

doi:10.3788/gzxb20154404.0430001

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

涂闪, 张文涛, 熊显名, 陈涛

(桂林电子科技大学 电子工程与自动化学院 广西高校光电信息处理重点实验室, 广西 桂林 541004)

摘 要:运用太赫兹时域光谱技术对 8 种转基因棉花种子进行了探测研究. 测量了转基因棉花种子在 0.3~1.5 THz 频率范围的特征吸收谱, 基于光谱分析, 对比了 8 种转基因棉花种子吸收光谱的异同, 并采用主成分分析法对其做了定性分析, 得到了光谱分析的相关数据. 提取了三种主成分因子, 前三种主成分的累计方差贡献率达到了 71.537%, 从主成分得分图中可以明显区分不同类别的转基因棉花种子. 太赫兹时域光谱技术能够实现转基因棉花种子的无损检测, 在农业选种、农业安全等领域将有广阔的应用价值.

关键词:光谱学; 太赫兹时域光谱; 转基因; 吸光度; 主成分分析; 无损检测

中图分类号: O433.4

文献标识码: A

文章编号: 1004-4213(2015)04-0430001-6

Principal Component Analysis for Transgenic Cotton Seeds Based on Terahertz Time Domain Spectroscopy System

TU Shan, ZHANG Wen-tao, XIONG Xian-ming, CHEN Tao

(Guangxi Colleges and Universities Key Laboratory of Optoelectronic Information Processing, Institute of Electrical Engineering and Automation, Guilin university of Electronic Technology, Guilin, Guangxi 541004, China)

Abstract: Terahertz time-domain spectroscopy system was used to measure the characteristic absorption spectra of 8 transgenic cotton seeds in the frequency range from 0.3 to 1.5 THz. Comparing the similarities and differences of THz absorption spectrum of 8 kinds of transgenic cotton seeds, which based on spectral analysis, and used principal component analysis to perform qualitative analysis, the relevant data of spectral analysis was obtained. Three principal component factors were extracted in this paper, of which the cumulative percentage variance reached 71.537%. Therefore, different categories of transgenic cotton seeds were clearly distinguished according to principal component analysis score chart. The experimental results indicated that nondestructive testing of transgenic cotton seeds could be achieved by using Terahertz time-domain spectroscopy system technology, which could be widely applied in the fields of agricultural seed selection, agricultural security and so on.

Key words: Spectroscopy; Terahertz time domain spectroscopy; Transgenic; Absorbance; Principal component analysis; Non-destructive testing

OCIS Codes: 300.6495; 160.1435; 220.4840; 300.1030

0 引言

(波长为 3 000~30 μm)之间的电磁波, 处于电子学向光子学过渡的特殊区域, 具有非常重要的学术价值和
太赫兹 (THz) 辐射是指频率在 0.1~10.0 THz 应用价值^[1]. 近年来人们发现, 物质处于远红外波段的

基金项目:国家自然科学基金(No. 11064002)、广西自然科学基金(Nos. 2013GXNSFDA019002, 2014GXNSFGA118003)、广西科学研究与技术开发项目(No. 桂科合 1346010-5)、广西教育厅重点项目(No. ZD2014057)、桂林市科学研究与技术开发项目(Nos. 20130122-1, 20140127-1)和广西自动检测技术与仪器重点实验室基金项目(No. YQ15104)资助

