

# 基于光谱特征约束非负矩阵分解的轮胎橡胶太赫兹混叠光谱分离

殷贤华<sup>1,2</sup>, 刘昱<sup>1,2</sup>, 奉慕霖<sup>1,2</sup>, 李安<sup>3</sup>, 莫玮<sup>1\*</sup>

1. 桂林电子科技大学电子工程与自动化学院, 广西 桂林 541004
2. 广西自动检测技术与仪器重点实验室, 广西 桂林 541004
3. 国家橡胶及橡胶制品质量监督中心(广西), 广西 桂林 541004

**摘要** 随着“绿色轮胎”概念的深入, 轮胎橡胶的有效成分直接关系到橡胶合格与否, 而轮胎橡胶对于检测部门而言是一个“黑色”分析系统, 如何利用现有手段进行橡胶成分的准确检测至关重要。太赫兹时域光谱(THz-TDS)技术已经成功应用于物质检测分析当中, 但是在许多情况下, 从橡胶这个复杂样品中观察到的太赫兹光谱数据代表了几个相互关联的组分或特征成分共同作用的综合结果, 原始数据中包含的实际信息可能会重叠, 进而影响橡胶混合物中各组分的分析。为了解决太赫兹光谱重叠的问题, 利用太赫兹光谱矩阵连续平滑和浓度矩阵稀疏的特点, 将具有平滑特性的 2 范数和具有稀疏特性的 1/2 范数引入到非负矩阵分解方法, 提出基于光谱特征约束非负矩阵分解(constraint non-negative matrix factorization, CNMF)的太赫兹混叠光谱分离方法。首先通过 THz-TDS 获取丁腈橡胶(NBR)与硫化促进剂 2-巯基苯并噻唑(MBT)二组分混合物, 以及 NBR、MBT 和一硫化四甲基秋兰姆(TMTM)三组分混合物的时域光谱; 然后对时域光谱进行傅里叶变换得到其频域光谱, 进而获取混合物所对应的吸光度混叠光谱矩阵; 光谱分离之前对获得的光谱矩阵进行主成分分析, 初步判定混合物组分数, 最后利用非负矩阵分解算法(non-negative matrix factorization, NMF)、基于纯变量初始化的非负矩阵分解(purity non-negative matrix factorization, PNMF)和 CNMF 三种方法对混叠光谱进行光谱解析。对比发现, 相较于 NMF 和 PNMF 方法, CNMF 算法的分离效果更佳, 特征吸收峰对应结果更准确, 针对不同组分混合物分离结果的相关系数均高于 89%, 光谱角小于 0.5, 具有较高的纯净物光谱还原度。将带有约束的非负矩阵分解算法引入到太赫兹混叠光谱的分离上, 较好的提取出复杂混合物中单一成分的特征信息, 为后续的太赫兹多组分混合物的定性分析和定量计算奠定了较好的基础, 在绿色轮胎及橡胶的质量检测领域具有一定的研究前景。

**关键词** 太赫兹时域光谱; 光谱分离; 非负矩阵分解; 轮胎橡胶; 黑色分析体系

**中图分类号:** O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)12-3736-07

## 引言

太赫兹波是一种处于红外和微波之间的频率间隙中新兴的振动光谱分析技术, 通常覆盖 0.1 至 10 THz<sup>[1]</sup>。在过去的十年里, 该技术已经成为在医学诊断<sup>[2]</sup>、监测结晶度和多态性<sup>[3]</sup>以及多组分混合物的定量检测方面的研究工具。然而, 其研究对象的多为纯净物或者已知组分的混合物白色系统, 面对混合物成分完全未知的黑色系统, 如何从未知混合物测得的数据中提取出纯组分的光谱数据就成为进一步探索物理

或化学机制的首要任务。针对此问题, Ma<sup>[4]</sup>等将自模型曲线分辨法引入太赫兹混叠光谱分解, 但是该方法需要知道光谱的一些先验知识进行交互; Gan<sup>[5]</sup>等将基于纯变量初始化的非负矩阵分解方法应用于太赫兹混叠光谱分解, 相较于随机初始化的非负矩阵分解方法, 该方法在精度和速度上都有所提高, 但是该分解方法缺乏适当的约束, 对于三组分混合物的分解结果稍微较差。

轮胎橡胶是由胶料(生胶)和一些加工助剂(硫化剂、防老剂等)组成的多组分混合物。对于轮胎制造企业, 生产轮胎的橡胶及助剂混合物组分都是已知的, 属于白色分析体

收稿日期: 2019-11-07, 修订日期: 2020-03-21

基金项目: 国家自然科学基金项目(61841502), 广西自然科学基金项目(2018GXNSFAA281341), 广西自动检测技术与仪器重点实验室基金项目(YQ19102), 桂林电子科技大学研究生创新项目(YJXCS201560)资助

