

基于线性判别分析的室内声源定位方法

杨 悦, 顾晓瑜

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘 要: 在小信噪比和混响时间较长的恶劣环境下, 基于模式分类的手段能够有效克服传统的声源定位算法鲁棒能力不足的缺点, 其中朴素贝叶斯分类器定位的准确率高, 计算量小, 鲁棒能力强。在此基础上, 为了获得更好的定位性能, 提出使用线性判别分析(LDA)分类器进行声源定位。使用 Matlab 进行仿真, 截取声源信号的相位变换加权广义互相关函数(PHAT-GCC)作为特征向量, 通过投影变换, 找到最佳的特征空间来区分特征数据, 从而训练得到线性判别分析分类器。然后在不同的混响时间和信噪比的条件下, 进行定位测试, 比较了线性判别分析分类器和朴素贝叶斯分类器的性能。仿真结果表明, 在环境恶劣场合更宜使用线性判别分析分类器, 特别是混响严重时, 线性判别分析分类器的定位准确率比朴素贝叶斯分类器高 1% ~ 2%。

关键词: 声源定位; 相位变换加权广义互相关函数; LDA 分类器; 朴素贝叶斯分类器

中图分类号: TP301

文献标识码: A

文章编号: 1673-629X(2017)06-0187-04

doi: 10. 3969/j. issn. 1673-629X. 2017. 06. 039

Indoor Acoustic Source Localization Method with LDA

YANG Yue, GU Xiao-yu

(College of Telecommunications and Information Engineering, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China)

Abstract: The method based on pattern classification can overcome the deficiency of traditional acoustic source localization algorithms which has an insufficient robust ability in the harsh environment of small SNR and severe reverberation. Among them, Naive Bayes classifier has high location accuracy with a small amount of calculation and strong robustness. In order to achieve better localization performance, Linear Discriminant Analysis (LDA) classifier is adopted to locate acoustic source on the basis of former research. It has been tested by Matlab, while the Phase Transform Generalized Cross-Correlation (PHAT-GCC) function would be used as feature vector. LDA classifier has been trained through projection transformation which could help to find a better feature space to discriminate the feature data. Subsequently, the source would be located in different reverberation and noisy conditions to compare the performance with LDA classifier and Naive Bayes classifier. The simulation results have demonstrated that LDA classifier is a better choice in harsh environment and that the location accuracy of LDA classifier is higher than that of Naive Bayes classifier by 1% to 2%, especially in severe reverberation environment.

Key words: acoustic localization; PHAT-GCC; LDA classifier; Naive Bayes classifier

0 引 言

基于麦克风阵列的声源定位技术一直是国内外研究的一项热点。它的主要原理是: 通过具有一定几何拓扑结构的麦克风阵列采集声源信号, 通过阵列信号处理技术对语音信号进行处理和分析, 从而确定声源位置。基于麦克风阵列的声源定位技术已广泛应用于许多领域, 如视频会议、声音检测、语音增强、语音识别、说话人识别、智能机器人、监控监听系统、助听装置

等。传统的声源定位技术研究大体上分为三类: 基于最大输出功率可控波束形成技术、基于高分辨率谱估计技术、基于时延估计技术。其中, 基于广义互相关时延估计算法 (Generalized Cross-Correlation, GCC) 由于其计算量小、方法简单、容易实现, 得到了广泛应用^[1-3]。

由于房间墙壁的吸收, 声源发出的声波在各方向来回反射, 又逐渐衰减的现象, 称为室内混响。当声源

收稿日期: 2016-07-07

修回日期: 2016-10-13

网络出版时间: 2017-04-28

基金项目: 江苏省自然科学基金 (BK20140891); 声纳技术国防科技重点实验室开放研究基金 (KF201503)

作者简介: 杨 悦 (1989-), 男, 硕士研究生, 研究方向为语音处理、声源定位、机器学习。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170428.1703.058.html>

