主要部件之间的相对位置关系, 然后根据各个子块鉴别能力的差异, 将每个子块看成一弱分类器, 利用 Adaboost 学习算法组成一个强分类器, 提高最终的分类效果。实验结果表明该方法可以有效提高人脸的识别准确率并对人脸的表情和光照具有较好的鲁棒性。

**关键词:** 局部匹配; 统计学习; 人脸识别

**中图分类号:** TP391.4

**文献标志码:** A

## Face recognition method with local matching

He Guanghui<sup>a</sup>, Zhang Taiping<sup>b</sup>

(a. College of mathematics and Statistics;

b. Department of Computer Science, Chongqing University, Chongqing 400044, China)

**Abstract:** From human cognition, a face recognition method with local matching based on statistical learning is proposed. The image is divided into several subimages and each subimage is considered as a weak classifier. The Adaboost learning algorithm is used to train the weak classifiers and construct a strong classifier. As a result, all subimages are effectively combined together to explore the best discriminating power and improve the classification accuracy. Compared with the holistic matching methods, the local matching method is robust to variations in illumination, expression, and pose, etc. The experimental results show that the proposed method can improve the face recognition accuracy and is robust to variations in illumination and expression.

**Key words:** local matching; statistical learning; face recognition

人脸包含了很多的特征, 比如眼睛、鼻子、嘴、皮肤等特征, 还包括了 2 个眼睛之间的距离, 眼睛和嘴之间的距离, 甚至还有眼睛、鼻子、嘴之间的相对位置等等许多信息。把人脸包含的特征分成了两类: 形状特征(shape feature), 即眼睛、鼻子、嘴的形状等信息; 结构特征(structure feature), 即眼睛之间的距离, 眼睛和嘴之间的距离等信息。把人脸包含的特征分成了 2 类: 形状特征(shape feature), 即眼

睛、鼻子、嘴的形状等信息; 结构特征(structure feature), 即眼睛之间的距离, 眼睛和嘴之间的距离等信息。这 2 类特征是人脸识别最重要的判别依据。因此一个好的人脸识别算法必须同时利用这两种信息才能取得较好的识别率。人脸千差万别, 各自有各自的特点, 对不同的人来说, 不同的特征往往具有不同的判别能力, 不同的特征更是如此, 使得这两类特征在区分不同的人脸时的重要程度可能不一

**收稿日期:** 2012-06-12

**基金项目:** 中央高校基本科研业务费专项项目(0208005202019)

**作者简介:** 何光辉(1976-), 男, 重庆大学副教授, 博士, 主要从事模式识别、数据挖掘方向研究, (Tel)13594130031; (E-mail)ghhe@cqu.edu.cn。

样。从人在身份鉴别时的判别方法来看,人们在通过人脸进行身份识别是往往是通过人脸的最主要的特征来判别待识别人的身份,而不需要对面脸的所有特征进行判别。这说明人脸的各个特征具有不同的鉴别能力,在识别时的重要程度是不一样的。从信息融合的角度来讲,形状特征和结构特征属于不同类别的特征信息,如何找到一个合适的组合将这 2 种特征有机的结合起来是目前信息融合研究的热点,也是主要目的。

目前的人脸识别方法根据对面脸的图像的处理方法可以分为 2 类:一类是将人脸作为一个整体处理,进行特征提取以及模式匹配分类,称为整体匹配的方法;采用降维措施获得人脸图像最具有代表性的特征,著名的主成分分析 PCA<sup>[1-3]</sup>,线性判别分析 LDA<sup>[4]</sup>以及独立成分分析 ICA<sup>[5]</sup>都属于整体匹配方法。可以看出这类方法未严格地区分人脸的两类主要特征,将他们平等对待。因此,并没有将这两类特征有效地融合起来,发挥出他们最大的鉴别能力。然而,当人脸处于光照、表情、以及姿态等因素影响下,部分特征可能失去了应有的鉴别能力,基于整体匹配的方法往往不能取得满意的结果。

另一类是提取人脸的主要部件信息如眼睛、鼻子、嘴等作为输入,然后进行模式匹配<sup>[6-9]</sup>,这种方法称为局部匹配方法。他们是通过几何的手段找出人脸的主要特征,然后作为输入进行匹配识别。此类方法也仅考虑了人脸主要部件的形状特征,而忽视了主要部件间的相对位置,即结构特征。此外,人脸主要部件的定位与提取也是一件困难的事件。因此最近提出的非负矩阵因式<sup>[10]</sup>分解方法通过代数的手段(矩阵分解)来提取人脸的主要部件特征,然后进行分类匹配,但是他们没有区分提取的部件特征的重要性以及鉴别能力,而是将他们同等对待。

