

文章编号: 2095-4980(2016)01-0026-05

## 煤炭标准物质的太赫兹光谱聚类分析

詹洪磊<sup>a,b</sup>, 王玉霞<sup>a</sup>, 王雪松<sup>a</sup>, 赵 昆<sup>a,b\*</sup>

(中国石油大学 a.油气光学探测技术北京市重点实验室; b.全国石油和化工行业油气太赫兹波谱与光电检测重点实验室, 北京 102249)

**摘要:** 作为一种广泛应用的化石能源, 煤炭的在线监测非常必要。太赫兹技术是一种有效的在线检测方法, 可作为传统技术的补充手段应用于煤炭的表征。基于9种煤炭标准试样的太赫兹时域光谱, 计算了不同煤炭在0.1 THz~1.5 THz频段的折射率和吸收系数谱。将该频段内所有折射率和吸收系数数据作为输入变量, 分别进行聚类分析, 根据聚类结果, 对应煤炭标准物质的已知参量, 证明碳含量是影响煤炭折射率的主要因素, 氢氮含量是影响吸收效应的主要因素。结果表明, 太赫兹光谱结合聚类分析方法可对煤炭之间的相似性和相异性进行定性分析, 为煤炭工业中的实时在线监测提供技术支持。

**关键词:** 煤炭; 太赫兹光谱; 光学参数; 聚类分析

**中图分类号:** TN761

**文献标识码:** A

**doi:** 10.11805/TKYDA201601.0026

## Cluster analysis concerning the terahertz spectroscopy of coal materials

ZHAN Honglei<sup>a,b</sup>, WANG Yuxia<sup>a</sup>, WANG Xuesong<sup>a</sup>, ZHAO Kun<sup>a,b\*</sup>

(a.State Key Laboratory of Petroleum Resources and Prospecting; b.CPCIF Key Laboratory of Oil and Gas Terahertz Spectroscopy and Photoelectric Detection, China University of Petroleum, Beijing 102249, China)

**Abstract:** The on-line detection is necessary for the characterization of coal, a kind of fossil resource widely applied in industry. Terahertz(THz) spectroscopy is an efficient method for on-line determination and can act as the supplementary tool of the traditional methods in this field. The refractive index and absorption coefficient spectra of nine coal materials are initially calculated from 0.1 THz to 1.5 THz based on the terahertz time-domain spectra. Cluster Analysis(CA) is then employed with the input of refractive and absorption spectra over the whole range from 0.1 THz to 1.5 THz. According to the CA dendrogram and the tested physical properties of coals, a conclusion can be drawn that the carbon content is the main factor of refractive effect, and hydrogen as well as nitrogen play a key role in absorption, respectively. Therefore, the analysis of the similarity and diversity can be realized by THz technique combined with CA. This combination can support the constant and on-line monitoring in coal industry.

**Key words:** coal materials; THz spectroscopy; optical parameters; Cluster Analysis

煤炭是世界上分布最为广阔的化石能源。中国是世界上最早利用煤进行生产、生活的国家, 煤炭资源丰富, 煤炭探明储量占我国化石能源探明储量的90%以上。同时, 煤炭也一直是我国能源消费的主体, 在一次能源生产和消费结构中, 煤炭的比例分别达到约80%和70%<sup>[1-2]</sup>。由于现代化的煤炭开采利用技术飞速发展, 煤炭曾是并将一直是工业生产和日常生活的重要能源<sup>[3]</sup>。

煤炭的成分主要分为有机质和无机质2部分, 有机质是燃烧的主体, 有机质的主要元素有碳、氢、氧、氮和硫等, 还含有少量的磷、氯等元素。在煤炭的工业生产和质量评价体系中, 碳含量是工业中煤发热量计算的重要参数, 同时, 碳含量和挥发分是表征煤变质程度的重要指标, 一般随变质程度的增高而增大, 在确定煤炭加工工艺时具有重要的指导作用<sup>[4]</sup>。氢和碳是有机质的主体, 是煤炭燃烧过程中产生热量的元素<sup>[5]</sup>。氮不产生热量, 但在高温下会变成氮氧化物, 排在空气中会对环境造成影响。燃烧时大部分硫元素会转化成二氧化硫, 既腐蚀设

