

水面油污染物差分偏振光谱信号预处理算法研究

袁越明 方勇华 崔方晓 李大成

(中国科学院安徽光学精密机械研究所 中国科学院通用光学定标与表征技术重点实验室, 安徽 合肥 230031)

摘要 采用差分偏振光谱法在 $3\sim 5\ \mu\text{m}$ 波段对水面溢油污染物进行被动遥感时, 所测差分偏振光谱是含有强大气吸收光谱信号与油污染物目标弱光谱信号的混合谱, 这给油污染物光谱特征识别带来了困难。另外, 受环境因素以及油膜自身张力影响, 水面油膜厚度分布以及油膜表面粗糙度在测量过程中发生变化, 从而使得连续测量的差分偏振光谱中油污染物光谱信息含量存在不同。利用这一特点, 基于固定点迭代的快速主成分分析算法 FastPCA 设计了水面溢油污染物差分偏振光谱信号预处理算法。实验结果表明, 该算法可以有效地将水面油污染物目标光谱特征信息从具有强大气吸收的混合差分偏振光谱信号中提取出来, 通过光谱重构得到油污染物光谱特征信号, 可用于进一步的定性、定量分析。

关键词 遥感; 差分偏振; 溢油污染; 主成分分析

中图分类号 O436.3; TN219 文献标识码 A doi: 10.3788/AOS201131.1128001

Research on Preprocessing Algorithm for Differential Polarization Spectrum of Oil Spills on Water

Yuan Yueming Fang Yonghua Cui Fangxiao Li Dacheng

(Key Laboratory of Optical Calibration and Characterization, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, Anhui 230031, China)

Abstract When the oil spills on water are passively detected by differential polarization Fourier transform infrared (FTIR) spectroscopy in the band of $3\sim 5\ \mu\text{m}$, the measured differential polarization spectrum will be a mixed spectrum of the strong atmospheric absorption signal and the minute signal of oil pollutant. It brings complexity to the identification of oil spills on water. In addition, the thickness distribution and the surface coarseness of oil film on water change with the influence of environmental factors and the tension of oil film. This makes the content of effective spectral information, which is contained in the differential polarization spectrum of oil spills on water, change continuously during the process of measurement. Using this characteristic, an algorithm based on a fast fixed-point algorithm for principal component analysis (FastPCA) is proposed for preprocessing the differential polarization spectrum of oil pollutant on water. Experimental results show that this algorithm can separate the spectral information of oil pollutant from the differential polarization spectrum with strong atmospheric absorption signal. The reconstructed spectral characteristic signal of oil pollutant can be used for the further qualitative and quantitative analyses.

Key words remote sensing; differential polarization; oil spill; principal component analysis

OCIS codes 300.6300; 070.4790

1 引 言

随着工业的发展, 人类对能源的需求急剧增加, 原油及其制品已成为目前最主要的能源之一。然而, 伴随着海洋运输业和海洋石油开采业的迅猛发

展, 溢油事故屡见不鲜, 严重危害到海洋生态环境, 造成海洋鱼类、鸟类、海藻和海洋哺乳动物的大量死亡。因此, 防治海面溢油污染已成为当今世界各国共同关注的问题。

收稿日期: 2011-04-12; 收到修改稿日期: 2011-05-30

基金项目: 国家自然科学基金(30860010)资助课题。

作者简介: 袁越明(1983—), 男, 博士研究生, 主要从事红外偏振遥感方面的研究。E-mail: billion_yy@hotmail.com

导师简介: 方勇华(1966—), 男, 研究员, 博士生导师, 主要从事遥感光电信息技术方面的研究。E-mail: yhfang@aiofm.