以上两类方法前者考虑了整个人脸图像的信息,或是主要考虑了人脸图像的结构特征,即人脸图像中每个特征之间的拓扑关系,但当人脸的中关键部位如眼睛、嘴在有表情时会产生很大的变化,此时提取出来的表示与自然表情下的表示有较大的差别。后者则主要考虑了特征信息,即重点处理眼睛、鼻子、嘴等特征信息但忽视了这些部件之间的拓扑关系,即结构特征,因此这两类方法均无法有效地将这两类特征有效地结合起来。

从人类认知方式出发,根据中国传统美术理论“三庭五眼”<sup>[11]</sup>,提出了一种基于统计学习的局部匹配人脸识别方法,实验结果表明该方法可以有效的提高人脸的识别准确率并对人脸的表情和光照具有

较好的鲁棒性。

## 1 人脸图像划分

在中国的美术理论中最基本的一个关于人脸结构的理论是“三庭五眼”。在面部正中作一条垂直的通过额部-鼻尖-人中-下巴的轴线;通过眉弓作一条水平线;通过鼻翼下缘作一条平行线。这样,两条平行线就将面部分成 3 个等分:从发际线到眉间连线;眉间到鼻翼下缘;鼻翼下缘到下巴尖,上中下恰好各占 1/3,称之为“三庭”;“五眼”是指是指眼角外侧到同侧发际边缘,刚好一个眼睛的长度,两个眼睛之间也是一个眼睛的长度,另一侧到发际边是一个眼睛长度(如图 1)。从图 1 中可以看出“三庭五眼”理论将人脸图像分为 15 个具有相同的竖直和水平距离的区域。这些区域中每个部分都包含了形状特征和结构特征,也就是说这些区域有些包含了形状特征,如眼睛、鼻子、嘴巴或者这些特征的一部分,也包含了结构特征,这个 15 个区域相对与每个人的人脸图像是位置固定的,亦即保持了结构特征。这个 15 个区域的划分方法如图 2。

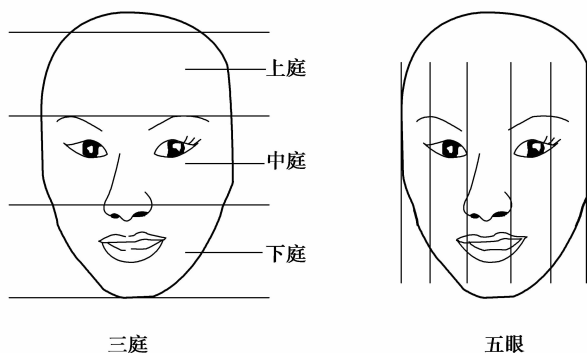


图 1 “三庭五眼”理论示意图

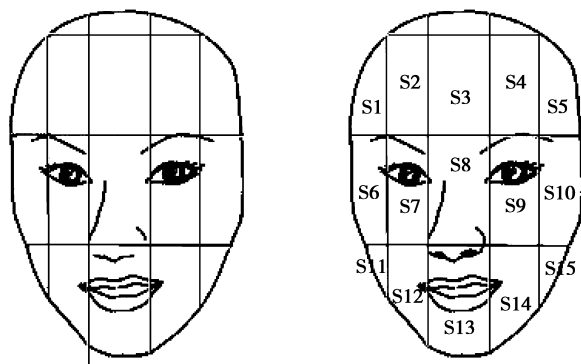


图 2 图像划分及其相应编号

对 AR<sup>[12]</sup> 数据库中部分人脸图像进行划分后的图像如图 3 所示。

设  $S$  是人脸图像,按照从左到右从上到下的方式

对子图像  $S^i$  进行编号  $i=1,2,\dots,15$ 。图 3 中上面一行是原始 AR 数据库中的人脸图像,下面一行是根据“三庭五眼”理论进行划分的 AR 数据库人脸图像。