收稿日期: 2015-04-22; 修回日期: 2015-05-31

基金项目: 国家重大科学仪器开发专项基金资助项目(2012YQ140005); 国家自然科学基金资助项目(11574401,61405259)

\*通信作者: 赵 昆 zhk@cup.edu.cn

备，也污染大气，危害动植物健康。灰分主要是煤炭中不能燃烧的物质，也是煤质指标评价体系中的重要一员，灰分越高，煤炭燃烧的热效率往往越低。

在煤炭的生产和加工利用过程中，对于几种重要指标的监控既需要准确的实验室分析，也需要快速的在线检测，两者有机结合，缺一不可。因此，需要多种方法、多种技术联合使用。近年来，作为一项在线检测技术<sup>[6]</sup>，太赫兹光谱已被应用到生物、石油等领域，具有良好的实用性和发展前景<sup>[7-10]</sup>。滕学明、许长虹等人利用太赫兹时域光谱对煤炭标准物质进行初步的探索，证明了不同煤质样品在太赫兹波段具有不同的响应，且太赫兹常数与煤炭的某几种成分具有一定的对应关系<sup>[11-12]</sup>。本文在测得太赫兹光学参数谱的基础上，引入一种数学统计分析方法——聚类分析，对折射率和吸收系数谱进行数据再处理，并把聚类结果和煤炭的主要物理性质联系起来，由此确定煤炭中各种成分对其太赫兹响应的影响。

## 1 实验材料与方法

实验的测试装备为购置于美国 Zomega 公司的透射式太赫兹时域光谱仪，其光源为 Maitai 钛蓝宝石飞秒激光器产生的飞秒激光，中心波长为 800 nm。通过衰减片获得适当功率后，一部分激光作为泵浦光来激发 GaAs(砷化镓)晶体产生太赫兹脉冲，太赫兹脉冲通过样品后与另一束探测激光共同作用于 ZnTe(碲化锌)晶体，将光信号转换为电信号，采用锁相放大器对信号进行放大，通过软件即可采集带有样品信息的太赫兹时域光谱<sup>[8]</sup>。为消除空气中水蒸气的太赫兹波的影响，光谱测试始终在氮气气氛中进行，测试温度为 21 °C±0.3 °C，湿度小于 3%。实验样品为煤炭标准物质，其基本物理性质参数已知，由国家煤炭质量监督中心提供，见表 1。大块样品经过粉碎得到粒径小于 200 μm 的煤炭粉末。采用压片机对煤炭样品进行压片，为保证样品的厚度(厚度太小样品易破碎)，压片前掺入在太赫兹波段吸收率极小的聚乙烯粉末，并按聚乙烯/煤炭质量比为 2:1 的比例混合，压片得到厚度为 1.5 mm，直径 30 mm 的圆片<sup>[12]</sup>。

表 1 9 种煤炭标准物质的标号及对应的性质参数  
Table 1 The physical properties and number of 9 coal materials

content(m%)	coal number								
	101p	101q	103h	104g	109g	110g	111f	113e	126a
sulfur	0.51	0.49	0.40	1.19	2.81	4.43	1.77	3.05	0.26
Ash	8.48	9.63	13.16	13.8	28.78	26.31	20.36	24.45	14.56
volatile component	33.97	22.19	9.99	5.58	29.43	18.34	29.96	10.96	5.82
carbon	77.03	79.21	78.73	79.77	56.98	61.14	64.08	64.84	80.00
hydrogen	4.70	4.21	2.66	2.14	3.91	2.98	3.97	2.87	0.95
nitrogen	1.40	1.40	0.98	0.92	0.94	0.96	1.11	1.06	0.24

