

# 基于太赫兹时域光谱技术的红木检测方法

张文涛<sup>1,2</sup>, 王思远<sup>1,2</sup>, 占平平<sup>1,2</sup>, 韩莹莹<sup>1,2</sup>

<sup>1</sup>桂林电子科技大学电子工程与自动化学院, 广西, 桂林 541004;

<sup>2</sup>广西高校光电信息处理重点实验室, 广西, 桂林 541004

**摘要** 提出了一种基于随机森林预测模型的太赫兹时域光谱的木材鉴别方法。对 4 种木材(2 种红木、2 种非红木)在 0.2~1.2 THz 频率范围的吸收光谱的差异进行分析;对得到的光谱吸光度数据进行主成分分析的数据降维处理,并提取方差贡献率最高的五种主成分(总贡献率高达 99.65%);将其代入随机森林预测模型预测鉴别红木的真伪,得出相应训练集和测试集的认识率。实验结果表明,与传统的支持向量机预测模型和单一决策树模型比较,使用时域光谱技术结合随机森林预测模型能够得到更高的认识率,认识率可达 91.25%,能够准确对红木和非红木进行检测。

**关键词** 光谱学; 太赫兹时域光谱; 木材; 主成分分析; 随机森林预测模型; 检测

**中图分类号** O433.4 **文献标识码** A

**doi:** 10.3788/AOS201737.0230006

## Method of Identifying Red Wood Based on Terahertz Time-Domain Spectroscopy

Zhang Wentao<sup>1,2</sup>, Wang Siyuan<sup>1,2</sup>, Zhan Pingping<sup>1,2</sup>, Han Yingying<sup>1,2</sup>

<sup>1</sup>*Institute of Electrical Engineering and Automation, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China;*

<sup>2</sup>*Key Laboratory of Optoelectronic Information Processing of Guangxi Colleges and Universities, Guilin, Guangxi 541004, China*

**Abstract** A method to identify wood based on random forests prediction model of terahertz time-domain spectroscopy is proposed. We analyzed the differences of the absorption spectrum of four kinds of the woods (two kinds of real red wood and two kinds of false red wood) in the frequency range of 0.2~1.2 THz. Then the principal component analysis was applied to decrease the dimension of the spectral absorbance data, and the five principal components with top cumulative variance contribution rates were extracted (the total contribution rate is up to 99.65%). The processed spectral data were substituted into the random forest prediction model to identify real red wood and false red wood, and then the recognition rate of the training set and test set were obtained. The experimental results show that the terahertz time-domain spectroscopy combined with random forest prediction model can obtain a higher recognition rate, the recognition rate can reach 91.25%, when comparing with that using the traditional support vector machine prediction model and single decision tree model. The research results show that it is feasible to apply the terahertz time-domain spectroscopy combined with random forest prediction model into the identification of red wood.

**Key words** spectroscopy; terahertz time-domain spectroscopy; wood; principal component analysis; random forests prediction model; identification

**OCIS codes** 300.6495; 160.2290; 100.5010; 220.4840

## 1 引 言

太赫兹(THz)波通常指频率在 0.1~10 THz(波长 30  $\mu\text{m}$ ~1 mm)的电磁波。THz 波是介于毫米波与

**收稿日期:** 2016-10-06; **收到修改稿日期:** 2016-10-27

**基金项目:** 国家自然科学基金(61565004)、广西自然科学基金(2013GXNSFDA019002, 2014GXNSFGA118003)、桂林市科学研究与技术开发课题(20140127-1, 20150133-3)

**作者简介:** 张文涛(1976—),男,博士,教授,主要从事纳米计量及激光技术方面的研究。E-mail: uestczrk@126.com

红外光之间相当宽范围的电磁波区域,在电磁波波段中占有很特殊的位置。THz 有很多优越的特性,有着非常重要的学术和应用价值,是一种前沿的科技。THz 技术在医学诊断、健康监测、通信、化学和生物传感、以及安全质量控制等方面有着极其重要的应用,使得全世界各国都给予极大的关注<sup>[1-6]</sup>。由于很多种大分子物质振动能级或转动能级间距正好处于 THz 的频带范围,因此很多高分子聚合物在 THz 波段会有特殊的“指纹”特征<sup>[7]</sup>。THz 技术用于检测红木源于 THz 波具备独特的优越性和技术特点,主要归纳为以下两点: 1) THz 波很容易穿透非金属和非极性物质; 2) THz 波具有极低的光子能量,安全可靠<sup>[8-9]</sup>。对探测而得的 THz 信号进行光谱分析可获得物质光学特性参数,这些特征信息也是进行物质检测和识别的重要前提。