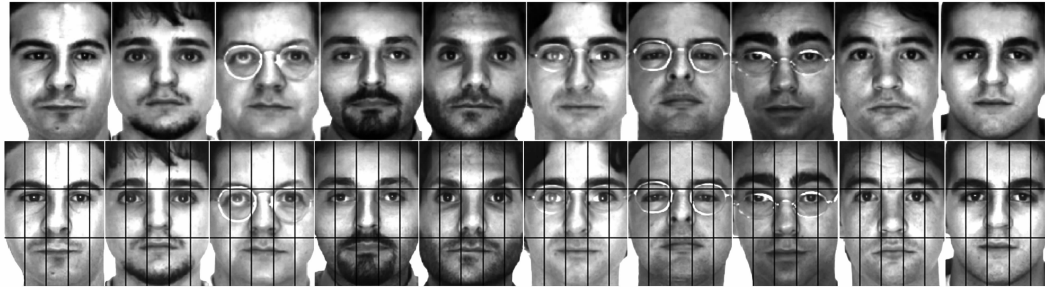


图 3 AR 数据库中人脸图像划分

## 2 Adaboost 算法思想

Boosting 是由 Schapire 于 1990 年提出来的<sup>[13]</sup>,是提高预测学习系统预测能力的有效工具,也是集成(组合)学习中最具代表性的方法,其将一组弱分类器训练成一个强分类器。Boosting 操纵训练样本以产生多个假设,从而建立通过投票组合的分类器集合。Adaboost 在训练样本上维护一套概率分布,在每一次迭代过程中 Adaboost 在每个样本上调整这种分布,每个弱分类器在训练样本上的错误识别率被计算出来并以此调整训练样本的概率分布。通过单个弱分类器的加权投票建立最终的强分类器。

AdaBoost 算法是 Boosting 家族的代表算法,该算法主要用来解决如下的 3 个基本的问题:

①从大量的原始特征中学习出最有效果的特征;

②利用被学习出的特征组建弱分类器;

③把这些弱分类器利用 boosting 方法组建成一个强分类器。

最基本的 AdaBoost 是用来解决 2 类分类问题的。一组  $N$  个有标记的训练样本被给出:  $(x_1, y_1), \dots, (x_N, y_N)$ , 其中  $y_i \in \{+1, -1\}$  是这些样本  $x_i \in \mathbf{R}^n$  的类别标记。AdaBoost 利用训练集学习出的最有效果的特征,并组成一个弱分类器  $h_m(x)$ , ( $m=1, 2, \dots, M$ )。根据训练样本的分布  $w_i^{(m)}$ , 利用  $M$  个弱分类器的线性组合,构建出一个强分类器,如式(1)

$$H_M(x) = \sum_{m=1}^M h_m(x), \quad (1)$$

Adaboost 是针对 2 类分类问题提出来的,其也可看成是在如下损失函数下的数据拟合的问题

$$L(y, f) = e^{-yf(x)}, \quad (2)$$

式中  $y_i \in \{+1, -1\}$  是数据的原始值,  $f(x)$  是数据拟合的值。根据式[14]中的描述,可以给出以下一个 AdaBoost 的算法流程。

算法 1 (AdaBoost 的算法流程)

1) 输入: 带标签的训练样本  $(x_1, y_1), \dots, (x_N, y_N)$ ,

2) 初始化分布权系数  $w_1(i) = 1/N, i=1, 2, \dots, N$ 。

3) 对  $t=1, 2, \dots, T$  进行循环计算:

a. 利用分布权系数  $w_t$  来训练弱分类器  $h_t$ ;

b. 计算错分率:  $e_t = \Pr_{i \sim w_t} [h_t(x_i) \neq y_i]$ ;

c. 计算:  $a_t = \frac{1}{2} \lg \left( \frac{1-e_t}{e_t} \right)$ ;

d. 更新权值:  $w_{t+1}(i) = w_t(i) \exp(-a_t y_i h_t(x_i))$ ;

e. 归一化权重  $w_{t+1}$ ;

4) 输出强分类器:  $H(x) = \text{sign} \left( \sum_{t=1}^T a_t h_t(x) \right)$

## 3 局部匹配的人脸识别

提出的局部匹配的人脸识别方法的基本思想是:将人脸划分后的各个子块看作为一个弱分类器,然后利用 Adaboost 算法将这些弱分类器训练成为一个强分类器,最终提高特征的鉴别能力。

Adaboost 算法是针对两分类问题提出来的,其不能直接用于多分类问题。为了解决多分类问题,需要做一些转换,第一种策略是采用支撑向量机方法解决多分类问题的方法,即 one-against-all 策略。它的思想是对每一个类构造一个分类模型,将多分类问题转换成多个 2 类分类问题来求解。对每一类样