第一作者:涂闪(1989-), 男, 硕士研究生, 主要研究方向为 THz 物质检测. Email: tushan0821@126.com

导师(通讯作者):张文涛(1976-), 男, 教授, 博士, 主要研究方向为光电检测、纳米计量及激光技术. Email: uestczrk@126.com

收稿日期:2014-10-14; **录用日期:**2014-12-29

<http://www.photon.ac.cn>

THz 光谱具有反映物质结构的“指纹”特征. THz 光谱对应着分子内的低频集体振动、分子间弱相互作用及晶格的声子振动模式等,在这一频段表现出强烈的吸收和色散特性,使得 THz 光谱对物质的成分含量分析成为可能;同时还表现出太赫兹波的光谱分辨率特性,使得太赫兹探测技术,特别是太赫兹光谱成像技术,不但能够辨别物体的形貌,而且能够鉴别物体的组成成分^[2]. 而且 THz 光子能量只有毫电子伏 (meV) 量级,与 X 射线相比,不会因为电离而破坏被检测的物质. THz 波的无损检测 (Non-Destructive Testing, NDT) 特性对生物大分子的检测显得尤为重要. Gen T^[3] 等人提出了衰减全反射 (Attenuated Total Reflection, ATR) 法提取药物的太赫兹光谱信息; Ge^[4] 等人提出了主成分分析-支持向量机 (Princial Component Analysis-Support Vector Machine, PCA-SVM) 法研究了玉米的太赫兹光谱特性; Wang^[5] 等人利用偏最小二乘 (Partial Least Squares, PLS) 法、常规间隔偏最小二乘 (interval PLS, iPLS) 法、反向常规间隔偏最小二乘 (Backward interval PLS, biPLS) 法和移动窗口偏最小二乘 (mwPLS) 法分别算出噻菌灵的太赫兹光谱的均方根误差,并作出了比较;陈涛^[6] 等利用模糊识别的方法研究了爆炸物太赫兹光谱.

转基因^[7] 棉花技术是指在棉花亲本的 DNA 序列中引入外援基因片段,目前用于棉花的外源基因中主要有抗虫基因^[8] 和动物角蛋白基因^[9],使得棉花产量得到了极大的提高.

THz 时域光谱技术在农业选种中有很大的应用空间,生物大分子的振动和转动频率位于太赫兹波段,运用太赫兹光探测种子能与其内部大分子产生共振,棉花种子内部生物大分子主要包含蛋白质分子^[10-11]、多糖^[12] 和油脂^[13]. 氨基酸在 THz 波段的吸收是由分子转动或扭动造成的,不同基因片段转义的氨基酸结构不同,氨基酸的不同结构造成了它们在 THz 波段不同的吸收峰位^[14],从而在特定的频率处出现吸收峰. 本文采用 8 种不同转基因棉花种子作为探测对象,运用太赫兹光谱技术结合主成分分析 (Principal Component Analysis, PCA) 方法得到了很好的鉴别和归类不同转基因棉花种子的效果.

1 材料和方法

1.1 实验设备

如图 1,采用 45° 入射发射极的透射式太赫兹系统. 泵浦和探测太赫兹所用的激光器是 Toptica Photonics AG 公司的超快飞秒激光器,激光器中心波长约为 780 nm,脉宽约为 100 fs,重复频率约为 80 MHz,激光器平均输出功率约为 140 mW. 激光束作为光源被 $\lambda/2$ 波片分为泵浦光和探测光,透过分束镜

的一束作为泵浦光,泵浦光经反射镜反射后透过斩波器 (光束平均功率约为 300 mW),经过两平面镜和机械平台组成的延迟器,再经过反射和聚焦后照射在太赫兹发射源上,激光束照射偏置光电导砷化镓 (GaAs) 天线 (透射模式) 后激发出太赫兹脉冲,太赫兹脉冲经准直后被离轴抛物面反射镜聚焦,照射在样品上,样品放置在离轴抛物面反射镜焦点处;由分束镜得到的另一束光为探测光,探测光经准直和聚焦照射在碲化锌 (ZnTe) 晶体 (一块约 2 mm 厚,另一块约 0.8 mm 厚) 上,通过电光取样探测太赫兹电场^[15]. 太赫兹波照射光电二极管调制探测光,改变探测光的偏振方向,产生的调制信号输入到锁相放大器. 其中,带有 USB 通讯的 LabVIEW 程序控制扫描台 (步长为 10 μm),锁相放大器 (时间常量为 100 ms) 采集太赫兹波信号.