红木起源于明朝 1405 年,郑和下西洋每次回国前都会用红木压舱,回国后木匠们将红木做成家具供帝后享用<sup>[10]</sup>。红木有着极为悠久的历史,并且体现着中国的传统文化。近年来随着人们对生活品质要求的提高,红木种类的家具和收藏品在市场上广受青睐。2000 年 8 月 1 日,国家颁布实施了国家标准(GB/T 18107-2000),将 5 属 8 类的 33 种珍稀木材列为红木。原料的稀缺,导致红木的市场价格不断飙升,不少商家为了谋取巨大的利益弄虚作假,将有相似特征的木材冒充高档红木或掺杂其中。据中国消费者协会调查近年来红木的市场合格率仅为 30%。

常用的红木检测方法大多是先由具有丰富经验的专家通过专业显微仪器观察木材切片,再与已知的红木样品的切片标本作对比,这种方法比较浪费时间和人力。王克奇等<sup>[11]</sup>提出由光泽度和色度等物理指标鉴别木材;李艳艳等<sup>[12]</sup>提出基于气色相谱-主成分分析的方法鉴别红木;李敏华等<sup>[13]</sup>提出由手持显微镜与便携式计算机连接的方法对红木进行无损检测;杨忠等<sup>[14]</sup>提出对红木进行近红外光谱分析。由于木材是高分子的聚合物,且这些大分子在 THz 波段有明显的吸收特性。因此,红木与非红木在分子结构上会存在差异,不同的分子结构和组成就会造成 THz 波段不同的吸收峰位。所以本文利用 THz 时域光谱技术可以对红木与非红木木材进行快速检测。

## 2 实验部分

### 2.1 实验装置

实验使用桂林电子科技大学太赫兹实验室 Z-3 透射式 THz 光谱系统,如图 1 所示。飞秒激光光束经过  $\lambda/2$  波片由分束镜(PBS)分为抽运光和探测光,透过分束镜,能量较强的光线为抽运光。抽运光通过时间延迟装置后入射到发射晶体上激发 THz 脉冲,由离轴抛物镜 PM1 和 PM2 准直聚集到样品,另一对离轴抛物镜 PM3 和 PM4 将带有样品信息的 THz 波聚集到探测器上<sup>[15]</sup>。探测光经准直后照射到探测晶体上,探测晶体因电光效应引起 THz 电场发生变化,由光电二极管探测光偏振方向,产生的调制信号输入到锁相放大器后传输到计算机。实验所用激光器为 Toptica Photonics AG 公司超快飞秒激光器,激光中心波长为 780 nm,脉宽约为 100 fs,重复频率约为 80 MHz,输出平均功率约为 140 mW。发射晶体为低温生长的