本,其本身样本标记为正样本,其余的所有样本标记为负样本,这样即称为一个 2 分类问题。

对于一个  $C$  分类问题,需要求解  $C$  个二分类问题,这样带来的问题是计算量成倍地增加,但其仍然是按照算法 1 来执行。

另一种策略是对 2 个类构造一个分类模型<sup>[15]</sup>,对于一个  $C$  分类问题,需要训练  $C(C-1)/2$  分类模型来将每类样本与其他类样本分开。此方法(算法 2)与算法 1 的区别在于计算  $a_i$  的方式不同,其他部分都相同,在此算法中  $a_i$  的计算方式为

$$a_i = \frac{1}{2} \lg \left( \frac{1-e_i}{e_i} \right) + \lg(C-1). \quad (3)$$

可以看出算法 1 是此算法 2 的一个特例,当  $C=2$  时,  $\lg(C-1)=0$ 。与算法 1 中的  $a_i$  相比,算法 2 中的  $a_i$  增加了一个正数项  $\lg(C-1)$ ,此附加项为 Adaboost 多分类性能带来了很大的改善。在算法 1 中,为了保证  $a_i$  为正数,要求  $1-e_i > 1/2$ ,即要求每个弱分类器的精度大于  $1/2$ ,这对于大多数情况是很困难的,但在此多分类 Adaboost 中,只要求  $1-e_i > 1/C$  即可。这说明只要弱分类器的精度比随机赋值好即可,大大地降低了对弱分类器精度的要求。根据权重的更新规则,Adaboost 算法会在更新权重时使错分样本得到更大的权重。此策略不需要将多分类问题转换成多个 2 分类问题,而是直接求解多分类问题,于是计算量相对于算法 1 大为减少。此算法相当于 Adaboost 中的损失函数扩展为

$$L(y, f) = \exp \left( -\frac{1}{C} (y_1 f_1 + \dots + y_c f_c) \right), \quad (4)$$

$$\text{其中: } y_k = \begin{cases} 1, & k = c; \\ -\frac{1}{C-1}, & k \neq c. \end{cases}$$

然后要求  $f_1 + f_2 + \dots + f_c = 0$ 。可以看出当  $C=2$  时,与两分类 Adaboost 的损失函数是一样的。

于是,局部匹配的人脸识别方法的主要步骤能够描述如下。

局部匹配的人脸识别方法主要步骤:

1) 训练阶段

a. 将训练样本图像按照“三庭五眼”的思想划分成 15 块。

b. 在每个子块上利用线性判别分析(LDA)提取特征。

c. 利用 Adaboost 算法学习各个块特征的组合系数,将他们有效地组合起来。

2) 测试阶段

a. 对于测试样本图像按照“三庭五眼”的思想划分成 15 块。

b. 然后每个子块上利用线性判别分析(LDA)提取特征。

c. 利用训练阶段第 3 步学习的组合系数,将测试样本的各个子块特征组合起来。

3) 分类

利用相关的分类准则,对测试样本按照组合特征进行匹配分类。

## 4 实验结果及分析

在 AR 人脸数据库以及 ORL 数据库的识别结果,并与相关方法进行了比较。在实验中,使用前面所述的 2 种多分类 Adaboost 算法来学习弱分类器的组合。算法 1 表示使用前一节中所说的策略一来学习各个弱分类器(图像子块)的组合系数。算法 2 表示使用前一节中所说的策略二来学习各个弱分类器的组合系数。

### 4.1 实验结果

AR 数据集是一个包含超过 100 人的人脸数据集,每幅图像是在不同的光照以及不同表情下获取的,部分图像还有少量遮挡。本次实验中使用其中一个子集,包含了 50 个人总共 1 300 幅人脸图像。所有的图像被缩放到大小  $55 \times 40$  像素。在此数据集进行人脸识别的主要困难是变化的表情以及遮挡。

在 AR 数据库中,随机选择 13 副照片进行训练,剩余的 13 副照片作为测试样本。为了得到稳定的识别结果,进行了 50 次随机选择实验,50 次实验的平均值作为最终的识别结果,如表 1 所示。提出的局部匹配的人脸识别方法算法 1 在欧氏距离与余弦距离度量下分别取得了 92.13% 和 92.19% 的识别率,而算法 2 在这两种距离度量下分别取得了 97.52% 和 97.61% 的识别率。提出的局部匹配的人脸识别方法的识别结果超过了经典的线性判别分析(LDA)方法。

表 1 AR 数据库中算法识别结果比较

方法	欧氏距离度量	余弦距离度量
LDA	81.81	81.83
PLDA	87.21	87.31
算法 1	92.13	92.19
算法 2	97.52	97.61