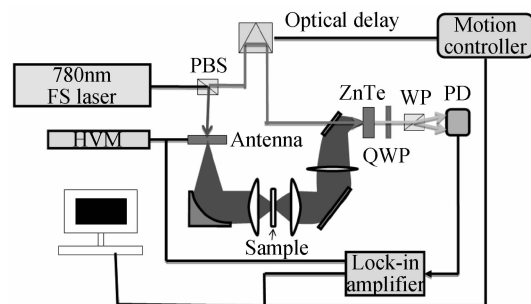


图 1 太赫兹时域光谱系统原理

Fig. 1 Schematic diagram of THz-TDS

为了消除探测空间中水分的影响,将系统中有太赫兹波的部分置于干燥空气罩中,样品周围的湿度小于 3%. 实验温度维持在 22°C ~ 25°C.

实验采用美国 Zomega 公司的 Z-3 太赫兹时域光谱系统 (Terahertz time-domain spectroscopy system, THz-TDS),光谱范围为 0.1 ~ 3.5 THz,频谱分辨率小于 5 GHz (FFT 后),最大延迟时间大于 1.3 ns,动态范围峰值大于 70 dB,数据采集时间约为 1 min,有透射和反射两种模式,本文采用 Z-3 系统的透射^[16] 模式,对应上位机界面为 Labview 编译构成.

1.2 光谱分析

八种转基因棉花种子的太赫兹时域光谱如图 2 (a). 参考信号定义为不放置样品时在自由空间中测得的太赫兹信号,用 Reference 表示. 8 种谱线相对 Reference 都发生了不同程度的频率红移^[17]. 原始数据点数较少 (仅 640 点),为了使 FFT 变换后曲线更光滑,需先对原始数据补零 (达到 32768 点). 图 2 (b) 为经过 FFT 变换后的频域光谱,八种样品频谱的强度相对 Reference 都有不同程度的降低,这是由于样品本身对太赫兹波的吸收所致;水对太赫兹波具有强烈吸收,为减少水蒸气对太赫兹波的吸收,测试环境选择在干燥的空气中;比较样品的频谱与 Reference 的频谱可以看出八种样品的吸收截止频率是不同的,鑫秋-k638

(xinqiuk638)截止频率高于 1.5 THz,其他 7 种样品的截止频率均在 1.1 THz 左右.光谱数据均是通过在每个样品扫描 3 次,取平均值所得,光谱数据的振荡是由偶然误差引起的.由于所测样品均是在相同环境下制备的,且样品厚度均为 1.000 mm 左右,误差小于 0.050 mm,不同的只是转基因棉花种子的型号不同,由此可以说明不同种类转基因棉花种子对太赫兹波的吸收是不同的.

为了进一步深入的分析,引入公式

$$\frac{E_{\text{trans}}(\omega)}{E_0(\omega)} = A \exp(-i\varphi(\omega)) \quad (1)$$

$E_0(\omega)$ 是太赫兹参考信号的幅度, $E_{\text{trans}}(\omega)$ 是透射太赫兹信号的幅度, $\varphi(\omega)$ 是样品与参考信号的相差, ω 是太赫兹波振动角频率.