停止后,从初始的声压降到 60 dB 所需的时间称为混响时间,在这里用 T_{60} 表示室内声音的混响时间^[1-2]。室内混响的存在,对时延估计影响很大。因此,对于声源定位算法来说,如何提高抗噪声和抗混响能力是长久以来的研究重心,很多改进方法的提出对克服糟糕的环境影响有一定的作用,如对 GCC 进行加权处理^[4-5],但是在信噪比很小,混响很严重的情况下,基于传统算法的改进措施很难有明显的效果;此外当麦克风无法接收到声源的直达声时,也难以定位。

近年来,随着统计机器学习的发展,很多传统的问题有了新兴的解决方式^[6],声源定位领域也是如此。文献[7-8]提出了针对头相关函数(HRTF)特征提取并通过 K -Means 聚类 and 神经网络进行声源定位的方法;文献[9-10]使用朴素贝叶斯分类器,通过鉴别互相关函数的方法进行声源定位。这类方法能有效解决传统声源定位算法的缺点,不仅有很好的鲁棒性,而且计算效率高。

在此基础上,为了能够进一步提高在恶劣条件下声源定位的效果,提出采用 LDA(线性判别分析)分类器鉴别相位变换加权广义互相关函数来进行声源定位的方法,并对其进行了实验。

1 PHAT-GCC

PHAT-GCC(PHase Transform Generalized Cross - Correlation)方法是经典的基于时延的声源定位技术之一。轻量级的计算复杂度和易于实现的特点使其应用广泛。

由于来自同一声源的信号具有一定的相关性,通过计算不同麦克风接收到的信号的相关函数,由相关函数的峰值可以确定时延差,从而计算出声源的位置。相关函数的计算如式(1):

$$R_i(\tau) = \int_{-\infty}^{\infty} X_1(\omega) X_2^*(\omega) e^{j\omega\tau} d\omega \quad (1)$$

其中, $X_1(\omega)$ 为麦克风 1 接收到的信号 $x_1(t)$ 的傅里叶变换; $X_2^*(\omega)$ 为麦克风 2 接收到的信号 $x_2(t)$ 的傅里叶变换的复共轭。

然而在实际环境中,由于噪声和混响的影响,相关函数的最大峰会被弱化,有时还会出现多个峰值,给实际峰值检测带来了困难。因此,对于互相关函数法,常使用加权函数突出时延峰值来减小混响和噪声的影响,其中 PHAT(相位加权)能在噪声较小时有效地抗混响。相位加权广义互相关函数的计算如式(2):

$$R_i(\tau) = \int_{-\infty}^{\infty} \psi_{1,2}(\omega) X_1(\omega) X_2^*(\omega) e^{j\omega\tau} d\omega \quad (2)$$

$$\psi_{1,2}(\omega) = \frac{1}{\sqrt{X_1(\omega) X_2^*(\omega)}} \quad (3)$$

万方数据

其中, $\psi_{1,2}(\omega)$ 为相位加权,相当于白化滤波,改进后的互功率谱是一个纯相位函数,它的幅值为 1,因此 PHAT 加权是用相位的信息来求时延。

但当噪声增大时,信噪比较低的频率逐渐增多,PHAT-GCC 的性能逐渐下降,这种方法就不管用了。对于 PHAT-GCC 来说,当噪声和混响严重时,可能会出现多个虚假谱峰,但是在相同的室内环境下,同一位置声源引起的混响对谱峰的干扰是类似的,如果使用模式识别的方法,将带有混响的信号的 PHAT-GCC 作为特征,而非峰值检测的方法。由混响产生的虚假谱峰也能看成是一种特征,可以一定程度上减小混响的影响,在小信噪比的情况下定位性能不至于迅速下降,比传统的时延定位算法有更强的鲁棒性。

2 线性判别分析

LDA 是一种经典的线性学习方法,也称“Fisher 判别分析”,常常作为特征提取和降维算法用于人脸识别、人脸检测中^[11-13],也可直接作为分类器使用^[14-16]。其思想是将高维的样本投影到最佳鉴别矢量空间,以达到抽取分类信息和压缩特征空间维数的效果,使投影后的样本在新的子空间有最大的类间距离和最小的类内距离,从而能更好地分辨模式类别^[17]。LDA 投影示例如图 1 所示^[18]。