为了分析不同煤炭标准物质之间的相似性，本文采用一种常用的统计分类方法——聚类分析(CA)，对不同煤炭样品的太赫兹参数谱进行数据再处理。聚类分析基于“物以类聚”的基本原理，是一种将物理或抽象对象集合分组为有类似对象组成的多个类的数学分析方法。在聚类过程中，不需要人为地给定一个分类标准，而是从所分析的数据出发，计算不同样品两两之间的欧氏距离，拥有最小欧氏距离的 2 个样品合并为一类，并作为一个新的样品参与计算，欧氏距离最小的样品再次成为一个新的类，依次进行，每次减少一类，直到所有的样品都合并为一类，可得到一个聚类树图，清晰显示样品的分类情况。欧氏距离的计算公式为： $d^2=(X_m-X_n)(X_m-X_n)'$ ，其中  $X_m$  和  $X_n$  分别为第  $m$  和第  $n$  个样品的光谱数据向量。

## 2 结果和讨论

透射式太赫兹时域光谱仪可直接测得样品和参照物的时域光谱，本课题的背景光谱为太赫兹波通过氮气环境后的时域信号。对背景和参照物时域光谱分别作快速傅里叶变换后，可求得样品的折射率和吸收系数谱。图 1 表示 9 种煤炭标准物质在 0.1 THz~1.5 THz 波段折射率随频率变化

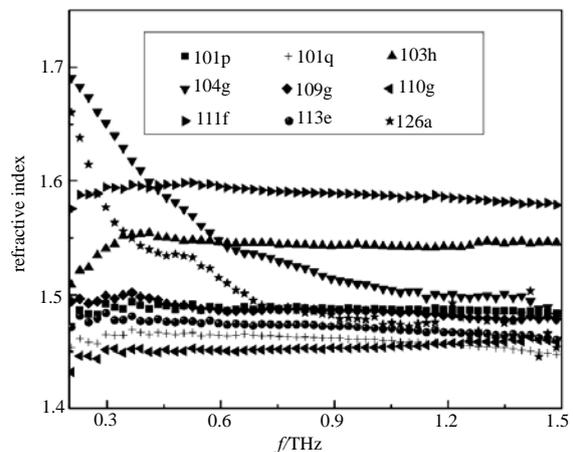


Fig.1 Refractive index spectra of nine coal materials from 0.1 THz to 1.5 THz  
图 1 9 种煤炭试样在 0.1 THz~1.5 THz 频段内的折射率谱

的光谱图,由图可知 9 种煤炭样品在太赫兹波段的折射率具有显著的差异,标号为 111f 的试样具有最大的折射率,整个频段内保持在 1.58~1.6 之间。其次是 103h,其折射率大约为 1.55。因此,111f 和 103h 的折射率相对于其他样品的区别最明显,在煤炭的快速识别中可用太赫兹技术对其首先进行标定。同时,104g 和 126a 2 种煤炭物质的折射率存在明显的先减小后不变的趋势,此外,126a 的折射率在 1.2 THz~1.5 THz 范围内出现了一定的波动,说明这 2 种样品在太赫兹波段存在较强的频散效应,可能与物质本身的性质和组成结构有关<sup>[9]</sup>。

太赫兹辐射穿过样品时,由于样品的内部分子振动对太赫兹光子产生吸收效应,使得样品的太赫兹信号相对于参考信号发生不同程度的衰减。图 2 表示 9 种样品在 0.1 THz~1.5 THz 频段内的吸收系数谱,在给定的频率区间内,未发现明显的吸收峰,但 126a 的吸收系数

在 1.3 THz 附近有一定的吸收增强效应,可能是由于该样品里存在微量 C/H 比较大的环状烷烃(相对于其他样品,126a 的 C 含量最大, H 含量最小)。9 种样品在各个频率处的吸收都存在不同程度差异,说明煤炭试样内部不同的元素组成和结构在太赫兹波段具有显著的体现。相比较而言,试样 126a 的吸收效应最为明显,在整个波段都具有最大的吸收系数。104g 次之,但其吸收谱较为平稳,吸收系数随频率的变化较小。其他煤炭试样的吸收系数都在 0~40  $\text{cm}^{-1}$  范围内,相比于 126a 和 104g,它们在太赫兹波段的吸收效应较为微弱。9 种试样在太赫兹各个频率处吸收的相对大小是基本一致的,例如在 1.0 THz 频率处的吸收从小到大依次为:101p,101q,113e,111f,109g,110g,103h,104g,126a。其他频率的吸收大小顺序也大抵如此。