作者简介: 殷贤华, 1974 年生, 桂林电子科技大学电子工程与自动化学院副教授 e-mail: yxh4417@guet.edu.cn

\* 通讯联系人 e-mail: wmo@miit.gov.cn

系, 对其检测可以防止轮胎生产过程中用错原材料成分而造成轮胎质量问题; 对于轮胎橡胶和助剂检测部门而言, 轮胎橡胶及助剂混合物组别则是全部未知或部分已知, 属于黑色分析体系或灰色分析体系, 对其检测可以知道有效成分是否达标, 从而实现轮胎产品的质量监督检查。丁腈橡胶(nitrile-butadiene rubber, NBR)具有耐油性好、耐磨性较高的特点, 在汽车、航空和石油等行业中得到广泛应用。2-硫醇基苯并噻唑(MBT)和一硫化四甲基秋兰姆(TMTM)是轮胎橡胶工业中常见的两种橡胶硫化促进剂。MBT 是橡胶工业中广泛使用的硫化促进剂, 常常作为其他硫化促进剂的母体原料<sup>[6]</sup>; TMTM使硫化橡胶拉伸强度偏低, 适用于工程车辆。因此对于丁腈橡胶、MBT 和 TMTM 组成的混合物体系进行光谱解析, 对于轮胎产品质量的检测具有重要意义。

针对轮胎橡胶黑色分析体系以及现有方法难以准确对太赫兹混叠光谱进行分离的问题, 本工作以丁腈橡胶、MBT 和 TMTM 组成的多组分混合物为研究对象, 结合太赫兹光谱曲线特征, 将带有连续约束和稀疏约束的非负矩阵分解方法引入到橡胶混合物的太赫兹光谱分离问题上, 针对橡胶混合物的混叠光谱问题展开研究。

## 1 实验部分

### 1.1 装置

装置采用美国 Zomega 公司的 Z-3 透射式太赫兹时域光谱系统和德国的 TOPTICA Photonics AG 公司的飞秒光纤激光器组成<sup>[7]</sup>。飞秒激光器进光功率约为 120 mW, 脉冲持续时间为 100 fs, 中心波长为 780 nm, 重复频率为 80 MHz。本系统主要由 THz 波辐射源、THz 波探测器和时间延迟控制装置组成, 系统示意图如图 1 所示<sup>[8]</sup>。

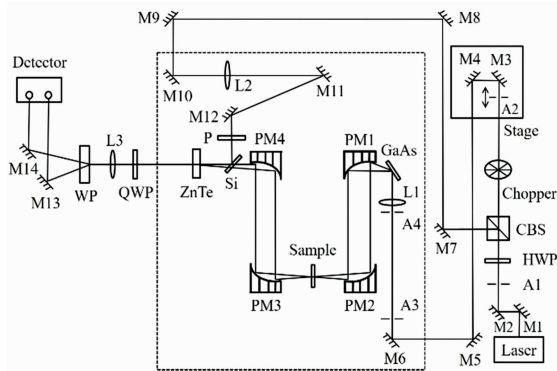


图 1 透射式 THz-TDS 原理图

Laser: 激光源; M1—M14: 反射镜; A1—A4: 光阑; HWP: 半波片; CBS: 分束镜; Chopper: 斩波器; L1—L3: 凸透镜; PM1—PM4: 离轴抛物面镜; QWP: 1/4 波片; WP: 渥拉斯顿棱镜; Detector: 探测器

Fig. 1 Schematic diagram of THz-TDS system

Laser: Laser source; M1—M14: Reflector; A1—A4: Diaphragm; HWP: Half wave plate; CBS: Cube beamsplitters; Chopper: Optical chopper; L1—L3: Convex lens; PM1—PM4: Off-axis parabolic mirror; QWP: Quarter wave plate; WP: Wollaston prism; Detector: Pulse detector

### 1.2 样品

实验对象为以丁腈橡胶为主要成分, 结合硫化促进剂 MBT 和 TMTM 组成的二组分和三组分混合物。丁腈橡胶购于天津明基金泰橡塑制品加工有限公司, 其纯度为 93.75%, 硫化促进剂 MBT 和 TMTM 均购于东莞市百亿塑料制品有限公司, 其纯度在 96% 以上, 所有实验样品均为固体粉末。样品制备方法采用压片法, 每个样本为厚度约 1 mm, 半径 6.5 mm, 重量约为 180 mg 的圆形样片。样本压制前后, 分别将原材料和压制好的样片放入真空恒温干燥箱中于 50 °C 干燥 2 h。每种比例的样片压制 6 个, 采用等梯度法进行浓度配比, 其各组分含量的具体比例见表 1 和表 2。