ORL 人脸数据库包含了由 40 人组成的 400 幅人脸图像,每人有 10 幅图像,大小为  $92 \times 112$  像素。此数据集图像主要涵盖了如下变化:面部表情、面部

倾斜(不超过 20 度)、少许遮掩(眼镜或胡须等)、轻微光照、以及年龄变化。随机选择其中的 5 副图片作为训练样本,其余的 5 副作为测试样本。为了得到稳定的识别结果,进行了 50 次随机选择实验,50 次实验的平均值作为最终的识别结果,如表 2 所示。我们提出的局部匹配的人脸识别方法算法 1 在欧氏距离与余弦距离度量下分别取得了 95.40% 和 95.50% 的识别率,而算法 2 在这 2 种距离度量下分别取得了 98.50% 和 98.55% 的识别率。提出的局部匹配的人脸识别方法的识别结果相对于经典的线性判别分析提高了近 3%。

#### 4.2 结果分析

提出的局部匹配的人脸识别方法根据“三庭五眼”的思想将人脸划分成不同特征子块,每个子块包含的信息重要性不同,应该赋予不同的权值,但如果根据个人经验设置,很难找到一个最佳的权值,必然会影响实验结果。将每个子块看成一个弱分类器,然后通过 Adaboost 方法将这些弱分类器训练成一强分类器,实现各个子块间的最佳组合。通过 Adaboost 方法在训练集上学习出的人脸各部分分类器的融合权值如表 3。从表 3 中不难发现表中左右脸对应位置的权重不同,这与实际情况相符,因为人脸本来就不是很完全对称的。对人脸的识别的运行时间主要来自于 LDA,因此识别速度较快,可以进行实时识别。由于可以在识别之前对人脸进行预处理,可以进行标准化处理,因此该算法具有一定的旋转、平移不变性。

表 2 ORL 数据库中算法识别结果比较

方法	欧氏距离度量	余弦距离度量
LDA	92.80	92.75
PLDA	93.40	93.40
算法 1	95.40	95.50
算法 2	98.50	98.55

表 3 Adaboost 方法学习后的人脸划分权重

人脸划分	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$
对应权重	0.153 4	0.281 6	0.205 7	0.278 9	0.149 7
人脸划分	$S_6$	$S_7$	$S_8$	$S_9$	$S_{10}$
对应权重	0.105 2	0.875 1	0.910 3	0.864 8	0.097 4
人脸划分	$S_{11}$	$S_{12}$	$S_{13}$	$S_{14}$	$S_{15}$
对应权重	0.178 5	0.224 3	0.679 2	0.231 6	0.168 9

AR 数据库中的人脸图像主要受光照以及遮挡的影响,使得局部的特征失去了应用的鉴别能力。提出的局部匹配的人脸识别方法将人脸划分成若干子块,利用 Adaboost 算法,根据各个子块的鉴别能力,赋予不同的子块以不同的权,将各个子块有效地组合起来。在人脸主要受光照以及遮挡的影响时,方法可以根据光照以及遮挡对各个子块的影响差异,对于影响较大的子块赋予较小的权值,对于影响较小的子块或没受影响的子块以较大的权值。实验结果表明这种局部匹配的思想能够有效地解决人脸图像受光照以及遮挡的影响。ORL 数据库上的图像主要受姿态变化的影响,局部匹配的方法仍然取得了很好的结果,表明提出的局部匹配的人脸识别方法对光照、遮挡、以及姿态变化有一定的健壮性。

## 5 结 论

从人类认知方式出发,提出了一种基于统计学习的局部匹配人脸识别方法,该方法首先根据中国传统美术构图中“三庭五眼”的思想,将人脸图像划分成若干小块,各个子块中包含不同的人脸形状特征,而不同的子块则描述了人脸主要部件之间的相对位置关系,根据各个子块包含的鉴别能力的差异,利用 Adaboost 学习算法来将这些弱分类器组成一个强分类器,将各个子块(特征)有效地组合起来,提高最终的分类效果。实验结果表明该方法可以有效地提高人脸的识别准确率并对人脸的表情和光照具有较好的鲁棒性。在本实验过程中对大角度姿势识别效果不是很理想,因此在下一步的工作中将进一步研究大角度人脸姿势变化的识别问题。