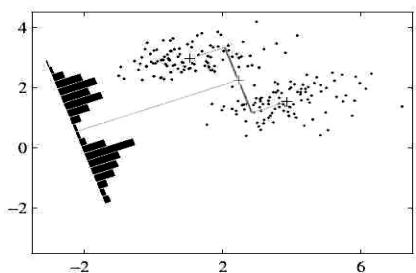


图 1 LDA 投影示例

对于二分类问题,设 $\mathbf{Y}_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$ 分别表示第 $i \in \{1, 2\}$ 类样本的集合、均值和协方差矩阵。若将样本数据投影到一条直线 \mathbf{w} 上,则两类样本的中心在直线上的投影分别为 $\mathbf{w}^T \boldsymbol{\mu}_1$ 和 $\mathbf{w}^T \boldsymbol{\mu}_2$,两类样本的协方差分别为 $\mathbf{w}^T \boldsymbol{\Sigma}_1 \mathbf{w}$ 和 $\mathbf{w}^T \boldsymbol{\Sigma}_2 \mathbf{w}$ 。为了使同类样本投影点尽可能近,可令同类样本投影点的协方差尽可能小,即 $\mathbf{w}^T \boldsymbol{\Sigma}_1 \mathbf{w} + \mathbf{w}^T \boldsymbol{\Sigma}_2 \mathbf{w}$ 尽可能小;为了使异类样本投影点尽可能远,可令类中心距离尽可能大,即 $\|\mathbf{w}^T \boldsymbol{\mu}_1 - \mathbf{w}^T \boldsymbol{\mu}_2\|_2^2$ 尽可能大。

则引入目标函数 J :

$$J = \frac{\|\mathbf{w}^T \boldsymbol{\mu}_1 - \mathbf{w}^T \boldsymbol{\mu}_2\|_2^2}{\mathbf{w}^T \boldsymbol{\Sigma}_1 \mathbf{w} + \mathbf{w}^T \boldsymbol{\Sigma}_2 \mathbf{w}} = \frac{\mathbf{w}^T (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \mathbf{w}}{\mathbf{w}^T (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{w}} \quad (4)$$

定义“类内散度矩阵”为:

$$S_w = \Sigma_1 + \Sigma_2 = \sum_{y \in Y_1} (y - \mu_1)(y - \mu_1)^T + \sum_{y \in Y_2} (y - \mu_2)(y - \mu_2)^T$$

(5)

定义“类间散度矩阵”为:

$$S_b = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

(6)

则式(3)可重写为:

$$J = \frac{\mathbf{w}^T S_b \mathbf{w}}{\mathbf{w}^T S_w \mathbf{w}}$$

(7)

通常可以使用拉格朗日乘子法和奇异值分解来求解参数 \mathbf{w} ,使得目标函数 J 最大,得到 $\mathbf{w} = S_w^{-1}(\mu_1 - \mu_2)$,即最佳投影方向。

计算出 LDA 模型参数后,即可将测试样本数据 \mathbf{y} 投影到 \mathbf{w} 上,得到投影结果 $Z = \mathbf{w}^T \mathbf{y}$ 。再由判决式(8)决定类别:

$$\begin{cases} Z > z_0 \Rightarrow y \in \text{class}_1 \\ Z < z_0 \Rightarrow y \in \text{class}_2 \end{cases}$$

(8)

$$z_0 = \frac{N_1 \mu_1 + N_2 \mu_2}{N_1 + N_2}$$

(9)

其中, N_1 和 N_2 分别是第一类训练样本和第二类训练样本的个数。

3 基于 LDA 分类器的声源定位方法

3.1 PHAT-GCC 特征提取

在房间中放置 2 个麦克风的麦克风阵,位于 l_s 的声源 S 到第 i 个麦克风的房间脉冲响应为 $h_i(l_s, t)$ ($i = 1, 2$)。设声源 S 发出的信号为 $s(t)$,则麦克风接收到的信号为 $x_i(t) = h_i(l_s, t) * s(t)$,其离散傅里叶变换为 $X_i(k)$ 。由式(2)可以求得长度为 K 帧的 PHAT-GCC: $R_j(\tau), j = 1, 2, \dots, K$ 。