近年来,遥感技术作为一种能够及时、准确、全面地监测河流、湖泊和海洋等水域油污染状况的手段,正受到越来越广泛的关注。将傅里叶变换红外(FTIR)光谱遥感技术与偏振探测技术相结合所产生的差分偏振光谱探测水面溢油污染的遥感方法,已经通过实验初步验证了其可行性^[1,2],并且该方法在实现水面溢油污染无背景探测和溢油种类鉴别方面,弥补了传统热红外遥感探测水面溢油污染的不足。

但是,采用差分偏振光谱法在 3~5 μm 波段对水面溢油污染物进行被动遥感时,为了保证目标特征信息的探测对比度,需要以太阳作为向下辐射源。这导致所获得的差分偏振光谱是含有油污染物光谱特征信息与太阳透过大气的辐射光谱中大气吸收光谱信息的混合谱,且油污染物光谱特征作为微弱信号叠加于大气吸收特征之上,因此无法从差分偏振光谱中直接识别油污染物光谱特征信息,给进一步定性、定量分析带来困难。本文基于固定点迭代的快速主成分分析算法 FastPCA,提出水面油污污染物的差分偏振光谱信号预处理算法。该算法能够有效地从差分偏振光谱中提取油污染物光谱特征信息,通过光谱重构得到油污染物光谱特征信号,可用于进一步定性、定量分析。

2 理论基础

2.1 差分偏振光谱探测模型

图 1 为室外差分偏振光谱法探测水面油膜时的辐射传输模型,其中 s 偏振方向为垂直于探测器观测平面的电场振动方向,p 偏振方向为平行于探测器观测平面的电场振动方向。假定油膜表面发生的反射为镜面反射。则探测器以某一观测角 θ 观测水面油膜时,油膜表面 s 和 p 偏振方向的反射率表示为 $R_{cont}^{s,p}$,透射率表示为 $\tau_{cont}^{s,p}$ 。在探测波段,反射率 $R_{cont}^{s,p}$ 和透射率 $\tau_{cont}^{s,p}$ 随波长的变化而变化。以天顶角

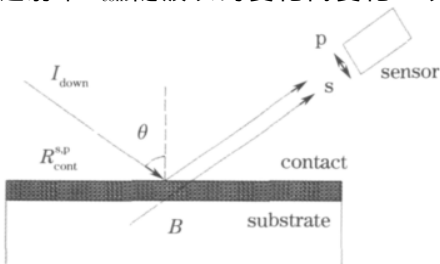


图 1 差分偏振光谱探测水面油膜时的辐射传输模型

Fig. 1 Model of radiation transmission with differential polarization FTIR spectroscopy for detection of oil film on water

θ 入射到油膜表面的辐射能量为 L_{down} 。由于油膜与水面相接触,假定温差 $\Delta T \approx 0$,故油膜表面温度 T 对应的黑体辐亮度 B ,可由普朗克公式计算求得^[3]。

当水足够深时,油膜和水可以看成是一个不透光的整体^[4,5]。此时 $\tau_{cont}^{s,p} \approx 0$,油膜表面沿观测角方向 s 和 p 偏振的发射率可表示为 $(1 - R_{cont}^{s,p})$,则探测器在 s 和 p 偏振方向上探到的辐射能量 L_{cont}^s 和 L_{cont}^p 为

$$L_{cont}^{s,p} = L_{sky} R_{cont}^{s,p} + B(1 - R_{cont}^{s,p}), \quad (1)$$

式中 $L_{down} R_{cont}^{s,p}$ 为天空辐射在油膜表面的反射能量, $B(1 - R_{cont}^{s,p})$ 为水体和油膜沿观测方向的自身辐射能量。

差分偏振光谱可定义为两个偏振方向上的辐射改变量,即

$$\Delta L^{pol} = L_{cont}^s - L_{cont}^p = (R_{cont}^s - R_{cont}^p)(L_{down} - B), \quad (2)$$

式中 $(R_{cont}^s - R_{cont}^p)$ 为油膜表面两个偏振方向上的反射率差值,包含了所探测目标的特征信息, $(L_{down} - B)$ 与目标特征的探测对比度有关。当 L_{down} 与 B 很接近时,会导致(2)式乘积 ΔL^{pol} 很小,以至于所得差分偏振光谱的信噪比很低而可能无法探测到目标的特征信息。因此需要增大 $|L_{down} - B|$,来保证目标特征信息的探测对比度。