#### 参考文献:

- [1] Yang J, Zhang D, Yang J Y. Constructing PCA baseline algorithms to reevaluate ICA-based face-recognition performance [J]. IEEE Transactions on Systems, Man, and Cybernetics, part B: cybernetics, 2007, 37(4): 1015-1021.
- [2] Ulfarsson M O, Solo V. Sparse variable PCA using geodesic steepest descent [J]. IEEE Transactions on Signal Processing, 2008, 56(12): 5823-5832.
- [3] Pang Y W, Tao D, Yuan Y, et al. Binary two-dimensional PCA [J]. IEEE Transactions on Systems, Man, and Cybernetics, part B: cybernetics, 2008, 38(4): 1176-1180.
- [4] Huang J, Yuen P C, Chen W H, et al. Choosing parameters of kernel subspace LDA for recognition of

- face images under pose and illumination variations[J]. IEEE Transactions on Systems, Man, and Cybernetics, part B: cybernetics, 2007,37(4): 847-862.
- [ 5 ] Kachenoura A, Albera L, Senhadji L, et al. Ica: a potential tool for bci systems [J]. IEEE on Signal Processing Magazine, 2008, 25(1): 57-68.
- [ 6 ] Bruneli R, Poggio T. Face recognition: feature versus templates[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(10): 1042-1052.
- [ 7 ] Li S Z, Hou X W, Zhang H J, et al. Learning spatially localized, parts-based representation[C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, December 8-14, 2001, Kauai, HI, USA. Piscataway: IEEE, 2001,1:207-212.
- [ 8 ] Heisele B, Ho P, Poggio T. Face recognition with support vector machines: global versus component-based approach[C]// Proceedings of the Eighth IEEE International Conference on Computer Vision, July 9-12, 2001, Vancouver, Canada. Piscataway: IEEE, 2001,2:688-694.
- [ 9 ] Ivanov Y, Heisele B, Serre T. Using component features for face recognition[C]// Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, May 17-19, 2004, Seoul, Korea. Piscataway: IEEE, 2004:421 - 426.
- [10] Zhang T P, Fang B, Tang Y Y, et al. Topology preserving non-negative matrix factorization for face recognition [ J ]. IEEE Transactions on Image Processing, 2008 , 17(4): 574-584.
- [11] <http://www.umedm.com/Article/2008/20080109231904-1294.html> [Z]
- [12] Martinez A M, Benavente R. The AR face database [EB/OL]. [2010-05-30]. <http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html>.
- [13] Schapire R E. The strength of weak learnability[J]. Machine Learning, 1990, 5(2):197-227.
- [14] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting: with discussion and a rejoinder by the authors[J]. Annals of Statistics, 2000,28(2):337-407.
- [15] Zhu J, Zou H, Rosset S, et al. Multi-class sdaboost[J]. Statistics and Its Interface, 2009,2:349-360.

(编辑 侯 湘)

~~~~~

(上接第 105 页)

- [ 7 ] 樊剑,袁涌,曾志和. 基于谐小波变换的地震波非平稳随机模型建立[J]. 华中科技大学学报:自然科学版, 2010,38(4):118-120.  
FAN Jian, YUAN Yong, ZENG Zhihe. Establishment of nonstationary stochastic model of earthquake record using harmonic wavelet transforms [J]. Journal of Huazhong University of Science and Technology: Natural Science Edition; 2010,38(4):118-120.
- [ 8 ] He J. An efficient numerical method for estimating reliabilities of linear structures under fully nonstationary earthquake [J]. Structural Safety, 2010, 32 ( 3 ): 200-208.
- [ 9 ] Cacciola P. A stochastic approach for generating spectrum compatible fully nonstationary earthquakes[J]. Computer & Structures, 2010,88(15/16):889-901.
- [10] Abbas A M, Manohar C S. Reliability-based vector nonstationary random critical earthquake excitations for parametrically excited systems [J]. Structural Safety, 2007,29(1):32-48.
- [11] Takewaki I. Critical envelope functions for non-stationary random earthquake input[J]. Computer & Structures, 2004,82(20/21):1671-1683.
- [12] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. The Royal Society, 1998, 454:903-995.
- [13] 楼梦麟,黄天立. 正交化经验模式分解方法[J]. 同济大学学报:自然科学版, 2007,35(3):293-298.  
LOU Menglin, HUANG Tianli. Orthogonal empirical mode decomposition[J]. Journal of Tongji University: Natural Science Edition,2007,35(3):293-298.
- [14] Husid R L. Analisis de terremotos: analisis general[J]. Revista del IDIEM, 1969, 8(1): 21-42.

(编辑 陈移峰)