对于任意一帧 $R_j(\tau)$,截取特征数据:

$$y_j \triangleq [R_j(-\tau_{\max}), R_j(-\tau_{\max} + 1), \dots, R_j(\tau_{\max} - 1), R_j(\tau_{\max})]^T$$

(10)

$$\tau = \tau_{\max} = \text{round}(\frac{\alpha D f_s}{c})$$

(11)

其中,round 为取整函数; D 为麦克风间距; f_s 为采样频率; c 为声速;因子 α 设为 1.67。

即截取 PHAT-GCC 谱线上所有可能取得谱峰的时延点的数据,转化为特征数据集,最终可以得到: $Y = [y_1, y_2, \dots, y_K]$ 。

3.2 训练 LDA 分类器模型

声源位于 L 个不同的方位角,即分类目标有多种类别,因此需要训练多分类 LDA 模型。对于 LDA 的多分类问题,可以直接从二分类问题的求解方式推广出来,也可以使用“One vs Rest”或“One vs All”方法^[17]将多分类任务^{互斥数据}并分成若干个二分类任务来实现。这里

通过直接多分类的方式求解模型。

对于 L 个类别的训练数据 $Y_i (i = 1, 2, \dots, L)$,使用 D 个投影向量 $\mathbf{W} = [w_1, w_2, \dots, w_D]$,对 K 帧训练数据进行投影,结果为 $Z = \mathbf{W}^T \mathbf{Y}_i$ 。此时类内散度矩阵由式(5)推广到:

$$S_w = \sum_{i=1}^L S_{w_i}$$

(12)

$$S_{w_i} = \sum_{y \in Y_i} (y - \mu_i)(y - \mu_i)^T$$

(13)

类间散度矩阵由式(6)推广到:

$$S_b = \sum_{i=1}^L N_i (\mu_i - \mu)(\mu_i - \mu)^T$$

(14)

其中, μ 为全体样本均值。

此时目标函数可选择:

$$J = \frac{\text{tr}(\mathbf{W}^T S_b \mathbf{W})}{\text{tr}(\mathbf{W}^T S_w \mathbf{W})}$$

(15)

其中, $\text{tr}(\cdot)$ 表示矩阵的迹。

与二类问题类似,同样使用拉格朗日乘子法,最大化 J ,即求解 $S_w^{-1} S_b \mathbf{W} = \lambda \mathbf{W}$,即可解出最佳投影方向 \mathbf{W}_{opt} 。多分类 LDA 模型训练完成。

3.3 定位测试

对于 K 帧的测试数据,可以使用单帧 PHAT-GCC 数据测试一个方位,测试 K 次;也可以同时使用多帧数据进行定位。

则当对训练出的 LDA 模型输入单帧测试数据 y 时,则根据投影后的结果 $Z_L = \mathbf{W}_{\text{opt}}^T \mathbf{y}$ 决定判定的类别。

LDA 分类器的声源定位流程见图 2。

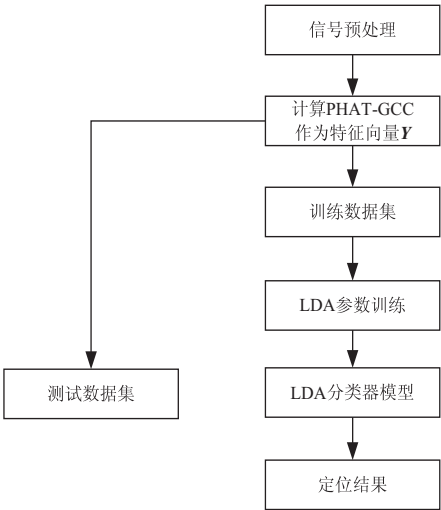


图2 LDA 分类器的声源定位流程

4 仿真实验

实验使用 Matlab R2014A 进行仿真。使用 Room-Sim 工具包^[19]模拟室内环境,使用 LDA 分类器在小信噪比和较严重的混响环境下进行定位,并将结果和朴素贝叶斯分类器的定位结果进行对比。