为了定性分析 9 种煤炭试样各样品之间的相似性以及不同成分对太赫兹参数的影响,将整个频段的所有光谱数据作为输入数据集,采用聚类分析方法对其进行分类计算。对 9 个样本进行聚类分析,经过 8 步逐步聚类,得到表征 9 种煤炭相似和相异性的聚类树形图。图 3 表示将折射率谱作为输入变量所得到的树形图,结果显示 9 种煤炭可分为 2 大类,103h 和 111f 作为第 1 大类,其余 7 个样本为第 2 类,第 1 类中 103h 和 111f 间的欧氏距离为 0.034,大于第 2 类中任何 2 个样品。在第 2 类中,101p 和 104g 的距离最小,它们所组成的新类与 109g 间的欧氏距离也较小,说明这 3 种样本在太赫兹波段的折射率较为相似。类似地,样本 110g 和 113e 之间有较小的欧氏距离,其新类与 126a 也保持着较相似的折射率。

通过比较聚类结果与样本已知的物理性质可知,碳含量是影响煤炭物质在太赫兹波段折射率的主要原因。101p 和 104g 的碳含量分别为 77.03% 和 79.77%,相对误差为 3.5%,其他成分之间的相对误差均在 20% 以上;同时,110g 和 113e 的碳含量分别为 61.14% 和 64.84%,相对误差为 5.9%;103h 和 111f 之间的欧氏距离稍大,其碳含量之间的相对误差也相较起来更大(20%)。各样品之间碳含量的相对值与折射率的差异具有一定的对应关系。此外,其他参数,如氢含量、灰分、挥发分、氮含量、硫含量等也在不同程度上影响着煤炭在太赫兹频段的折射率。101p 与 109g 的氢含量较为相近,挥发分的差值也较小;104g 和 109g 的氮含量几乎相等;101q 与 113e 的氮含量相差较小;113e 和 126a 的欧氏距离较小,但它们的各个物理性质均具有较大的差异,证明其相近的折射率是各物理参数综合作用的结果。因此,煤炭在太赫兹频段的折射率是由所有物理性质综合决定的,其中,碳元素含量是最重要的影响因素。

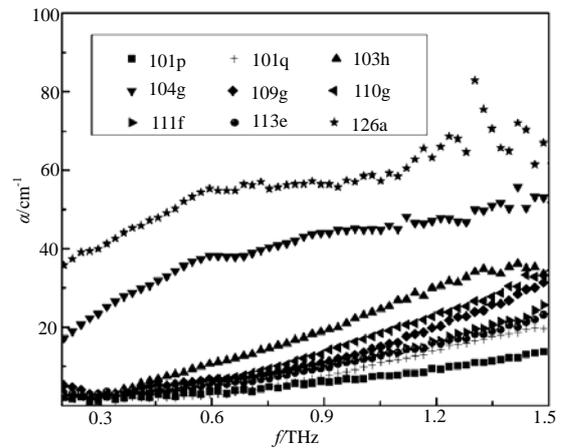


Fig.2 Absorption coefficient( $\alpha$ ) spectra of nine coal materials from 0.1 THz to 1.5 THz

图 2 9 种煤炭试样在 0.1 THz ~1.5 THz 频段内的吸收系数( $\alpha$ )谱

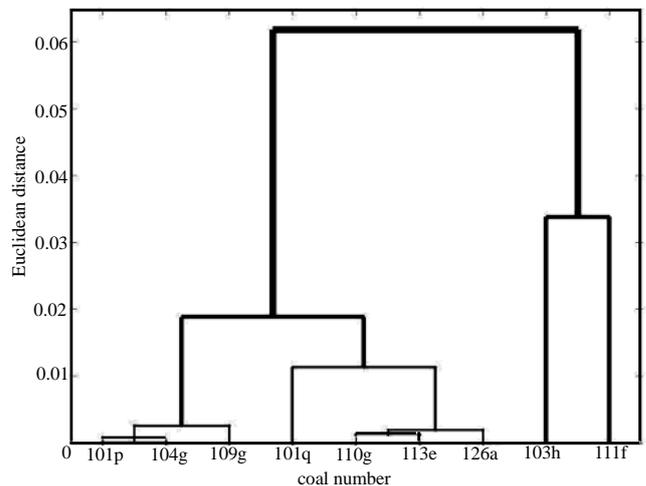


Fig.3 Euclidean distance dendrogram with the input of refractive index spectra from 0.1 THz to 1.5 THz