2.2 主成分分析

主成分分析是一种建立在统计最优基础上的分析方法^[6,7],其目的是将原变量进行变换,使数目较少的互不相干的新变量(即主成分)成为原变量的线性组合,新变量最大限度地表征原变量的数据结构特征,且不丢失有用信息。其过程是对原坐标系进行平移和旋转变换,使得新坐标系的原点与样本点集合的中心重合,新坐标系的第一轴与数据变异的最大方向对应,新坐标系的第二轴与第一轴正交,并且对应于数据变异的第二大方向,依此类推。这些主轴分别被称为第一主轴 a_1 ,第二主轴 a_2 等。若由新主轴 a_1, a_2, \dots, a_m 构成的子空间能够十分有效地表征原数据的变异情况,且 $m \ll p$,则原来 p 维空间就被降至 m 维。新主轴构成的 m 维子空间被称为 m 维主超平面。

对于一个样本数据集 X ,包含 n 个 p 维样本,即 $X = (x_{i,j})_{n \times p}$ 。样本数据集在主超平面的 m 个主轴上的投影所得综合变量即为主成分, m 个主成分构成主成分矩阵,即 $F = [F_1, F_2, \dots, F_m]$ 。代表各个主轴的向量 a_1, a_2, \dots, a_m 构成主向量矩阵 $a = [a_1, a_2, \dots, a_m]$ 。主成分矩阵与主向量矩阵之间满足关系

$$F = Xa, \quad (3)$$

如果令 $\hat{X} = (\hat{x}_{i,j})_{n \times p}$ 为 n 个 p 维样本的重构数据矩阵, 则有

$$\hat{X} = Fa^T. \quad (4)$$

为了满足光谱实时处理的应用要求, 本文利用一种基于固定点迭代的快速主成分分析算法 FastPCA 来完成对数据的主成分分析^[8~12]。

2.3 基于 FastPCA 的水面油污染物差分偏振光谱信号预处理算法

前文已经提到, 在 $3 \sim 5 \mu\text{m}$ 波段以太阳作为向下辐射源探测到的水面油污染物的差分偏振光谱是油污染物特征光谱信息与太阳透过大气的辐射光谱中大气吸收光谱信息的混合谱, 且油污染物光谱特征作为微弱信号叠加于大气吸收特征之上, 因此无法从差分偏振光谱中直接识别油污染物光谱特征信息。而研究表明主成分分析技术具有从很强的混合光谱信号中提取微弱目标光谱特征的能力^[8]。鉴于连续测量一组差分偏振光谱时, 受环境因素以及油膜自身张力影响, 水面油膜厚度分布以及油膜表面粗糙度变化导致所测差分偏振光谱中油污染物光谱信息含量的变化, 因此如对一组连续测量的差分偏振光谱做主成分分析, 则油污染物光谱信息含量变化引起的数据变异会体现在某个主成分上, 油污染物光谱特征就体现在该主成分对应的主向量中, 这样就有可能从具有强大气吸收的混合差分偏振光谱信号中提取出微弱的水面油污染物目标光谱特征。

FastPCA 算法核心思想如下所述。由于主成分分析的目的是在数据空间中找到一组主向量 $a = [a_1, a_2, \dots, a_m]$ 能够满足重建数据和原始光谱的均方误差最小, 因此可以得到主成分分析的目标函数

$$E_{MS} = \|X - \hat{X}\|^2 = \|X - Fa^T\|^2. \quad (5)$$

当重建误差最小时, 原始光谱与重建光谱的均方差对主向量 a 取导数应该等于零, 即

$$\frac{\partial}{\partial a} E_{MS} = -F^T(X - Fa^T) = 0. \quad (6)$$

从而由固定点迭代方法可得主成分分析的两步迭代快速算法:

1) 更新主向量:

$$a^T = \frac{F^T X}{F^T F}. \quad (7)$$

根据固定点迭代方法原理, 提取每个主向量时, 任取单位向量 $a_i(0)$ 作为迭代初始值, 代入(7)式迭代直到收敛为止。

2) 满足各个主向量之间正交的要求。由于主

向量 a 是正交归一化矩阵, 所以可采用 Gram-Schmidt 正交分解法使各主向量之间正交化。假设已经提取了 $i-1$ 个主向量 a_1, a_2, \dots, a_{i-1} , 则在提取第 i 个主向量 a_i 时, 进行下轮迭代前应将 a_i 按下式作正交化:

$$a_i - \sum_{j=1}^{i-1} \langle a_i, a_j \rangle a_j \rightarrow a_i, \quad (8)$$