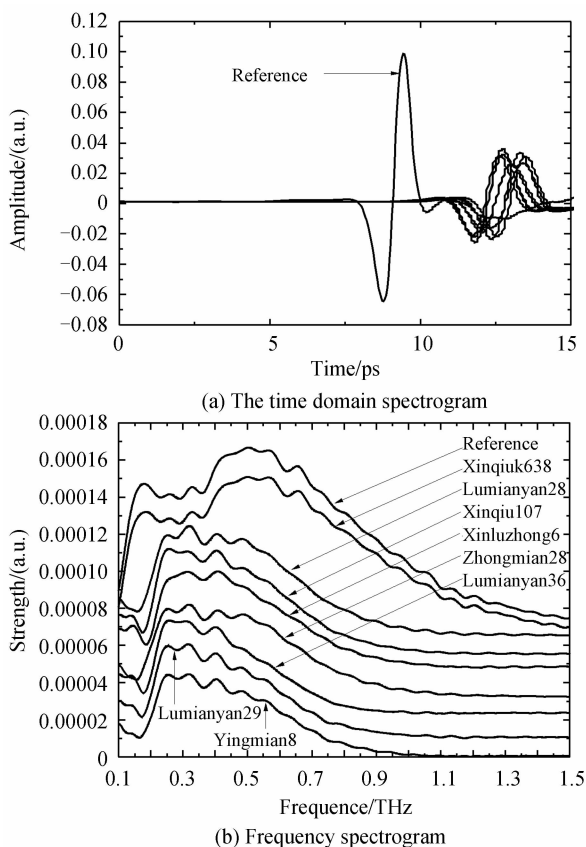


图 2 八种转基因棉花种子的时域光谱和频域光谱
(为清楚起见,光谱曲线向上做了平移)

Fig. 2 The time domain spectrogram and frequency spectrogram of eight kinds of transgenic cotton seeds (the spectra have been vertically shifted for clarity)

材料的吸收系数为

$$\alpha = \frac{2\omega\kappa}{c} \quad (2)$$

κ 是消光系数,表示材料的吸收, c 是光速.

式(1)可近似为

$$\frac{E_{\text{trans}}(\omega)}{E_0(\omega)} = A \exp(-i\varphi(\omega)) \approx \frac{4n}{(n+1)^2} \cdot$$

$$\exp\left[\frac{i\omega(n-1)d}{c}\right] \exp\left(-\frac{\alpha d}{2}\right) \quad (3)$$

材料的折射率和吸收系数为

$$n = \frac{c\varphi(\omega)}{\omega d} + 1 \quad (4)$$

$$\alpha = \frac{2}{d} \ln\left[\frac{4n}{A(n+1)^2}\right] \quad (5)$$

本文中,样品的纯度能达到 98% 以上,但并不是纯净物,且厚度有一定差异,对于同型转基因棉花种子,得不到准确的折射率和吸收系数.为了减小厚度对光谱分析的影响,采用吸光度(Absorbance)处理实验数据,吸光度是相对量,无量纲,表示光波被材料吸收的程度.吸光度公式

$$\text{Absorbance} = -\lg\left[\frac{E_{\text{trans}}(\omega)^2}{E_0(\omega)^2}\right] \quad (6)$$

由于人、光谱仪器自身存在的问题,相对中间的点,频域光谱的两端比较粗糙,两端的信噪比较低,因此将两端的信息删除,只采用中间信噪比较高的信息进行光谱分析.本文使用的 Z-3 系统,有效光谱探测范围为 0.1~3.5 THz,取 0.3~1.5 THz 频段进行分析.

如图 3,八条测试曲线依次为 Lumianyan28(鲁棉 28)、Lumianyan29(鲁棉 29)、Luminayan36(鲁棉 36)、Xinqiu107(鑫秋 107)、Xinqiuk638(鑫秋-k638)、Xingluzhong6(新陆中 6)、Yingmian8(银棉 8)、Zhongmian28(中棉 28)的吸收曲线,由于棉花的转基因型号不同,得到的吸光度曲线是不同的,曲线在更高频段出现剧烈振荡,这是由于样品对 THz 波过饱和和吸收所致;由图 3 八种样品的吸光度光谱可知,在 0.1~1.0 THz 频段吸光度都随着频率的增加近似线性增加,可做线性回归处理^[18],便于进一步研究这一频段转基因棉花对 THz 波的吸收特性.