表 1 二组分混合物浓度配比信息

Table 1 Concentration ratio information of binary system

编号	NBR/%	MBT/%	样品数量
1	10	90	6
2	20	80	6
3	30	70	6
4	40	60	6
5	50	50	6
6	60	40	6
7	70	30	6
8	80	20	6
9	90	10	6

表 2 三组分混合物浓度配比信息

Table 2 Concentration ratio information of ternary system

编号	NBR/%	TMTM/%	MBT/%	样品数量
1	50	5	45	6
2	50	10	40	6
3	50	15	35	6
4	50	20	30	6
5	50	25	25	6
6	50	30	20	6
7	50	35	15	6
8	50	40	10	6
9	50	45	5	6

### 1.3 光学参数提取

THz-TDS 检测的主要目标是获得表征材料特征的宏观光学常数, 一般包括复折射率、介质损耗、电导率、吸收系数等<sup>[9]</sup>。利用透射式太赫兹时域光谱系统在获得参考信号  $E_{\text{ref}}(t)$  和样品信号  $E_{\text{sam}}(t)$  后, 经过快速傅里叶变换, 得到参考信号和样本信号的频域信号  $E_{\text{ref}}(\omega)$  和  $E_{\text{sam}}(\omega)$ 。吸光度表示太赫兹波被材料吸收的程度, 其为无量纲的相对量。使用吸光度来描述样品的吸收特性, 可以减少样本厚度不均匀对光谱分析产生的影响。吸光度的计算公式如式(1)所示<sup>[10]</sup>

$$A(\omega) = -\lg \left| \frac{E_{\text{sam}}(\omega)}{E_{\text{ref}}(\omega)} \right|^2 \quad (1)$$

其中  $\omega$  为太赫兹波振动的角频率。

### 1.4 模型评价方法

采用相关系数(correlation coefficient,  $r$ )和光谱角(spect-

tral angle, SA)作为模型的评价指标,  $r$  表征变量之间线性相关的程度, SA 从向量夹角方面比较了光谱在形状上的相似性。相关系数越大, 光谱角越小, 说明模型的性能越好。如式(2)和式(3)所示

$$r = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

$$SA(x, y) = \cos^{-1} \left( \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \right) \quad (3)$$

其中  $n$  为样本数,  $x_i$  和  $y_i$  分别表示不同组分第  $i$  个样本的参考值,  $\hat{y}_i$  为第  $i$  个样本的预测值,  $\bar{y}$  为  $n$  个样本参考值的均值。

## 2 结果与讨论

### 2.1 光谱分析

通过使用太赫兹透射系统分别对 NBR, MBT 和 TMTM 三种纯净物 and 不同比例组成的混合物样本进行检测, 由于实验仪器两侧的信噪比较低, 选取包含吸收峰的 0.3 ~ 1.5 THz 波段进行分析。各样品的太赫兹吸光度谱如图 2 和图 3 所示。

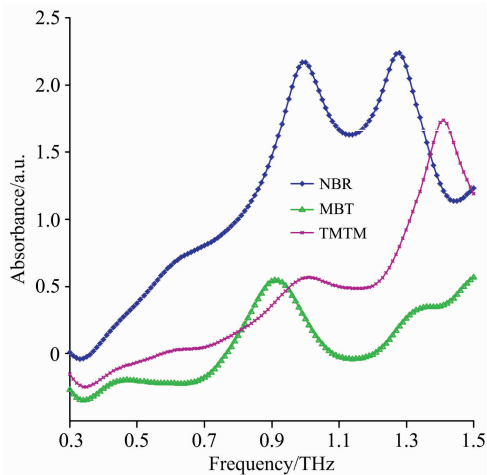


图 2 NBR, MBT 和 TMTM 纯净物吸光度

Fig. 2 THz absorbance spectrum of NBR, MBT and TMTM

由图 2 可知, 丁腈橡胶在 0.3 ~ 1.5 THz 频段内存在两个明显的特征吸收峰, 分别位于 0.99 和 1.29 THz, 防老剂 MBT 在 0.91 THz 处有较为明显的特征吸收峰, TMTM 在 1.01 和 1.41 THz 有两个特征吸收峰。随着频率的增加, 吸光度曲线呈现上升的趋势, 其中 NBR 峰最强, TMTM 次之, MBT 峰最弱。