再将所得 a_i 归一化。

基于 FastPCA 的水面油污染物差分偏振光谱信号预处理算法可归纳为以下两步: 1) 将实验中获取的一组水面油污染物的差分偏振光谱信号利用 FastPCA 算法进行主成分分析, 获取含有污染物目标特征光谱的主向量及相应的主成分; 2) 利用主向量和主成分重构水面油污染物的特征光谱信号, 从而达到特征提取的目的。

3 实验与讨论

3.1 水面油污染物的差分偏振光谱实验

差分偏振光谱探测系统由光谱采集和偏振信息提取两部分组成。光谱采集部分由 Bruker 公司生产的傅里叶红外光谱仪 OPAG33 实现, 而偏振信息提取部分由 Throlabs 公司生产的 ZnSe 金属线栅偏振片实现。偏振片装卡在偏振片调整架上, 调整架与光谱仪之间螺纹连接, 通过旋转调整架上的套筒可以在 $0^\circ \sim 360^\circ$ 范围内改变偏振片方向。

实验中选用汽油和机油作为水面油污染物探测目标, 汽油和机油在 $2000 \sim 3300 \text{ cm}^{-1}$ 波段的吸光度谱分别如图 2 和图 3 所示, 其特征峰均在 $2800 \sim 3000 \text{ cm}^{-1}$ 之间。

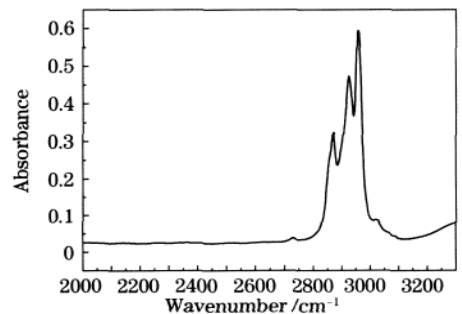


图 2 汽油吸光度谱

Fig. 2 Absorbance spectrum of gasoline

在室外对水面汽油油膜和机油油膜进行差分偏振光谱探测实验。在两个直径 $\phi = 42 \text{ cm}$, 深度 $H = 12 \text{ cm}$ 的水槽中接满水, 分别滴入 40 mL 汽油和 40 mL 机油, 待其自然展开后油膜厚度约为 $290 \mu\text{m}$ 。

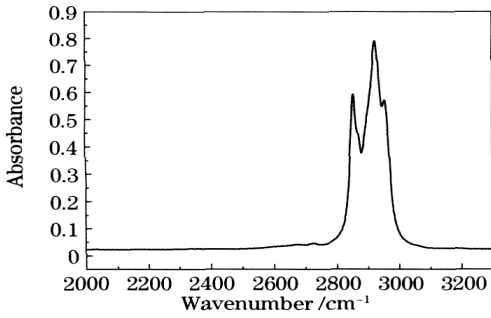


图 3 机油吸光度谱

Fig. 3 Absorbance spectrum of engine oil

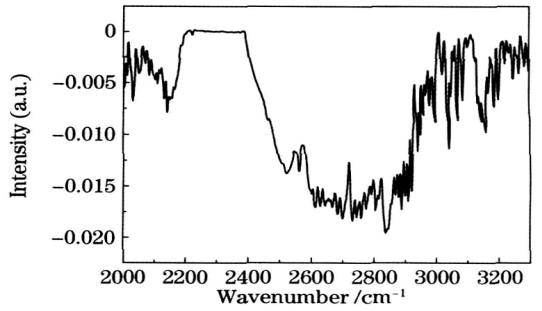


图 5 水面机油油膜差分偏振光谱

Fig. 5 Differential polarization spectrum of engine oil film on water surface

设置光谱采集范围为 $2000 \sim 3300 \text{ cm}^{-1}$ 波段, 探测器为 InSb 探测器, 光谱分辨率为 4 cm^{-1} , 光谱扫描平均次数为 10 次。为了保证目标特征信息的探测对比度, 实验过程中以太阳光作为向下辐射源, 即探测时太阳天顶角与仪器观测角相同, 仪器观测方向相对太阳入射方向的方位角为 180° 。分别测量水面汽油油膜和机油油膜在 s 和 p 偏振方向的偏振强度谱, 并通过(2)式计算得到差分偏振光谱, 如图 4, 5 所示。