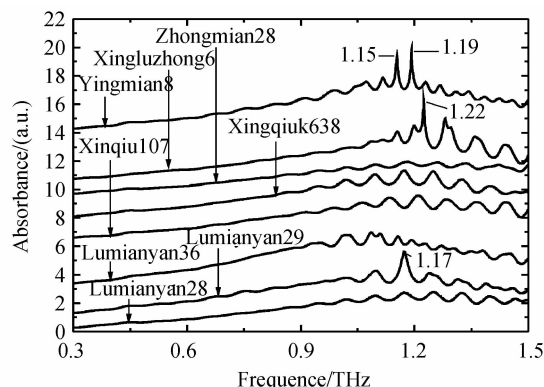


图 3 八种转基因棉花种子频率-吸光度图谱
(为清楚起见,光谱曲线向上做了平移)

Fig. 3 Frequency-absorbance spectrogram of eight kinds of transgenic cotton seeds (the spectra have been vertically shifted for clarity)

由图 3 可知,鲁棉 28、鲁棉 36、鑫秋 107、鑫秋-

k638 和中棉 28 无明显的吸收峰,鲁棉 29 的特征吸收峰为 1.17 THz,新陆中 6 的特征吸收峰为 1.22 THz,银棉 8 有两个明显的吸收峰,分别为 1.15 THz、1.19 THz.8 种样品的特征吸收峰峰位非常接近,不易观察区分,为了更精细地区分,建立数学模型,运用主成分分析^[19-20]法进行区分.

鲁棉 29、新陆中 6 和银棉 8 的吸收峰很明显,其吸收峰峰值和相关峰峰值相差比较大,但特征吸收峰峰位非常靠近,不易区分.其他 5 种样品无明显吸收峰,不易区分,这些波动较大的相关峰,主要的原因可能是 THz 光束入射到样品会发生多次反射,并与入射光束发生干涉,对判定结果产生较大影响;或是样品制备过程中研磨准确度偏低,样片中有一部分颗粒的粒度偏

大,产生较大的散射;或是由于有机大分子转动能级过高,在设定温度范围内无法迁移;或是样品的特征吸收峰不在仪器探测范围之内.

1.3 主成分分析

主成分分析法是一种典型的数据处理方法,能对数据集降维,是基于原始特征的一种线性变换,并最终得到能有效反映原始数据集特征的新数据.主成分分析法适合于数据之间没有相关性或相关性较小的数据.一般当少数新变量的累积方差贡献率约大于 65.000%时,这部分新变量可以近似作为原始数据集的主成分.

表 1~表 4 中,LMY28、LMY29、LMY36、XQ107、XQk638、YM8、ZM28、XLZ6 分别代表鲁棉 28、鲁棉 29、

表 1 相关系数矩阵

Table 1 The correlation coefficient matrix

	LMY28	LMY29	LMY36	XQ107	XQk638	YM8	ZM28	XLZ6
LMY28	1.000	0.322	0.559	0.263	0.283	0.331	0.237	0.237
LMY29	0.322	1.000	0.511	0.153	0.235	0.343	0.192	0.313
LMY36	0.559	0.511	1.000	0.168	0.234	0.337	0.106	0.388
XQ107	0.263	0.153	0.168	1.000	0.654	0.458	0.598	0.376
XQk638	0.283	0.235	0.234	0.654	1.000	0.530	0.629	0.334
YM8	0.331	0.343	0.337	0.458	0.530	1.000	0.367	0.423
ZM28	0.237	0.192	0.106	0.598	0.629	0.367	1.000	0.257
XLZ6	0.293	0.313	0.388	0.376	0.334	0.423	0.257	1.000

表 2 公因子方差

Table 2 Communalities

	Initial	Extract
LMY28	1.000	0.796
LMY29	1.000	0.558
LMY36	1.000	0.780
XQ107	1.000	0.740
XQk638	1.000	0.763
YM8	1.000	0.602
ZM28	1.000	0.733
XLZ6	1.000	0.752

表 4 成分矩阵

Table 4 Component matrix

Composition	1	2	3
LMY28	0.599	0.416	0.513
LMY29	0.545	0.496	-0.122
LMY36	0.591	0.644	0.128
XQ107	0.725	-0.426	0.028
XQk638	0.770	-0.404	0.083
YM8	0.734	-0.034	-0.250
ZM28	0.663	-0.497	0.216
XLZ6	0.634	0.158	-0.571