图 3(a) 对应丁腈橡胶和 MBT 按照不同比例混合而成的二组分混合物吸光度图谱。从图中可以看出随着 MBT 含量的降低, 0.93 THz 处的吸收峰在丁腈橡胶的作用下吸收峰逐渐向高频移动, 在 1.1 ~ 1.5 THz 区间内, 不同浓度配比

的吸收曲线交织在一起, 看不到明显的规律; 图 3(b) 为对应丁腈橡胶、TMTM 和 MBT 按照不同比例混合而成的三组分混合物吸光度图谱。从图中可以看出, 三组分混合物主要在 0.93 和 1.37 THz 附近存在两个比较突出的吸收峰, 相比于二组分混合物的吸收光谱, 三组分混合物的吸收曲线重叠效应更加明显, 没有明显的梯度关系, 因此有效的混叠光谱分离算法意义重大。

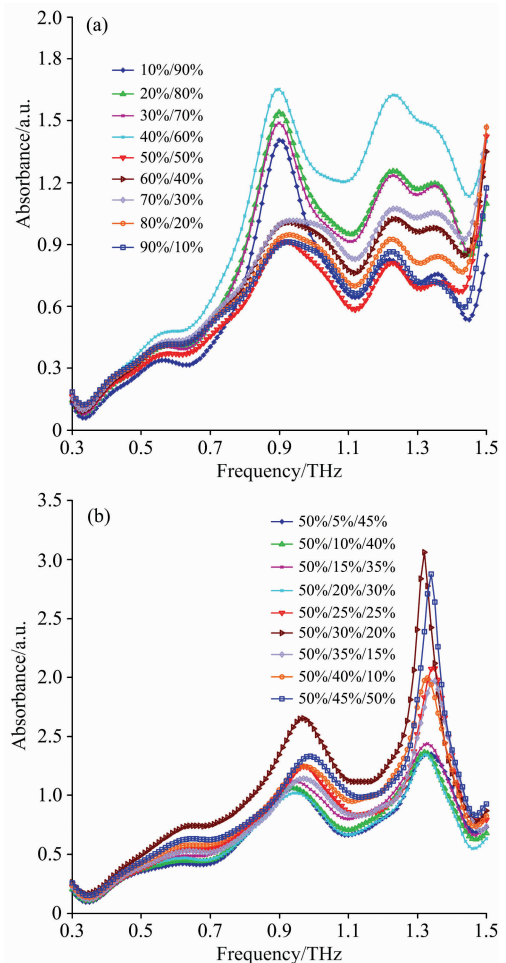


图 3 九种不同浓度配比混合物吸光度谱

(a): 二组分混合物; (b): 三组分混合物

Fig. 3 THz absorbance spectrum of nine different concentration mixtures

(a): Binary system; (b): Ternary system

### 2.2 混叠光谱剥离

常规的混合物检测分析中, 都是在混合物组分已知的前提下, 针对光谱曲线与浓度之间的关系进行建模; 然而, 当分析对象为复杂的混合物时, 由于光谱重叠或失真的影响, 或者当混合物组分完全未知时, 面对测得的光谱结果则无法进行有效的光谱解析。因此, 有效的混叠光谱分离算法至关重要。

非负矩阵分解<sup>[11]</sup> (non-negative matrix factorization, NMF), 自 1999 年由 Lee 和 Seung 提出以来, 被广泛应用在图像处理 and 模式识别。本研究结合太赫兹光谱平滑性和矩阵

稀疏性的特点, 在原始非负矩阵分解的基础上加入特定约束项, 利用欧氏距离构造逼近误差的目标函数, 实现太赫兹混叠光谱的有效分离, 其数学过程如式(4)

$$\min E = \frac{1}{2} \| \mathbf{V} - \mathbf{WH} \|_2^2 + \frac{1}{2} \alpha \| \mathbf{W} \|_2^2 + 2\beta \| \mathbf{H} \|_{1/2}$$

$$\mathbf{W}, \mathbf{H}, \alpha, \beta \geq 0 \quad (4)$$

其中,  $\mathbf{E}$  为矩阵分解前后的逼近误差,  $\mathbf{W} \in R^{m \times n}$  表示  $m$  行  $n$  列的原始吸光度数据矩阵,  $\mathbf{W} \in R^{m \times k}$  表示混合物中纯组分的吸光度数据矩阵,  $\mathbf{H} \in R^{k \times n}$  表示  $\mathbf{W}$  所对应的浓度矩阵,  $k$  表示待分类组分数。  $\alpha$  和  $\beta$  表示正则化参数, 将逼近误差作为适应度函数, 利用粒子群算法(PSO)进行参数寻优,  $\| \cdot \|_2$  表示 2 范数, 用于光谱矩阵的连续性约束,  $\| \cdot \|_{1/2}$  表示 1/2 范数, 用于浓度矩阵的稀疏性约束。