由于以太阳光作为向下辐射源, 因此入射到油膜表面的向下辐射 L_{down} 为太阳透过大气的辐射光谱, 其中含有明显的大气吸收特征, 由(2)式可知 ($R_{\text{cont}}^s - R_{\text{cont}}^p$) 中含有汽油(机油)光谱信息, 而 ($L_{\text{down}} - B$) 中含有大气吸收光谱信息, 因此图 4 和图 5 中的差分偏振光谱是含有大气吸收光谱信息和油污染物光谱信息的混合谱, 并且从差分偏振光谱中无法直接识别 $2800 \sim 3000 \text{ cm}^{-1}$ 之间的油污染物光谱特征, 需要通过算法对其进行光谱预处理, 将微弱的油污染物光谱特征信息从大气吸收背景信息中分离出来。

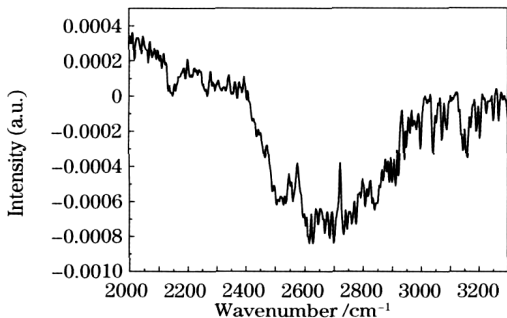


图 4 水面汽油油膜差分偏振光谱

Fig. 4 Differential polarization spectrum of gasoline film on water surface

3.2 水面油污物的差分偏振光谱预处理

对一组实测水面汽油油膜的差分偏振光谱利用 FastPCA 算法进行主成分分析, 获取主向量及其相应的主成分, 前 4 个主成分对应的主向量如图 6 所示。

从图 6 可以看到, 第一主向量 PV1 和原始光谱相似, 它与参与主成分分析的光谱的均值相关; 第二主向量 PV2 是低频的基线漂移信号, 经过分析表明, 该主向量体现的是由于受风等环境因素以及油膜自身张力的影响, 在连续测量 s 和 p 两个偏振方向偏振光谱的过程中, 水面油膜厚度分布以及油膜表面粗糙度存在变化, 即两个偏振方向测量的目标本身存在变化, 从而在计算差分偏振光谱时产生基线, 此时目标变化导致的数据变异信息已经大于污染物光谱特征幅度变化导致的数据变异; 第三主向量 PV3 为水面汽油油膜的光谱特征; 第四主向量 PV4 为随机噪声。

这里以参与主成分分析的一条水面汽油油膜差分偏振光谱为例进行分析, 采用 FastPAC 算法获得的 4 个主向量对应的主成分依次为 0.010056 , -0.00081052 , 0.00067172 和 -0.00019208 。利用主成分分析的重建表达式(4)式重构光谱, 与原谱比较, 如图 7 所示。重构信号几乎与原始信号重合, 说明上述分析结果是可靠的。但由于主成分分析过程是取少数几个主成分以达到降维的目的, 因此重建信号与原始信号之间存在误差。利用主成分分析中第三主向量 PV3 和它对应的主成分, 可以重构水面汽油污染物的光谱特征信号, 如图 8 所示。

图 8 即为通过主成分分析从水面汽油油膜差分偏振光谱中提取的水面汽油污染物光谱特征信号, 其特征峰与图 2 汽油吸光度谱中的特征峰位置一致。与图 7 中原光谱信号相比, 汽油特征峰的相对强度相比差分偏振光谱中大气吸收线的强度相当甚至更低, 另外在 $2500 \sim 3000 \text{ cm}^{-1}$ 波段大气透射率较高, 因此太阳透过大气的辐射光谱中在这一波段有一个较强的光谱突起, 这也直接反映在图 7 的差分偏振光谱中, 突起波段的相对强度约为汽油特征