4.1 实验数据

仿真的房间尺寸为 8 m×6 m×3 m,使用 2 个麦克风组成麦克风阵,麦克风阵位于(4,2.5,1.5)处,麦克风间距 0.3 m;对接收到的信号分帧加汉宁窗,每帧 512 点,帧移 256 点,即每帧信号长度 32 ms,使用 100 帧训练数据 $Y_{train} = [y_1, y_2, \cdots, y_{100}]$ 和 400 帧测试数据 $Y_{test} = [y_{101}, y_{102}, \cdots, y_{500}]$;声源信号采用麻省理工学院 TIMIT 语音库的语音。声源位于以麦克风阵中点为圆心,半径为 2 m 的圆弧上,方向角为 10°、30°、50°、70°、90°、110°、130°、150°、170° 共 9 个位置,如图 3 所示。

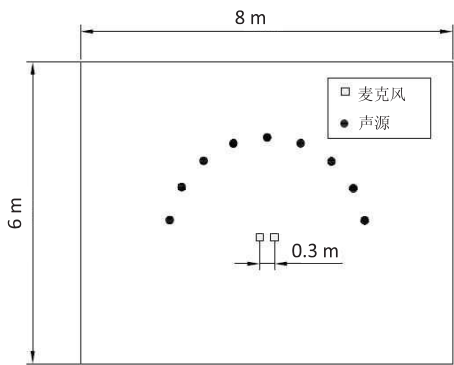


图 3 仿真的房间平面图

实验参数如下:中等的混响时间 $T_{60} = 300$ ms 和严重的混响时间 $T_{60} = 600$ ms;加性白噪声信噪比为 clean (无噪)、25 dB、20 dB、15 dB、10 dB 和 5 dB;特征属性使用单帧 PHAT-GCC。

4.2 结果与分析

在混响时间 $T_{60} = 300$ ms 和 $T_{60} = 600$ ms 这两种情况下,分别比较了 LDA 算法和朴素贝叶斯算法在 6 种信噪比情况下的定位准确率,如图 4 和图 5 所示。

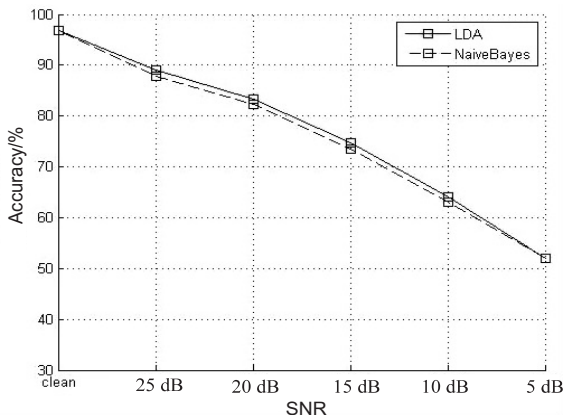


图 4 $T_{60} = 300$ ms 时的定位准确率比较

从实验结果中可以看出,定位准确率随着信噪比的降低而降低,随着混响的加强而降低。在 $T_{60} = 300$ ms 时,信噪比为 25 dB、20 dB、15 dB、10 dB 的情况下,使用 LDA 算法定位的准确率相比朴素贝叶斯算法有 1% 左右的提升。当 $T_{60} = 600$ ms 时,6 种信噪