图 3 基于 0.1 THz~1.5THz 频段折射率谱的聚类树

为进一步讨论和验证煤炭碳含量对样品折射率的影响,随机选择 2 个频率 0.11 THz 和 1.0 THz 处的折射率,并把它们与碳含量联系起来。见图 4,在 0.11 THz 频率处,随着碳含量的增加,折射率缓慢上升,当碳含量较高时,折射率急速上升。在 1.0 THz 频率处,随着碳含量的增多,折射率依然缓慢增加,但其增幅相对 0.11 THz 处较小,原因是在太赫兹频段内,部分煤炭样品存在较强的频散效应,2 个频率处的变化趋势并不完全一样,这与图 1 的现象是完全一致的。此外,碳含量增加,折射率增加,证明碳含量是影响折射率的重要元素。但在碳含量较低时,折射率的增幅较小,说明碳含量并非是影响折射率的唯一要素。因此,碳含量应是影响煤炭折射率的主要因素,同时,折射率的相对大小是多种成分共同作用的结果。

与图 3 的计算过程类似,将 0.1 THz~1.5 THz 频段内的所有吸收数据作为聚类分析的输入变量,可计算得到基于吸收系数谱的聚类树形图,如图 5 所示。从图 2 可知,126a 和 104g 在太赫兹波段具有最大的吸收系数。对应于图 5 的树形图,104g 和 126a 被归为一类,且具有较大的欧氏距离。在 9 种煤炭样品中,103h 和 110g 的欧氏距离最小,相似性最高,103h 和 110g 组成的新类与 109g 次之,111f 和 113e 紧随其后,101q 和 101p 也被依次与上一步的新类聚集。由于样品的吸收系数数值相对较大,因此聚类树中吸收系数的最大距离(16.4)远大于折射率的最大距离(0.062)。

对于在吸收系数聚类树中最相近的 103h 和 110g,它们的氢含量和氮含量都极为相近(氢含量分别为 2.66% 和 2.98%,氮含量分别为 0.98% 和 0.96%),对 109g 而言,其氮含量与 103h 和 110g 也极为相近(109g 的氮含量为 0.94%),氢含量的差值也较小,此外,109g 和 110g 的灰分含量的差值也在一个较小的范围;对于拥有较小欧氏距离的 111f 和 113e 而言,它们的氮含量、灰分和氢含量之间的差异均较小,特别是氮含量,其差值几乎可以忽略(分别为 1.11% 和 1.06%);类似地,101q 的氮含量为 1.4%,氢含量为 4.21%,都与 111f 所对应的参数值极为相近;101p 的情形与 101q 基本一致,同时,101p 和 101q 的硫含量和灰分含量的差值也保持在一个较小的范围;对于归属同类但欧氏距离较大的 104g 和 126a 而言,它们的氢含量、氮含量差异较大,但其灰分和挥发分的含量极为相似。综合 9 个样品的吸收系数聚类结果可知,氢含量和氮含量是影响煤炭在太赫兹波段吸收效应大小的关键因素,这可能是由于氢含量和氮含量在某种程度上反映了煤炭中水分及有机质的百分比(尽管碳是有机质的重要元素,但煤炭中的碳成分既来源于有机质,又来源于无机质成分,如碳酸盐体系( $\text{CaCO}_3$ ,  $\text{Fe}_2(\text{CO}_3)_2$  等),因此碳元素不能完全反映有机质的含量),有机物特别是极性物质对太赫兹辐射往往都具有较强的吸收。同时,其他成分对煤炭的太赫兹吸收效应也具有不同程度的影响。