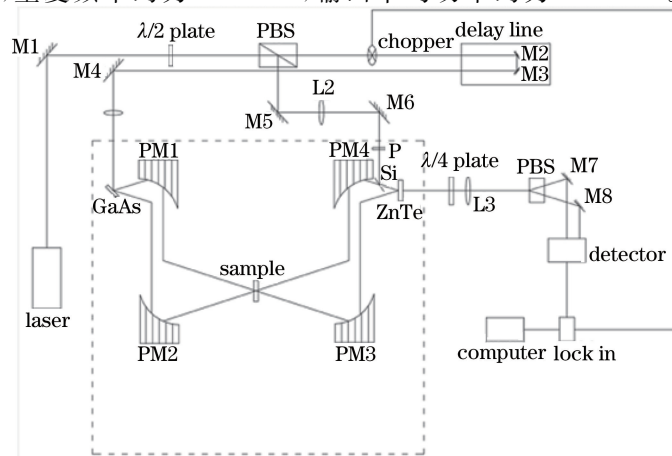


图 1 THz 时域光谱系统原理图

Fig. 1 Schematic of terahertz time-domain spectroscopy system

GaAs 晶体,探测晶体为 ZnTe 晶体。

## 2.2 样品制备和数据采集

实验选用 4 种木材分别为:檀香紫檀(*pterocarpus santalinus*, 紫檀属红木)、交趾黄檀(*dalbergia cochinchinenses*, 黄檀属红木)、虎斑木(*dracaena goldieana*, 虎斑木属非红木)、降真香(*acronychia pedunculata*, 山油柑属非红木),样品均取自中国林科院热林中心。采用压片法制备样品样本,将样品经粉碎机粉碎后经 200 目(75  $\mu\text{m}$ )筛网过滤后,每次取 0.2 g 经 6 t 压片机压制成厚度为 1 mm 左右(误差小于 0.05 mm),直径为 13 mm 内部均匀表面互相平行的片状样品。每种样品有 15 个样片,测量时每个样片得到 8 组数据,每 2 组数据取平均值作为一组实验数据,即每种样品获得 60 组实验数据。

实验前先将制备好的样品至于恒温干燥机干燥 1~2 h,实验时保证样品周围空气相对湿度低于 4%,温度为 23~25  $^{\circ}\text{C}$ 。每种样品取不同样片测试 60 组时域信号作为实验数据。记录并保存 4 种样品共 240 组数据。

## 3 理论介绍

### 3.1 主成分分析

主成分分析(PCA)是由 Karl Pearson 于 1901 年发明的一种分析简化数据集的技术<sup>[16]</sup>。主要应用于减少数据维数,同时保证数据集对方差贡献的最大特征,是通过保留低阶主成分忽略高阶主成分而做到的。主成分分析的主要思想是移动坐标轴,将  $n$  维特征映射到  $k$  维上( $k < n$ ),这  $k$  维是全新的正交特征,而不是简单地从  $n$  维特征中去除其余的  $n-k$  维特征。一般认为,当方差累计贡献率高于 85% 时,所选主成分可以包含大部分原始数据的信息。主成分分析的一般模型为

$$\begin{cases} Y_1 = \mu_{11} X_1 + \mu_{12} X_2 + \cdots + \mu_{1p} X_p \\ Y_2 = \mu_{21} X_1 + \mu_{22} X_2 + \cdots + \mu_{2p} X_p \\ \vdots \\ Y_p = \mu_{p1} X_1 + \mu_{p2} X_2 + \cdots + \mu_{pp} X_p \end{cases}, \quad (1)$$

式中  $X_1, X_2, \dots, X_p$  为样品的原始数据; $\mu_{1i}, \mu_{2i}, \dots, \mu_{pi}$  为  $X$  协方差矩阵的特征值对应的特征向量,可以理解为每个指标的权重; $Y_1, Y_2, \dots, Y_p$  为提取的  $p$  个主成分。上述模型还需满足条件如下:

1)  $Y_i$  与  $Y_j$  ( $i \neq j, i, j = 1, 2, 3, \dots, p$ ) 不相关;

2)  $Y_1$  是  $X_1, X_2, \dots, X_p$  的一切线性组合中方差最大的,  $Y_2$  是与  $Y_1$  不相关的  $X_1, X_2, \dots, X_p$  的一切线性组合中方差最大的,  $Y_p$  是与  $Y_1, Y_2, \dots, Y_{p-1}$  都不相关的  $X_1, X_2, \dots, X_p$  的一切线性组合中方差最大的;

3)  $\mu_{1i}^2 + \mu_{2i}^2 + \cdots + \mu_{pi}^2 = 1$  ( $i = 1, 2, 3, \dots, p$ )<sup>[17-18]</sup>。

### 3.2 随机森林

随机森林由 Breiman<sup>[19]</sup> 在 2001 年提出。顾名思义,随机森林是用随机的方式建立一个森林,森林由多棵决策树组成。它是一种多功能的机器学习算法,能够执行分类和回归的任务。当在基于某些属性对一个新的对象进行分类判别时,随机森林中的每一棵决策树都会被用来判断其归属分类,并由此进行“投票”,森林整体的输出结果将会是这些决策树输出结果的众数,其实质是对决策树算法的一种改进。在训练每棵树的节点时,使用的特征是从所有特征中按照一定比例随机地无放回地抽取的<sup>[20]</sup>。随机森林的具体实现过程如下:

1) 假设原始训练集为  $M$ ,应用重采样法有放回地随机抽取  $N$  个自助样本集,这  $N$  个样本集用来训练一棵决策树,作为决策树根节点处的样本;