比的情况下,使用 LDA 算法定位的准确率相比朴素贝叶斯算法都有明显提升,能提高 2% 左右的准确率。显然,混响严重的情况下更宜使用 LDA 分类器进行定位。

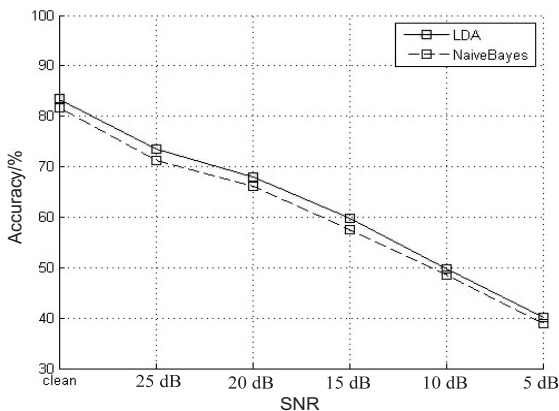


图 5 $T_{60} = 600$ ms 时的定位准确率比较

5 结束语

为了提高恶劣环境下声源定位的性能,提出了一种基于 LDA 分类器的声源定位方法,使用 PHAT-GCC 作为特征。实验中仅使用了单帧(32 ms) PHAT-GCC 作为特征属性,实际环境下可以使用多帧 PHAT-GCC 数据,能得到更高的准确率。实验结果表明,LDA 分类器在声源定位的性能上优于朴素贝叶斯分类器,特别是在混响严重的情况下有明显优势,显著改善了在恶劣环境下声源定位的效果。

参考文献:

[1] 万新旺. 基于阵列信号处理与空间听觉的声源定位算法研究[D]. 南京:东南大学,2011.

[2] 周 峰. 室内麦克风阵列声源定位算法研究和实现[D]. 上海:复旦大学,2009.

[3] 刘 超. 基于麦克风阵列的声源定位算法研究[D]. 南京:南京大学,2015.

[4] 崔玮玮,曹志刚,魏建强. 声源定位中的时延估计技术[J]. 数据采集与处理,2007,22(1):90-99.

[5] 夏 阳,张元元. 基于矩形麦克风阵列的改进的 GCC-PHAT 语音定位算法[J]. 山东科学,2011,24(6):75-79.

[6] 檀何凤,刘政怡. 模式分类方法比较研究[J]. 计算机技术与发展,2015,25(2):99-102.

[7] 马 浩,吴镇扬,张 杰,等. 与头相关传递函数的双耳特征提取与分类[J]. 电路与系统学报,2007,12(5):58-64.

[8] 马 浩,周 琳,胡红梅,等. 基于与头相关传递函数的神经网络声源定位方法研究[J]. 电子与信息学报,2007,29(9):2058-2062.

[9] Wan Xinwang, Wu Zhenyang. Sound source localization based on discrimination of cross-correlation functions[J]. Applied

3 所示。

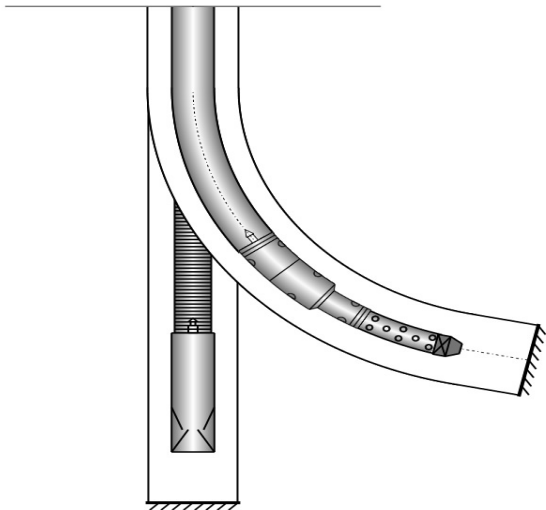


图 3 采用纹理贴图方法绘制的斜井管柱图

可以明显看出,图中斜井段所示的水力泵的每一点都与弧段轨迹的切线方向^[14]平行,与前后管柱工具图例良好地拼接在一起,实现了图例间的无缝拼接。

从实验结果的对比可以看出,采用一定的纹理贴图方法绘制的管柱图效果要明显好于矢量绘图法绘制的管柱图效果。

4 结束语

为改善管柱图绘制的视觉效果,针对目前在管柱图绘制过程中存在的问题,提出了一种管柱工具图例吸附于井眼轨迹的纹理贴图方法,有效解决了管柱图绘制过程中管柱工具图例在弯曲弧段处不能弯曲和管柱工具图例间的衔接出现断缝等问题。实验结果表明,与现在普遍采用的矢量绘图法相比,采用提出方法绘制的管柱图可以实现无缝衔接,具有更好的视觉效果。