使用乘法迭代规则对光谱矩阵和浓度矩阵  $\mathbf{H}$  进行迭代更新, 如式(5)和式(6)

$$\mathbf{W} \leftarrow \mathbf{W} * \frac{\mathbf{VH}^T + \alpha \mathbf{W}}{\mathbf{WHH}^T + \alpha \mathbf{WW}^T} \quad (5)$$

$$\mathbf{H} \leftarrow \mathbf{H} * \frac{\mathbf{W}^T \mathbf{V}}{\mathbf{W}^T \mathbf{WH} + \beta \mathbf{H}^{-1/2}} \quad (6)$$

由于乘法迭代更新规则对初始矩阵的选择较敏感, 随机初始化难以达到要求, 所以本文利用 Gan 提出的基于纯变量初始化方法初始化光谱矩阵和浓度矩阵<sup>[5]</sup>, 在利用主成分分析(PCA)确定混合物组分数之后, 进行迭代更新, 实现混合光谱矩阵的解析。

在进行矩阵分解之前, 针对获得的混合物数据矩阵进行主成分分析, 通过主成分的方差贡献率来预测混合物组分数。为消除噪声的影响, 减小误差, 对表 2 和表 3 中每种浓度配比对应制作 6 个样片, 每个样品正反各测三次, 对每种

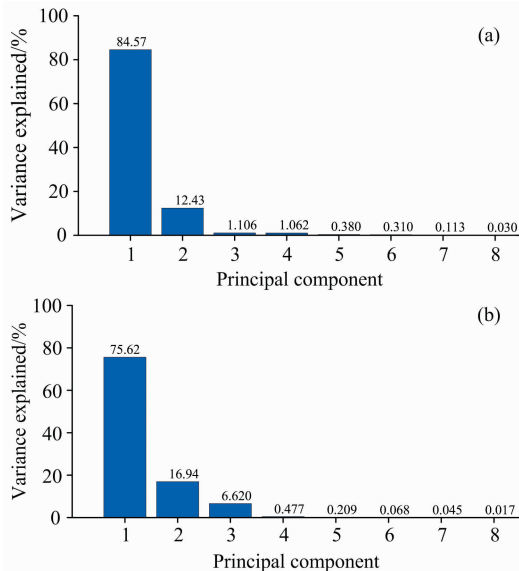


图 4 基于方差贡献率的纯组分估计  
(a): 二组分混合物; (b): 三组分混合物

Fig. 4 A plot of relative contribution versus factor number by PCA

(a): Binary system; (b): Ternary system

浓度配比获得 36 个光谱数据取平均后得到最终的吸光度光谱矩阵。针对 9 组不同浓度配比组成的数据矩阵进行主成分分析, 确定混合物组分数, 组分估计如图 4 所示。

从图 4(a)可以看出, 第一个主成分的方差贡献率为 84.57%, 第二个主成分的方差贡献率为 12.43%, 前两个主成分的方差贡献率累计达到 97.00%; 从图 4(b)可以看出, 前两个主成分的方差贡献率为 92.56%, 加入第三个主成分的方差贡献率累计达到 99.18%。依照经验, 主成分方差贡献率累计达到 95% 基本可以包含主要信息, 所以图 4(a)中纯组分数可初步认定为 2 种, 图 4(b)中可初步认定纯组分数的种类为 3 种。

确定混合物纯组分种类数之后, 利用基于光谱 CNMF, 与 NMF 和 PNMF 算法进行对比分析, 验证本算法的有效性。其中, 利用 CNMF 方法得到的混合物光谱解析如图 5 和图 6 所示。