2) 如果当前节点上达到终止条件(下一个节点选出的属性刚好是其父节点分裂时用过的属性),则设置当前节点为叶子节点。该叶子节点的预测输出为当前节点样本集中数量最多的那一类,整个决策树形成过程中无需剪枝;

3) 重复操作使得每个节点都训练过或者被标记为叶子节点,并保证每棵决策树都被训练过;

4) 将生成的多棵决策树组成随机森林,用随机森林分类器对新的数据进行判别与分类,分类结果按树分类器的投票多少而定。

将测得木材样品的 THz 光谱数据处理后,先应用主成分分析法对光谱数据降维,再将提取的主成分数

据代入到随机森林模型中,通过上述方法训练后的模型鉴别红木和非红木。得出对训练集和测试集数据的识别率。

## 4 实验结果分析

### 4.1 光谱分析

实验中每种样品获得 60 组测试数据,共 240 组。将测得每种样品原始实验数据取平均值并且整理后导入 Origin 软件得到 4 种木材的时域光谱图,如图 2(a)所示。4 条谱线较参考信号幅值都有所衰减并发生一定的峰位移动,幅值衰减是由于样品表面的散射和吸收造成的,峰位移动则是由于样品的厚度的差别以及 THz 波在样品的折射率不同引起。可以看到两种红木的振幅较高而且红移幅度较大。对测得的时域信号去回波后,进行快速傅里叶(FFT)变换可以得到相应的频域光谱,如图 2(b)所示。相比于参考信号,4 种木材的频谱强度有所下降。4 种样品对 THz 波都有吸收并且吸收强度和位置存在差异,两种红木的吸收大致相同并且略强度高于非红木。

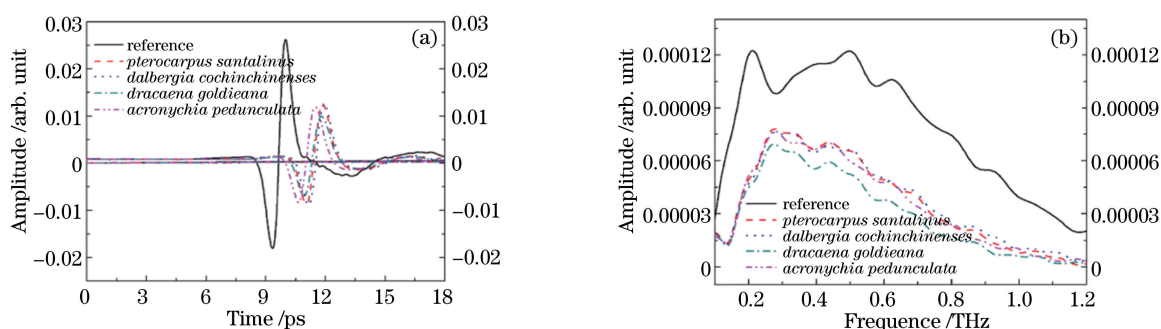


图 2 4 种木材的光谱图。(a)时域波形图;(b)频域光谱图

Fig. 2 Spectra of four kinds of wood. (a) Time-domain waveforms; (b) frequency-domain spectra

为了避免由于样品厚度不均匀造成的误差,进行进一步深入的分析,引入吸光度表示光波被材料吸收的程度<sup>[21]</sup>。吸光度是指光线通过物质前的入射光强与该光线通过物质后的透射光强度比值的对数,计算公式为

$$A = -\lg \frac{E_{\text{sam}}(\omega)^2}{E_{\text{ref}}(\omega)^2}, \quad (2)$$

式中  $A$  为吸光度,  $E_{\text{sam}}(\omega)$  为被测样品频域光谱的信号强度,  $E_{\text{ref}}(\omega)$  为参考信号频域光谱的信号强度,吸光度为无量纲。

处理后得到图 3(a)。时域光谱系统的有效范围是 0~3 THz,由于实验仪器光谱两端信噪比较低,取 0.2~1.2 THz 进行分析。可以发现 4 种样品在 0.3~1.2 THz 区间呈上升趋势。为了进一步观察,将吸光度图谱在 0.9~1.2 THz 进行放大,如图 3(b)所示。可以发现檀香紫檀与交趾黄檀两种红木样品在 1.17 THz 有吸收峰出现。但由于水蒸气在 1.165 THz 处会存在较强的吸收峰,如果实验干燥不够彻底或者样品在实