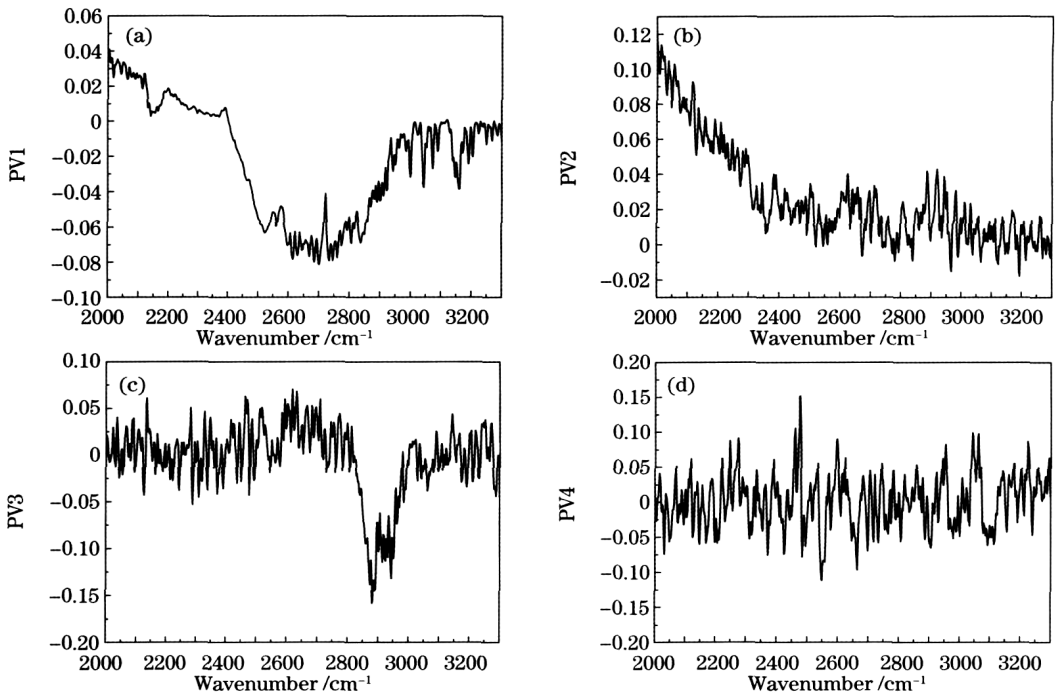


图 6 对水面汽油油膜差分偏振光谱做主成分分析所得主向量

Fig. 6 Principal vectors calculated from the differential polarization spectra of gasoline film water with FastPCA

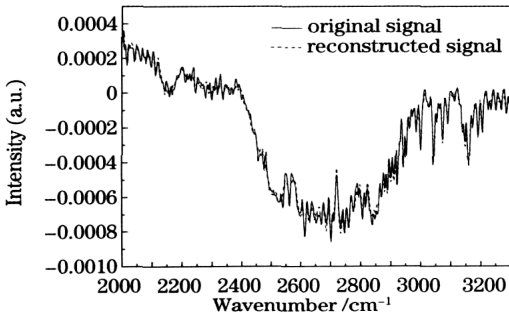


图 7 原始信号和重构信号

Fig. 7 Original signal and reconstructed signal

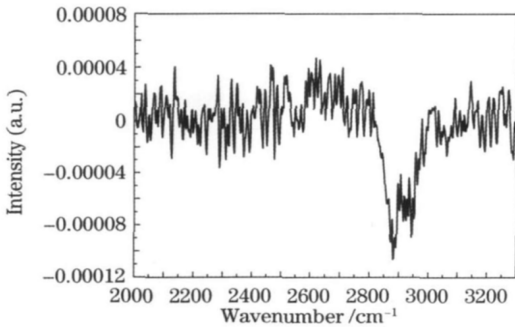


图 8 重构的汽油光谱特征信号

Fig. 8 Reconstructed characteristic signature signal of gasoline

峰相对强度的 10 倍。由此可以说明,汽油光谱特征信息相对于水面汽油差分偏振光谱信号中的大气吸

收光谱信息来说是弱信号,并且很难从差分偏振光谱中直接识别。而采用 FastPCA 算法对实验测得的差分偏振光谱进行主成分分析,可以将该微弱信号从混合光谱信号中分离出来,从而达到特征提取的目的,提取出的汽油污染物的光谱信息可用于进一步的定性、定量分析。