下一步的研究重点是图库的扩充,需要完善的是增加管柱工具图例纹理图片库编辑功能,使用户可以按

照图例标准手动绘制纹理图片。

参考文献:

- [1] 姜翰青,王博胜,章国锋,等. 面向复杂三维场景的高质量纹理映射[J]. 计算机学报,2015,38(12):2349-2360.
- [2] 隋义勇,刘超,岳宝林,等. 特殊结构井管柱图绘制方法[J]. 油气田地地面工程,2013,32(7):28-29.
- [3] 柯新利. OpenGL 三维管线绘制中的坐标变换研究[J]. 地理空间信息,2008,6(3):53-55.
- [4] 王金辉,陈冰,王建庄. 实时图像仿射变换系统的研究与实现[J]. 机械与电子,2012,41(2):59-62.
- [5] 王魁生,刘蒙. 浅谈图像仿射变换的应用[J]. 信息技术与信息化,2015,41(1):155-156.
- [6] 肖源源,王子牛. 基于 VC++ 的 OpenGL 纹理映射技术的研究与实现[J]. 贵州大学学报:自然版,2008,25(2):158-160.
- [7] 张洁. 基于纹理映射真实感图形的研究[D]. 西安:西安电子科技大学,2007.
- [8] 江巨浪. 纹理映射技术的研究及实现[D]. 合肥:合肥工业大学,2003.
- [9] Liu Y, Lin W, Hayes J. Near regular texture analysis and manipulation[J]. ACM Transactions on Graphics, 2004, 23(3): 368-376.
- [10] 刘静. Photoshop 通道解析及应用[J]. 电子技术与软件工程,2013,24(19):93-94.
- [11] 周志宏,刘湘瑜. 石油管柱仿三维绘制软件研究[J]. 长江大学学报:自然科学版,2010,7(3):113-116.
- [12] 张文东,黄超,潘云燕. 基于 Flex 的管柱图绘制系统设计与实现[J]. 计算机工程,2012,38(21):237-240.
- [13] Lee S, Kim C, Chin S. An optimized sub-texture mapping technique for an arbitrary texture considering topology relations[J]. Network and Parallel Computing, 2012, 75(13):559-564.
- [14] Li J, Guo L, Qian Z, et al. Determination of the controlled landing trajectory of Chang'E-1 satellite and the coordinate analysis of the landing point on the Moon[J]. Chinese Science Bulletin, 2010, 55(13):1240-1245.
- [15] pain recognition based on LDA classifier[C]//International symposium on instrumentation & measurement, sensor network and automation. [s. l.]:[s. n.], 2012:380-382.
- [16] Zhang Y, Zhou X, Witt R M, et al. Automated spine detection using curvilinear structure detector and LDA classifier[J]. Neuroimage, 2007, 36(2):346-360.
- [17] 陈惠勤,骆德汉. 基于扩散映射和 LDA 的辛味中药材鉴别研究[J]. 计算机技术与发展, 2015, 25(5):192-195.
- [18] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016.
- [19] Bishop C. Pattern recognition and machine learning[M]. [s. l.]:Springer, 2007.
- [20] Campbell D. ROOMSIMtoolbox[EB/OL]. 2012-08-08. http://media.paisley.ac.uk/~campbell/Roomsim/.
- [21] Acoustics, 2013, 74(1):28-37.
- [22] Wan Xinwang, Liang Juan. Improved sound source localization using classifier in reverberant noisy environment[J]. Journal of Applied Science, 2013, 13(21):4897-4901.
- [23] 李道红. 线性判别分析新方法研究及其应用[D]. 南京:南京航空航天大学, 2004.
- [24] 王建国, 杨万扣, 郑宇杰, 等. 一种基于 ICA 和模糊 LDA 的特征提取方法[J]. 模式识别与人工智能, 2008, 21(6):819-823.
- [25] 谢永林. LDA 算法及其在人脸识别中的应用[J]. 计算机工程与应用, 2010, 46(19):189-192.
- [26] Mansor M N, Rejab M N, Syam S, et al. Automatically infant

(上接第 190 页)