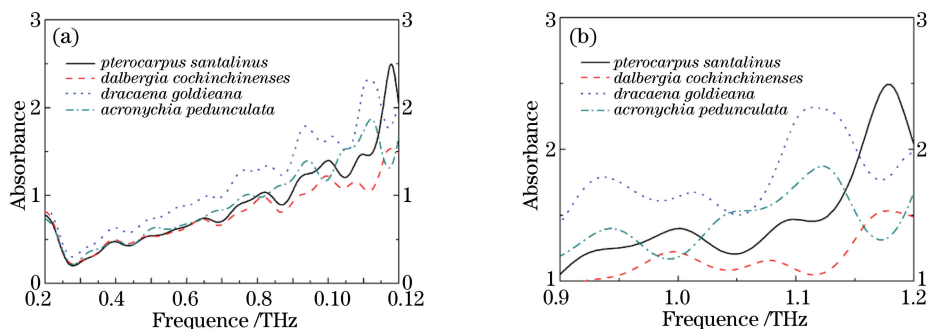


图 3 4 种木材吸收光谱图比较。(a)0.2~1.2 THz;(b)0.9~1.2 THz

Fig. 3 Comparison of the absorbance spectra of four kinds of wood. (a) in the range of 0.2-1.2 THz; (b) in the range of 0.9-1.2 THz



验过程中受到水蒸气影响,就会有吸收峰出现在该位置,所以该吸收峰可能不具备说服力。而且可以观察到虽然虎斑木和降真香两种非红木样品在 0.94 THz 和 1.13 THz 两个位置有吸收峰,但是吸收峰不明显,并且可以发现 4 种样品的吸收峰位置相近而且不易观察。吸收峰的出现有很多原因,如入射光与经样品表面的反射光发生干涉,或者是由于存在粒度较大的粒子发生散射,还有可能是温度设定范围内无法发生迁移等原因。于是为了结果更加精确,要通过数学建模来进一步鉴别红木与非红木。

#### 4.2 主成分分析结果

为了减少数据维度过大而造成误差,选用主成分分析的方法将提取的 0.2~1.2 THz 的吸光度  $240 \times 83$  的原始矩阵数据减少到  $240 \times 5$ (选取方差累计贡献率最高的 5 个主成分),各主成分的贡献率以及累计贡献率如表 1 所示。提取方差累计贡献率最高的 5 个主成分时累计贡献率已达到 99.65%,已经包含原始数据的绝大部分信息。

表 1 各主成分贡献率及累计贡献率

Table 1 Contribution rate of principal component and accumulation

Component	Variance /%	Cumulative rate /%
PC1	76.8252	76.8252
PC2	12.4294	89.2646
PC3	9.2547	98.5193
PC4	0.6388	99.1581
PC5	0.4945	99.6526

#### 4.3 随机森林模型鉴别结果

将主成分分析后的数据划分为两个数据集:训练集和测试集。抽取每种样品实验数据中的前 40 组作为训练集后 20 组为测试集。并且将檀香紫檀与交趾黄檀两种红木定义为标签“1”,虎斑木与降真香两种非红木定义为标签“2”,如表 2 所示。

表 2 木材样品数据分类及标签

Table 2 Data classification and sample label of the wood

Wood name (type)	Number of train data	Number of test data	Sample label
<i>Pterocarpus santalinus</i> (red wood)	40	20	1
<i>Dalbergia cochinchinenses</i> (red wood)	40	20	1
<i>Dracaena goldieana</i> (false red wood)	40	20	2
<i>Acronychia pedunculata</i> (false red wood)	40	20	2

将加好标签的实验数据代入随机森林预测模型中,提取贡献率最高的 5 个主成分数据作为识别的特征数据。分别得到训练集和测试集红木与非红木的识别率,并与单一决策树模型<sup>[22]</sup>,遗传算法(GA)优化的 SVM 模型以及粒子群优化(PSO)的 SVM 模型<sup>[23]</sup>的识别率作对比得到如表 3 所示的结果。这里选取的 SVM 的核函数为径向基核函数,迭代次数为 200 次。