同样利用上述方法,对一组实测水面机油油膜的差分偏振光谱利用 FastPCA 算法进行主成分分析,前 4 个主成分对应的主向量如图 9 所示,其中第三主向量 PV3 为机油光谱特征信息。利用其中一条参与主成分分析的差分偏振光谱所得 4 个主向量对应的主成分 0.24852, 0.032997, 0.00022735, -0.0016518 重构机油光谱特征信号,如图 10 所示。

对比图 3 中的机油吸光度谱发现,在 $2800 \sim 3000 \text{ cm}^{-1}$ 波段,重构的机油光谱特征信号与机油吸光度谱特征峰位置一致,除 $3000 \sim 3300 \text{ cm}^{-1}$ 波段仍有一些大气吸收特征没有扣除干净外,其他波段已不再含有大气吸收信息。

利用差分偏振光谱法探测水面溢油污染时,需要根据油污污染物特征峰所属波段,选择相应的探测波段和该探测波段上探测灵敏度高的探测器。然而,对于大部分石油及其制品来说,其主要特征峰均在 $2800 \sim 3000 \text{ cm}^{-1}$ 之间,由甲基和亚甲基伸缩振动产生。因此,本研究中选取汽油和机油作为水面

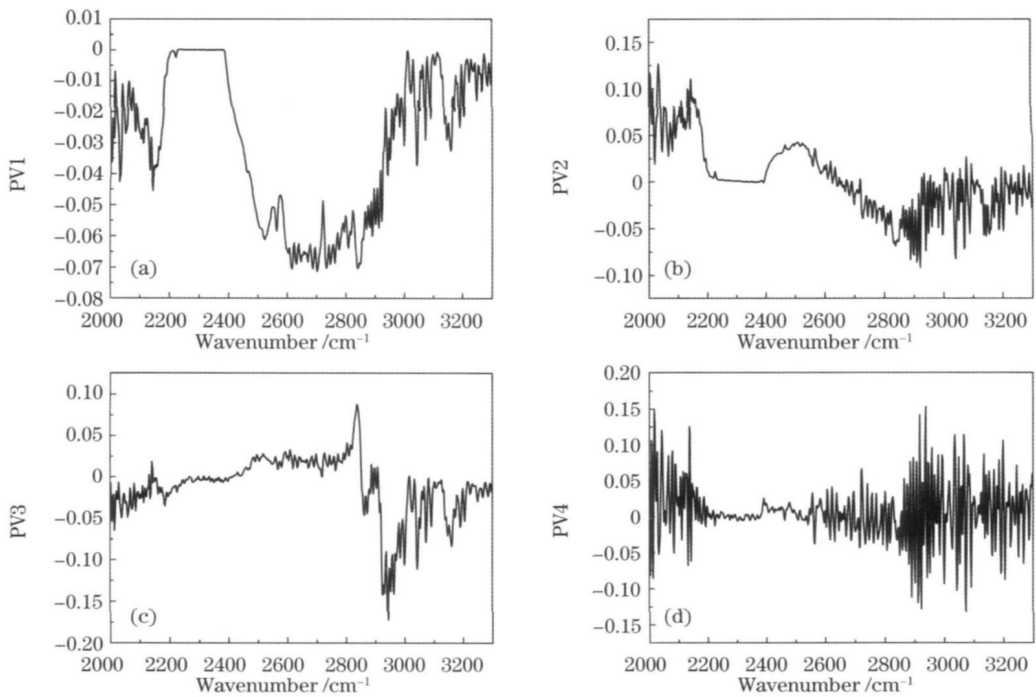


图 9 对水面机油油膜差分偏振光谱做主成分分析所得主向量

Fig. 9 Principal vectors calculated from the differential polarization spectra of engine oil film on water with FastPCA