表 3 各成分解释的总方差

Table 3 Total variance explained

Component	Initial eigenvalues			Extraction of sum of squared loadings		
	Total	Variance /%	Cumulative /%	Total	Variance /%	Cumulative /%
1	3.503	43.783	43.783	3.503	43.783	43.783
2	1.483	18.543	62.326	1.483	18.543	62.326
3	0.737	9.211	71.537	0.737	9.221	71.537
4	0.672	8.399	79.936			
5	0.552	6.904	86.840			
6	0.395	4.941	91.781			
7	0.354	4.430	96.211			
8	0.303	3.789	100.000			

Extraction method: principal component analysis

鲁棉 36、鑫秋 107、鑫秋-k638、银棉 8、中棉 28、新陆中 6 八种不同转基因棉花种子.对吸光度频谱的原始数据进行标准化,得到进行主成分分析的变量的相关系数矩阵,如表 1.

表 2、表 3 和表 4 分别给出了原始数据的公因子方差、各个主成分的总方差和主成分的计算结果.在主成分综合分析过程中,新变量所代表的方差贡献率和原变量变为新变量的线性变换系数是研究的重点.前三个新变量的累计方差贡献率为 71.537%,大于 65.000%,因此只需分析第一主成分、第二主成分和第三主成分.

从表 4 成分矩阵可以看出,与第一主成分密切相关的是 XQ107、XQk638 和 YM8,它们与第一主成分的相关性绝对值都超过了 65.000%,这说明它们的第一

主成分含量较高,在 0.3~1.5 THz 频段对 THz 波有明显吸收,影响它们对此波段 THz 波吸收的是第一主成分;与第二主成分密切相关的是 LMY36,它与第二主成分的相关性绝对值超过了 50.000%,这说明它的第二主成分含量较高,在 0.3~1.5 THz 频段对 THz 波有明显吸收,影响它们对此波段 THz 波吸收的是第二主成分;与第三主成分密切相关的是 LMY28、XLZ6,它们与第三主成分的相关性绝对值都超过了 50.000%,这说明它们的第三主成分含量较高,在 0.3~1.5 THz 频段对 THz 波有明显吸收,影响它们对此波段 THz 波吸收的是第三主成分。

从表 3 各成分解释的总方差可以看出各个成分对总方差的贡献率,第一主成分的方差贡献率为 43.783%,远大于第二、第三主成分的 18.543%、9.211%,因此可以得出对 0.3~1.5 THz 频段的 THz 波有强烈吸收的是第一主成分的 XQ107、XQk638 和 YM8,其次是 LMY36。

如图 4,1~3 号样品是 LMY28,4~6 号样品是 LMY29,7~9 号样品是 LMY36,10~12 号样品是 XQ107,13~15 号样品是 XQk638,16~18 号样品是 XLZ6,19~21 号样品是 YM8,22~24 号样品是 ZM28。图 4(a)、(b)分别为二维得分图和三维得分图,Component 1 代表第一主成分,Component 2 代表第二主成分,Component 3 代表第三主成分。八种转基因棉