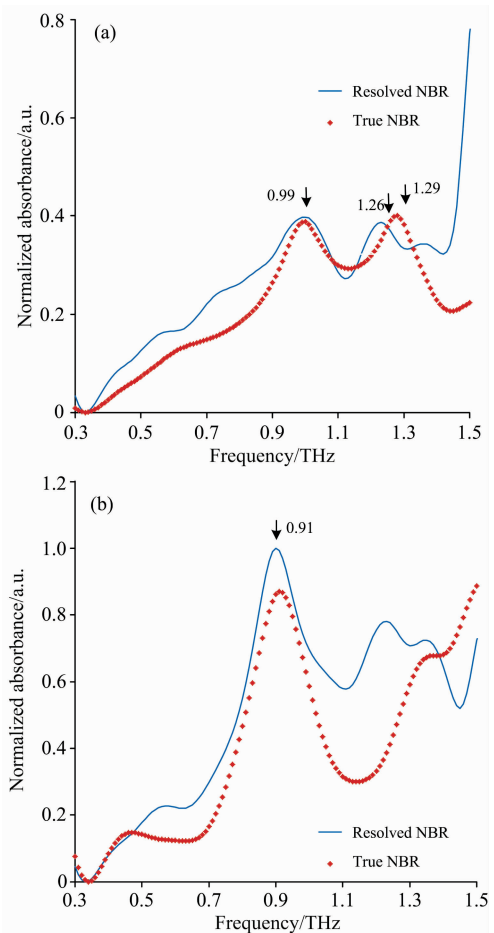


图 5 对二组分混合物使用 CNMF 的结果

(a): 丁腈橡胶; (b): 二巯基苯并噻唑

Fig. 5 Best resolved spectra of binary system extracted by CNMF

(a): NBR; (b): MBT

经过纯组分数预测和矩阵的初始化后, 经过多次迭代, 正则化参数和分别取 1 和 0.1 时迭代误差最小。针对二组分混合物进行光谱解析, 结果如图 5 所示, 丁腈橡胶和 MBT

的主要特征吸收峰都能较好的提取出来。从图 5(a)可以看出, 丁腈橡胶在 0.99 THz 处的特征吸收峰可以较好的分离出来, 但是由于两物质相互作用的影响, 1.29 THz 处的吸收峰存在左移的现象; 从图 5(b)可以看出, 相较于丁腈橡胶, MBT 特征吸收峰可以较好的分离出来, 且两者的趋势基本一致。但是在 1.1 ~ 1.5 THz 频段, 受丁腈橡胶的影响,