表 3 4 种模型识别率及运行时间

Table 3 Recognition rate and running time of four kinds of models

Identification model	Recognition rate of train data /%	Recognition rate of text data /%	Comprehensive recognition rate /%	Identifying time /s
Random forest	100	91.25	97.08	0.36
Decision tree	99.375	76.25	91.67	0.39
GA SVM	100	88.75	96.25	10.55
PSO SVM	100	86.25	95.42	7.36

从表 3 中可以看出,随机森林模型的训练集识别率为 100%,测试集识别率为 91.25%,综合识别率为 97.08%,运行时间为 0.36 s。其他三种模型的测试集识别率均低于 90%。虽然 SVM 模型的综合识别率也达到了 95%以上,但是运行时间较长。由此得出结论,相比于其他三种模型,随机森林预测模型可以用更短的时间并且更加准确地对红木和非红木进行鉴别。

## 5 结 论

以取自中国林科院热林中心的两种常见红木以及两种非红木为研究对象。应用 THz 时域光谱技术,先宏观上对样品光谱进行观察分析,得到样品在光谱图上的差异。后将计算得到的吸光度数据(取 0.2~1.2 THz)进行主成分分析,提取贡献率达到 99.65% 的 5 个主成分。再代入随机森林的模型中预测鉴别,训练集识别率为 100%,测试集识别率为 91.25%,综合识别率为 97.08%,能够较好鉴别红木与非红木的种类。并与单一决策树模型以及 SVM 模型作比较,在识别率和识别时间上都有所提高。但是,由于木材是高分子的化合物,不宜进行仿真实验,且 THz 波照射时容易出现散射现象对鉴别会结果造成一定影响。这种通过 THz 时域光谱技术结合主成分分析以及随机森林的预测模型的方法,为木材种类鉴别提供了一种全新的思路。

## 参 考 文 献

- [1] Kang Shengwu, Wang Jiping, Liu Kan, *et al.* Analysis of the spectral characters of terahertz-wave[J]. *Acta Optica Sinica*, 2012, 32(6): 0612001.  
康胜武, 汪继平, 刘 侃, 等. 太赫兹波光谱特性分析[J]. *光学学报*, 2012, 32(6): 0612001.
- [2] Siegel P H. Terahertz technology in biology and medicine[C]. *MTT-S International Microwave Symposium Digest*, 2004: 1575-1578.
- [3] Zhang X C. Terahertz wave imaging: horizons and hurdles[J]. *Physics in Medicine and Biology*, 2002, 47(21): 3667.
- [4] Yao Jianquan, Chi Nan, Yang Pengfei, *et al.* Study and outlook of terahertz communication technology[J]. *Chinese J Lasers*, 2009, 36(9): 2213-2233.  
姚建铨, 迟 楠, 杨鹏飞, 等. 太赫兹通信技术研究与发展[J]. *中国激光*, 2009, 36(9): 2213-2233.
- [5] Sun Yaru, Shi Tonglu, Liu Jianjun, *et al.* Terahertz label-free bio-sensing with EIT-like metamaterials[J]. *Acta Optica Sinica*, 2016, 36(3): 0328001.  
孙雅茹, 史同璐, 刘建军, 等. 太赫兹超材料类 EIT 谐振无标记生物传感[J]. *光学学报*, 2016, 36(3): 0328001.
- [6] Ferguson B, Zhang X C. Materials for terahertz science and technology[J]. *Nature Materials*, 2002, 1(1): 26-33.
- [7] Li N, Shen J L, Sun J H, *et al.* Study on the THz spectrum of methamphetamine[J]. *Optics Express*, 2005, 13(18): 6750-6755.
- [8] Mickan S, Abbott D, Munch J, *et al.* Analysis of system trade-offs for terahertz imaging[J]. *Microelectronics Journal*, 2000, 31(7): 503-514.
- [9] Shen Jingling, Zhang Cunlin. Terahertz nondestructive imaging technology and its application[J]. *Nondestructive Testing*, 2005, 27(3): 146-147.  
沈京玲, 张存林. 太赫兹波无损检测新技术及其应用[J]. *无损检测*, 2005, 27(3): 146-147.
- [10] 王 露. 正确解析红木标准——访红木流通专业委员会秘书长车畅[J]. *中国市场*, 2012 (34): 16-17.
- [11] Wang Keqi, Yang Shaochun, Dai Tianhong, *et al.* Research on wood classification using uniform color space[J]. *Computer Engineering and Design*, 2008, 29(7): 1780-1784.  
王克奇, 杨少春, 戴天虹, 等. 基于均匀颜色空间的木材分类研究[J]. *计算机工程与设计*, 2008, 29(7): 1780-1784.
- [12] Li Yanyan, Sun Duoyong, Zhu Zhongliang, *et al.* Study on the classification and recognition of mahogany based on gas chromatograph-principle component analysis[J]. *Computers and Applied Chemistry*, 2010, 27(2): 237-240.  
李艳艳, 孙多永, 朱仲良, 等. 基于气相色谱-主成分分析的红木分类识别方法研究[J]. *计算机与应用化学*, 2010, 27(2): 237-240.
- [13] Li Minhua, Liu Hongqing, Li Guilin, *et al.* A study on wood non-destructive testing method of hongmu furniture and crafts[J]. *Furniture & Interior Design*, 2013(9): 96-98.  
李敏华, 刘红清, 李桂兰, 等. 红木家具与工艺品木材无损检测方法研究[J]. *家具与室内装饰*, 2013(9): 96-98.
- [14] Yang Zhong, Jiang Zehui, Lü bin. Investigation of near infrared spectroscopy of rosewood[J]. *Spectroscopy and Spectral Analysis*, 2012, 32(9): 2405-2408.  
杨 忠, 江泽慧, 吕 斌. 红木的近红外光谱分析[J]. *光谱学和光谱分析*, 2012, 32(9): 2405-2408.
- [15] Chen Tao, Li Zhi, Mo Wei. Identification of terahertz absorption spectra of explosives based on fuzzy pattern recognition[J]. *Chinese Journal of Scientific Instrument*, 2012, 33(11): 2480-2486.  
陈 涛, 李 智, 莫 玮. 基于模糊模式识别的爆炸物 THz 光谱识别[J]. *仪器仪表学报*, 2012, 33(11): 2480-2486.
- [16] Liu X, Chen X, Wu W, *et al.* Process control based on principal component analysis for maize drying[J]. *Food control*, 2006, 17(11): 894-899.