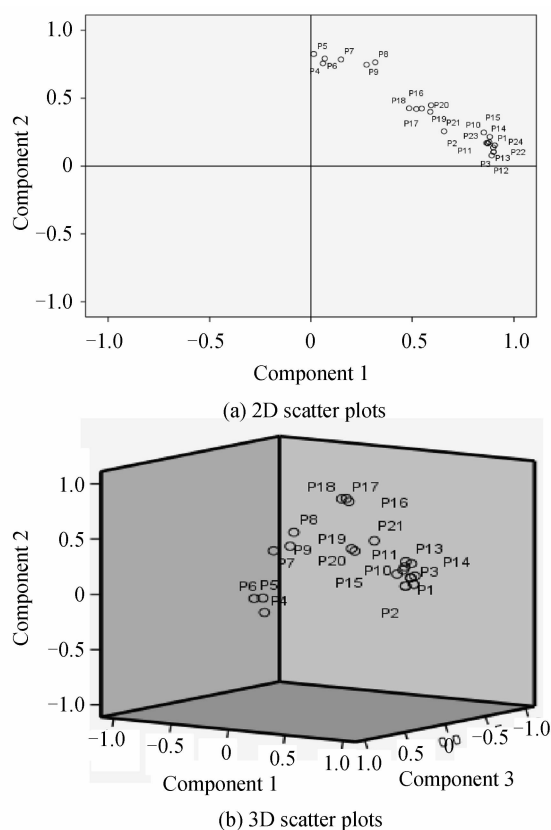


图 4 24 个样品主成分得分图
Fig. 4 PCA scores for 24 samples

花种子各有 27 个样片,对每 9 个样片的原始光谱数据取平均并标准化,分别得到 3 组数据,最后对 24 组数据做主成分分析,得到图 4 所示的得分图。两种或三种主成分代表了八种样品的 THz 光谱特性,24 组数据能被有效区分。由此可见,主成分分析法能有效提取转基因棉花种子样品的 THz 光谱特性并将相似的特性聚集。

Liu^[21]等运用太赫兹光谱技术检测了不同转基因棉花种子。结果表明,太赫兹光谱技术结合 GA-SVM 方法可以有效识别不同的转基因棉花种子。本文运用太赫兹光谱技术结合 PCA 方法亦得到了很好的鉴别和归类不同转基因棉花种子的效果,前三种主成分的累计方差贡献率达到了 71.537%,从主成分得分图中可以明显区分不同转基因棉花种子。由此可知,太赫兹光谱技术可以无损、快速、可靠地区分不同的转基因棉花种子。

2 结论

本文采用太赫兹时域光谱系统研究了 8 种转基因棉花种子吸收光谱,对比其在 0.3~1.5 THz 频段太赫兹波的吸收峰,并通过主成分分析,提取了三种主成分因子。结果表明,由于外源基因片段不同,不同转基因棉花种子具有不同的吸收峰;由于不同转基因棉花种子的三种主成分含量比重不同,不同转基因棉花种子在 THz 频段对 0.3~1.5 THz 波的吸收程度不同。由此可对不同的转基因棉花种子进行无损检测,为检测其他转基因品种奠定实验基础。

参考文献

- [1] LIU Sheng-gang. Recent development of terahertz science and technology[J]. *China Basic Science*, 2006, **8**(1): 7-12.
刘盛刚. 太赫兹科学技术的新发展[J]. *中国基础科学*, 2006, **8**(1): 7-12.
- [2] 许景周, 张希成. 太赫兹科学技术和应用[M]. 北京: 北京大学出版社, 2007.
- [3] GEN T, YOICHI K, KOICHIRO A, *et al.* Evaluation of drug crystallinity in aqueous suspension using terahertz time-domain attenuated total reflection spectroscopy[J]. *Pharmaceutical Sciences*, 2013, **102**(11): 4065-4071.
- [4] GE H Y, JING Y Y, XU Z H, *et al.* Identification of wheat quality using THz spectrum[J]. *Optics Express*, 2014, **22**(10): 12533-12544.
- [5] WANG Q, MA Y H. Qualitative and quantitative identification of nitrofen in terahertz region[J]. *Chemometrics and Intelligent Laboratory Systems*, 2013, **127**(8): 43-48.
- [6] 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. 基于模糊识别的爆炸物 THz 光谱识别[J]. *仪器仪表学报*, 2012, **33**(11): 2480-2486.
- [7] DOY C H, GRESSHOFF P M, ROLFE B G. Biological and molecular evidence for the transgenesis of genes from bacteria to plant cells[J]. *Proceedings of the National Academy of*