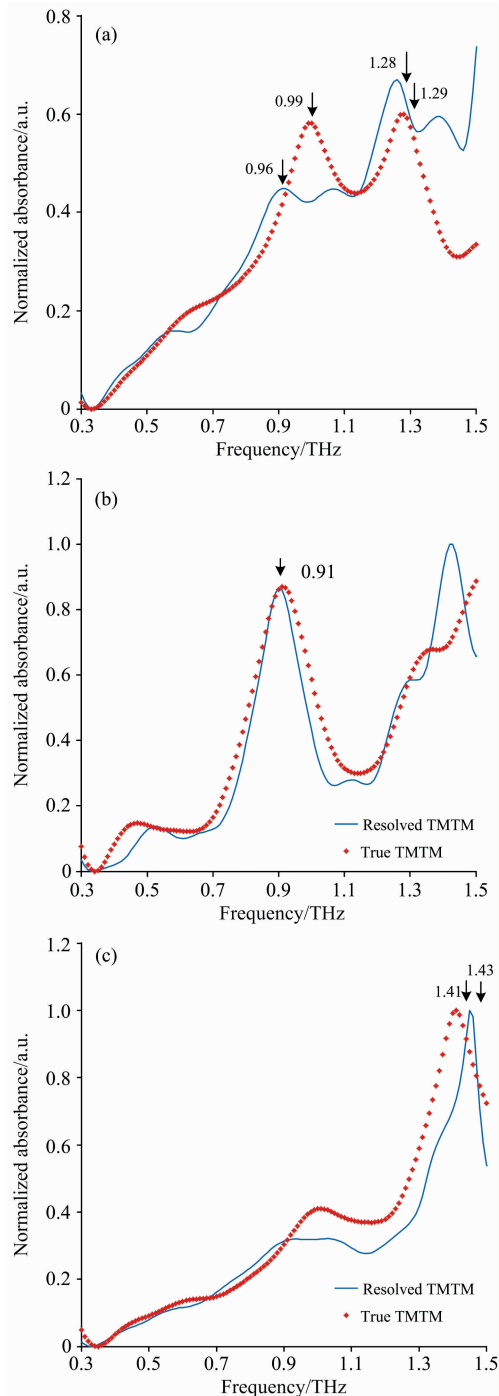


图 6 对三组分混合物使用 CNMF 的结果

(a): 丁腈橡胶; (b): 二巯基苯并噻唑; (c): 硫化四甲基秋兰姆

Fig. 6 Best resolved spectra of ternary system extracted by CNMF

(a): NBR; (b): MBT; (c): TMTM

MBT 产生了两个多余的较弱吸收峰。

经过纯组分预测和矩阵的初始化后, 经过多次迭代, 正则化参数和分别取 0.7 和 0.32 时迭代误差最小。针对三组分混合物进行光谱解析, 结果如图 6 所示, MBT 和 TMTM 的特征吸收峰均可以较好的分离出来, 丁腈橡胶的吸收峰分离结果相对存在一些偏差。从图 6(a)可以看出, 丁腈橡胶分解的结果和实测结果趋势基本一致, 但是受 MBT 吸收峰的影响, 两个主要的吸收峰都存在一定的低频移动; 从图 6(b)可以看出, MBT 的特征吸收峰可以较好的对应起来, 但是受 TMTM 的影响, 在 1.45 THz 处存在一个较明显的多余吸收峰; 从图 6(c)可以看出, TMTM 的分离结果和实测结果基本一致, 在 1.41 THz 处的吸收峰存在微弱的高频移动, 其主要原因还是三种物质混合物作用的结果, 导致分离结果存在一定的偏差。

采用相关系数和光谱角两个评价指标对实验结果做进一步分析, 结果如表 3 和表 4 所示。

表 3 二组分混合物光谱分离结果

Table 3 Spectral separation result of binary system			
组分	方法	相关系数 $r/\%$	光谱角 SA
NBR	NMF	79.65	0.573
	PNMF	90.73	0.392
	CNMF	<b>91.21</b>	<b>0.357</b>
MBT	NMF	83.32	0.473
	PNMF	95.78	0.213
	CNMF	<b>96.89</b>	<b>0.102</b>

表 4 三组分混合物光谱分离结果

Table 4 Spectral separation result of ternary system			
组分	方法	相关系数 $r/\%$	光谱角 SA
NBR	NMF	66.23	0.593
	PNMF	87.25	0.476
	CNMF	<b>89.33</b>	<b>0.433</b>
MBT	NMF	82.57	0.533
	PNMF	<b>92.49</b>	<b>0.419</b>
	CNMF	92.37	0.493
TMTM	NMF	80.32	0.558
	PNMF	95.27	0.201
	CNMF	<b>98.88</b>	<b>0.089</b>

从表 3 可以看出, 针对丁腈橡胶和 MBT 分解结果的相关系数和光谱角中, CNMF 算法的效果最佳, 结果分别为 91.21% 和 0.357, PNMF 算法次之, NMF 最差; 相较于丁腈橡胶的结果, MBT 的分离结果更佳, CNMF 算法的相关系数和光谱角分别为 96.89% 和 0.102。总体而言, CNMF 算法效果更好。

如表 4 所示, 丁腈橡胶分离结果的相关系数和光谱角中, CNMF 算法最佳, 分别为 89.33% 和 0.439, 但是由于三组分混合物的复杂性, 其结果相较于二组分混合物分离结果有所下降; MBT 分离结果的相关系数和光谱角中, PNMF

算法效果更佳, 分别为 92.49% 和 0.419, 然而 CNMF 算法分离结果与 PNMF 算法结果较接近, 效果相对较好; TMTM 分离结果的相关系数和光谱角中, CNMF 算法效果最佳, 分别为 98.88% 和 0.089。综合三种物质的分离结果, 虽然 CNMF 在 MBT 的分解上稍差于 PNMF, 但是总体而言, CNMF 算法效果更佳。

### 3 结 论

太赫兹时域光谱技术在物质检测的广泛应用, 使得未知

混合物吸收峰构成的准确解析变得尤其重要。针对轮胎橡胶的太赫兹混叠光谱问题, 提出基于光谱特征约束非负矩阵分解的太赫兹混叠光谱分离方法, 分别对丁腈橡胶和 MBT 组成的二组分混合物, 丁腈橡胶、MBT 和 TMTM 组成的三组分混合物进行光谱分离实验。实验结果表明, 相较于 NMF 和 PNMF 算法, CNMF 方法分离的准确性得到了提升, 效果最佳。本研究较好的遵循了混合物分析中“从整体到局部, 先定性后定量”的分析准则, 为太赫兹混叠光谱分离提供了更优的方法, 也为绿色轮胎的无损、快速检测提供了新的技术参考, 具有一定的应用与研究价值。