- [17] Chen Yang, Zhang Taining, Guo Peng, *et al.* Quantitative analysis for nonlinear fluorescent spectra based on principal component analysis[J]. *Acta Optica Sinica*, 2009, 29(5):1285-1291.  
陈 扬, 张太宁, 郭 澎, 等. 基于主成分分析的复杂光谱定量分析方法的研究[J]. *光学学报*, 2009, 29(5): 1285-1291.
- [18] Wu Yiquan, Zhou Yang, Long Yunlin. Small target detection in hyperspectral remote sensing image based on adaptive parameter SVM[J]. *Acta Optica Sinica*, 2015, 35(9): 0928001.  
吴一全, 周 杨, 龙云淋. 基于自适应参数支持向量机的高光谱遥感图像小目标检测[J]. *光学学报*, 2015, 35(9): 0928001.
- [19] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [20] Cai Jiabin, Feng Guocan, Tang Xin, *et al.* Human action recognition based on local image contour and random forest [J]. *Acta Optica Sinica*, 2014, 34(10): 1015006.  
蔡加欣, 冯国灿, 汤 鑫, 等. 基于局部轮廓和随机森林的人体行为识别[J]. *光学学报*, 2014, 34(10): 1015006.
- [21] Tu Shan, Zhang Wentao, Xiong Xianming, *et al.* Principal component analysis for transgenic cotton seeds based on terahertz time domain spectroscopy system[J]. *Acta Photonica Sinica*, 2015, 44(4): 0430001.  
涂 闪, 张文涛, 熊显名, 等. 基于太赫兹时域光谱系统的转基因棉花种子主成分特性分析[J]. *光子学报*, 2015, 44(4): 0430001.
- [22] Ryniec R, Zagrajek P, Pałka N. Terahertz frequency domain spectroscopy identification system based on decision trees [J]. *Acta Physica Polonica A*, 2012, 122(5): 891-895.
- [23] Alba E, García-Nieto J, Jourdan L, *et al.* Gene selection in cancer classification using PSO/SVM and GA/SVM hybrid algorithms[C]. *IEEE Congress on Evolutionary Computation*, 2007: 284-290.