基于煤炭冶炼工业和环境污染治理中所涉及的关键指标和关键问题,研究了具有代表性的煤炭类型和参数指标。本文采用聚类分析法,以整个频段内的太赫兹光谱为输入变量,研究了煤的主要成分与其在太赫兹波段响应之间的相互关系。这种计算方法考虑了煤炭样品在太赫兹整个波段的折射和吸收响应,所得结果可信度高。尤其对于没有特征吸收峰的样本,考虑其在整个波段的太赫兹响应显得更加重要。

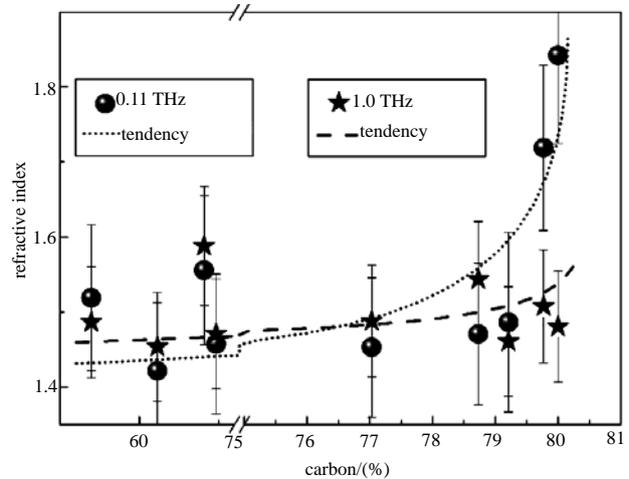


Fig.4 Tendency of carbon dependent refractive index at 0.11 THz and 1.0 THz  
图 4 0.11 THz 和 1.0 THz 频率处折射率随碳含量的变化趋势

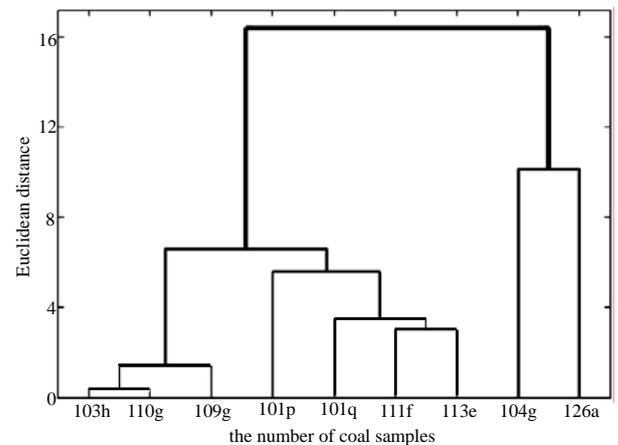


Fig.5 Euclidean distance dendrogram with the input of absorption coefficient spectra from 0.1 THz to 1.5 THz  
图 5 基于 0.1 THz~1.5 THz 频段吸收系数谱的聚类树

### 3 结论

综上所述,本文对 9 种煤炭标准物质的太赫兹光谱进行了测试,分别计算了煤炭样本的折射率谱和吸收系数谱,结果表明成分含量不同的煤炭在太赫兹光学参数谱中表现出明显的差别。基于所得到的折射率谱和吸收系数谱,采用聚类分析方法,分别对煤炭样品在太赫兹波段的折射率和吸收效应进行了进一步讨论。结合各样本的已知参数及聚类分析结果,表明各成分均对煤炭在太赫兹波段的折射和吸收效应具有不同程度的影响,其中碳含量是影响煤炭折射率的主要因素,氢含量和氮含量是影响煤炭吸收系数的主要因素。太赫兹光谱对不同煤炭的成分及结构的微小差异非常敏感,该差异在太赫兹折射率和吸收谱上能得到充分的体现,聚类分析技术更是能从大量的数据中判别煤炭之间的相似相异性。通过增加煤炭的种类数量,可利用数学分析方法对样本的太赫兹光学参数进行定性和定量分析,建立数据库,推广成为煤炭冶炼和环境污染治理过程实时在线监测的重要参考依据。