### References

- [1] Rolere S, Liengprayoon S, Vaysse L, et al. *Polymer Testing*, 2015, 43: 83.
- [2] Reid C B, Fitzgerald A, Reese G, et al. *Physics in Medicine and Biology*, 2011, 56(14): 4333.
- [3] Virira F S, Pasquini C. *Analytical Chemistry*, 2014, 86(8): 3780.
- [4] Ma Y, Huang P, Hou D, et al. *Chemometrics and Intelligent Laboratory Systems*, 2016, 150: 65.
- [5] Gan J, Qin B Y, Li Y, et al. *Optik*, 2019, 176: 600.
- [6] AN Jing, HAN Yi, LIU Bing-yan(安静, 韩毅, 刘秉言). *China Rubber(中国橡胶)*, 2015, (10): 46.
- [7] Chen T, Li Z, Yin X H, et al. *Spectrochim Acta A*, 2016, 153: 586.
- [8] CHEN Tao, CAI Zhi-hua, HU Fang-rong, et al(陈涛, 蔡治华, 胡放荣, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2019, 39(3): 687.
- [9] WANG Qi-chao, WANG Jia-chun, WANG Ke-wei, et al(王启超, 汪家春, 王科伟, 等). *Chinese Journal of Luminescence(发光学报)*, 2016, 37(6): 737.
- [10] Chen Z, Zhang Z, Zhu R, et al. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2015, 167: 1.
- [11] Lee D D, Seung H S. *Nature*, 1999, 401: 788.

## Separation of Tire Rubber Overlapping Terahertz Spectra Using Non-Negative Matrix Factorization of Spectral Feature Constraints

YIN Xian-hua<sup>1, 2</sup>, LIU Yu<sup>1, 2</sup>, FENG Mu-lin<sup>1, 2</sup>, LI An<sup>3</sup>, MO Wei<sup>1\*</sup>

1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China

2. Guangxi Key Laboratory of Automatic Detection Technology and Instrument, Guilin 541004, China

3. National Rubber and Rubber Products Quality Supervision and Inspection Center(Guangxi), Guilin 541004, China

**Abstract** With the deepening of the concept of “green tires”, the effective composition of tire rubber is directly related to the qualification of rubber. But tire rubber is a “black” analysis system for the inspection department, and it’s exceedingly crucial to accurately detect rubber components by the existing methods. Terahertz time-domain spectroscopy (THz-TDS) technology has been successfully applied to material detection and analysis, but the terahertz spectral data observed from a complex sample of rubber represents the comprehensive results of several interrelated components or interaction of characteristic components in many cases, where as the actual information contained in the raw data may overlap, which will conversely affect the analysis of the components in the rubber mixture. In order to solve the problem of terahertz spectral overlap, the characteristics of continuous smoothing of terahertz spectral matrix and sparse concentration matrix are combined in this paper, then the 2 norms with smoothing characteristics and the 1/2 norm with sparsity characteristics into the non-negative matrix factorization method is introduced, which are applied to the separation of terahertz aliased spectra, so as separation method of terahertz aliasing spectral based on spectral feature Constrained Non-negative Matrix Factorization (CNMF) is proposed. Firstly, nitrile-butadiene rubber combined with vulcanization accelerator 2-Mercaptobenzothiazole(MBT) to form a binary mixture in diverse proportions, and it

combined with vulcanization accelerators MBT and tetramethyl thiuram monosulfide (TMTM) to form a ternary mixture in different proportions. Then the terahertz time domain spectrum of all samples is measured by terahertz spectroscopy system, which the measured data is subjected to obtain a corresponding absorbance spectrum. Further, principal component analysis is performed on the obtained spectral matrix to initially determine the number of components of the mixture. Finally, the Non-negative Matrix Factorization (NMF), Non-negative Matrix Factorization based on pure variables initialization (PNMF) and CNMF methods are used to the decomposition of the mixture data matrix and spectral analysis of the aliased spectrum. The results show that the separation effect of the CNMF algorithm is better than that of NMF and PNMf method, and the corresponding results of the characteristic absorption peak are accurate. In addition, the correlation coefficients of separation results for different component mixtures are higher than 89%, and the spectral angles are less than 0.5 with a higher reduction degree of purity spectrum. Therefore, the constrained non-negative matrix factorization algorithm is introduced into the separation of terahertz aliasing spectra, which is preferable to extract the characteristic information of single components in complex mixtures and provides a better foundation for the qualitative analysis and quantitative calculation of subsequent terahertz multi-component mixtures as well as the considerable research prospects in the field of quality testing of green tires and rubber.

**Keywords** THz-TDS; Spectral separation; Non-negative matrix factorization; Tire rubber; Black analysis system

(Received Nov. 7, 2019; accepted Mar. 21, 2020)

\* Corresponding author