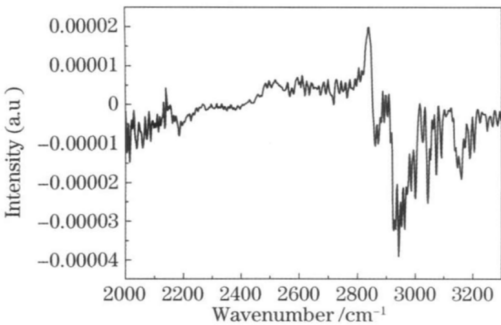


图 10 重构的机油光谱特征信号

Fig. 10 Reconstructed spectral characteristic signal of engine oil

油污染物目标进行光谱预处理算法研究具有代表性意义。

本文仅是对单一厚度的水面油膜进行差分偏振光谱探测,并开展差分偏振光谱预处理算法研究。今后研究中还需考虑不同的水面油膜厚度对油污染物差分偏振光谱特征提取的影响,为进一步开展差分偏振光谱法探测水面溢油污染的实用化研究奠定基础。

4 结 论

基于固定点迭代的快速主成分分析算法 FastPCA,设计了 3~5 μm 波段水面油污染物红外差分偏振光谱信号的预处理算法。该算法利用环境

因素作用导致的差分偏振光谱中油污染物信息含量的变化,从含有强大气吸收信息的差分偏振光谱中提取出微弱的油污染物光谱特征信号,可用于进一步的定性、定量分析。

参 考 文 献

- 1 Yuan Yueming, Xiong Wei, Fang Yonghua *et al.*. Detection of oil spills on water by differential polarization FTIR spectrometry [J]. *Spectroscopy and Spectral Analysis*, 2010, **30** (8): 2129~2132
袁越明,熊伟,方勇华等.差分偏振 FTIR 光谱法探测水面溢油污染[J]. *光谱学与光谱分析*, 2010, **30**(8): 2129~2132
- 2 Yuan Yueming, Xiong Wei, Fang Yonghua *et al.*. Modeling analysis of detection of oil spills on water by differential polarization FTIR spectrometry [J]. *Infrared and Laser Engineering*, 2011, **40**(3): 408~412
袁越明,熊伟,方勇华等.差分偏振 FTIR 光谱法探测水面溢油污染的模型分析[J]. *红外与激光工程*, 2011, **40**(3): 408~412
- 3 Zhang Jianqi. *Infrared Physics*[M]. Xi'an: Xidian University Publishing House, 2004
张建奇. *红外物理*[M]. 西安:西安电子科技大学出版社, 2004
- 4 W. C. Shih, A. B. Andrews. Modeling of thickness dependent infrared radiance contrast of native and crude oil covered[J]. *Opt. Express*, 2008, **16**(14): 10535~10542
- 5 A. S. Joseph. Degree of linear polarization in spectral radiances from water-viewing infrared radiometers[J]. *Appl. Opt.*, 1999, **38**(15): 3157~3165
- 6 I. T. Jolliffe. *Principal Component Analysis*[M]. New York: Springer-Verlag, 1986
- 7 Gao Jun. *The Theory of Neural Network and Its Simulation*[M]. Beijing: China Machine Press, 2003
高隽. *神经网络原理及仿真实例*[M]. 北京:机械工业出版社, 2003

版社，2003

- 8 Yang Fusheng, Hong Bo. Independent Component Analysis and Application[M]. Beijing: Tsinghua University Press, 2006
杨福生, 洪波. 独立分量分析的原理与应用[M]. 北京: 清华大学出版社, 2006
- 9 A. Hyvarinen, E. Oja. A fast fixed-point algorithm for independent component analysis[J]. *Neural Comput.*, 1997, **9**(7): 1483~1492
- 10 A. Sharma, K. K. Paliwal. Fast principal component analysis using fixed point algorithm[J]. *Pattern Recognition Lett.*, 2007, **28**(10): 1151~1155
- 11 Lan Tiange, Xiong Wei, Fang Yonghua *et al.*. Research on preprocessing algorithm for infrared spectral signals of biological aerosols[J]. *Acta Optica Sinica*, 2010, **30**(9): 2742~2747
兰天鸽, 熊伟, 方勇华等. 生物气溶胶红外光谱信号预处理算法研究[J]. *光学学报*, 2010, **30**(9): 2742~2747
- 12 T. Hasegawa. Detection of minute chemical signals by principal component analysis[J]. *Trends in Analytical Chemistry*, 2001, **20**(2): 53~64

栏目编辑:何卓铭