#### 参考文献:

- [1] 谢和平,钱鸣高,彭苏萍,等. 煤炭科学产能及发展战略初探[J]. 中国工程科学, 2011,13(6):44-50. (XIE Heping, QIAN Minggao, PENG Suping, et al. Sustainable capacity of coal mining and its strategic plan[J]. Engineering Sciences, 2011,13(6):44-50.)
- [2] 申宝宏,雷毅,郭玉辉. 中国煤炭科学技术新进展[J]. 煤炭学报, 2011,36(11):1779-1783. (SHEN Baohong, LEI Yi, GUO Yuhui. Progress of coal science and technology in China[J]. Journal of China Coal Society, 2011,36(11):1779-1783.)
- [3] 谢和平,王金华,申宝宏,等. 煤炭开采新理念—科学开采与科学产能[J]. 煤炭学报, 2012,37(7):1069-1079. (XIE Heping, WANG Jinhua, SHEN Baohong, et al. New idea of coal mining: scientific mining and sustainable mining capacity[J]. Journal of China Coal Society, 2012,37(7):1069-1079.)
- [4] LEE B, SONG J, KIM R, et al. Simulation of the influence of the coal volatile matter content on fuel NO emissions in a drop-tube furnace[J]. Energy Fuels, 2010,24(8):4333-4340.
- [5] YAO H M, VUTHALURU H B, TADÉ M O, et al. Artificial neural network-based prediction of hydrogen content of coal in power station boilers[J]. Fuel, 2005,84:1535-1542.
- [6] 王波,张岩. 太赫兹超材料和超表面器件的研发与应用[J]. 太赫兹科学与电子信息学报, 2015,13(1):1-12. (WANG Bo, ZHANG Yan. Design and applications of THz metamaterials and metasurfaces[J]. Journal of Terahertz Science and Electronic Information Technology, 2015,13(1):1-12.)
- [7] DRAGOMAN D, DRAGOMAN M. Terahertz fields and applications[J]. Progress in Quantum Electronics, 2004,28(1):1-66.
- [8] ZHAN H L, WU S X, BAO R M, et al. Water adsorption dynamics in active carbon probed by terahertz spectroscopy[J]. RSC Advances, 2015(5):14389-14392.
- [9] ZHAN H L, WU S X, BAO R M, et al. Qualitative identification of crude oils from different oil fields using terahertz time-domain spectroscopy[J]. Fuel, 2015,143:189-193.
- [10] GE L, ZHAN H, LENG W, et al. Optical characterization of the principal hydrocarbon components in natural gas using terahertz spectroscopy[J]. Energy Fuels, 2015,29(3):1622-1627.
- [11] 滕学明,赵昆,赵卉,等. 利用太赫兹技术研究煤炭中的灰分含量与碳含量[J]. 现代科学仪器, 2011(6):19-21. (TENG Xueming, ZHAO Kun, ZHAO Hui, et al. Analysis of time-resolved terahertz spectroscopy of ash contents and carbon contents of coal[J]. Modern Scientific Instruments, 2011(6):19-21.)
- [12] 许长虹,滕学明,赵卉,等. 煤炭中氢含量与挥发分的太赫兹时域光谱研究[J]. 现代科学仪器, 2013(4):228-230. (XU Changhong, TENG Xueming, ZHAO Hui, et al. Analysis of hydrogen and volatile matter content in coal maceral using time-resolved terahertz spectroscopy[J]. Modern Scientific Instruments, 2013(4):228-230.)

#### 作者简介:



詹洪磊(1991-),男,贵州省铜仁市人,在读博士研究生,研究方向为太赫兹光谱计量和纳米岩石物理.email:hlzhan@126.com.

王玉霞(1988-),女,石家庄市人,在读硕士研究生,研究方向为光传感与光探测器件及技术。

王雪松(1990-),女,辽宁省阜新市人,在读硕士研究生,研究方向为燃煤电厂尾气脱硝处理。

赵昆(1971-),男,山东省聊城市人,博士,教授,博士生导师,研究方向为油气光学工程。