- Sciences, 1973, **70**(3): 723-726.
- [8] JOHNSON E T, CHRISTOPHER S, DOWD P F. Identification of a bioactive bowman-birk inhibitor from an insect-resistant early maize inbred[J]. *Journal of Agricultural and Food Chemistry*, 2014, **62**(24): 5458-5465.
- [9] KINUE K, MINORI N, KAORI N, *et al.* Identification of a feather beta-keratin gene exclusively expressed in pennaceous barbule cells of contour feathers in chicken[J]. *Gene*, 2014, **542**(1): 23-28.
- [10] ANDRA M, SCOTT W, JAY H, *et al.* THz time domain spectroscopy of biomolecular conformational modes [J]. *Physics in Medicine and Biology*, 2002, **47**(21): 3797-3805.
- [11] WHITMIRE S E, WOLPERT D, MARKELZ A G, *et al.* Protein flexibility and conformational state: A comparison of collective vibrational modes of wild-type and D96N bacteriorhodopsin[J]. *Biophysical Journal*, 2003, **85**(2): 1269-1277.
- [12] NWOKOCHA L M, WILLIAMS P A. Isolation and characterization of a novel polysaccharide from seeds of *Peltophorum pterocarpum*[J]. *Food Hydrocolloids*, 2014, **41**(12): 319-324.
- [13] ZHANG W W, XIAO H N, QING L Y. Enhanced water vapour barrier and grease resistance of paper bilayer-coated with chitosan and beeswax [J]. *Carbohydrate Polymers*, 2014, **101**(9): 401-406.
- [14] HOSOBUCHI M, XIE X, CHENG Y, *et al.* 2013 IEEE conference on electrical insulation and dielectric phenomena: measurements of THz absorption peaks in photo-degraded polyethylene and their assignment by quantum chemical calculations[C]. Shenzhen:IEEE, 2013: 20-23.
- [15] WU Q, LITZ M, ZHANG X C. Broadband detection capability of ZnTe electro-optic field detectors[J]. *Applied Physics Letters*, 1996, **68**(21): 2924-2926.
- [16] PLATTE F, HEISE H M. Substance identification based on transmission THz spectra using library search[J]. *Journal of Molecular Structure*, 2014, **1073**(SD): 3-9.
- [17] LIU X F, LI Q Z, LI R, *et al.* The structure, properties, and nature of HARF - benzene complex; Redshift and blueshift of Ar-H stretch frequency and rare gas atomic number dependence of hydrogen bonds[J]. *Spectrochimica Acta Part A*, 2011, **84**(8): 68-73.
- [18] QI Hui, LI Yu-wu, LIU Xian-de, *et al.* Identification and estimate of polychlorinated biphenyls mixture by multiple linear regression analysis[J]. *Research of Environmental Sciences*, 2005, **18**(4): 91-95.
祁辉, 李玉武, 刘咸德, 等. 用多元线性回归分析法定量判别 PCBs 污染物类型[J]. 环境科学研究, 2005, **18**(4): 91-95.
- [19] CAI Xiao-na, HUANG Da-zhuang, SHEN Zu-rui, *et al.* Research on artificial neural network method used for insects classification and identification principal component analysis and mathematical modeling[J]. *Journal of Biomathematics*, 2013, **28**(1): 23-33.
蔡小娜, 黄大庄, 沈佐锐, 等. 用于昆虫分类鉴定的人工神经网络方法研究: 主成分分析与数学建模[J]. 生物数学学报, 2013, **28**(1): 23-33.
- [20] CAO Fang, WU Di, HE Yong, *et al.* Variety discrimination of grapes based on visible-near reflection infrared spectroscopy[J]. *Acta Optica Sinica*, 2009, **29**(2): 537-540.
曹芳, 吴迪, 何勇, 等. 基于可见-近红外反射光谱技术的葡萄品种鉴别方法的研究[J]. 光学学报, 2009, **29**(2): 537-540.
- [21] LIU J J, LI Z, HU F R, CHEN Tao, ZHU A J. A THz spectroscopy nondestructive identification method for transgenic cotton seed based on GA-SVM[J]. *Optical and Quantum Electronics*, 2014, **46**(5): 9914.