

文章编号 1004-924X(2017)04-0884-07

基于 PCA 的时间分辨油荧光光谱分析及优化

李 杰^{1,2}, 李晓龙^{3*}, 唐秋华², 赵朝方¹, 王炯炯¹

(1. 中国海洋大学 信息科学与工程学院, 山东 青岛 266100;

2. 国家海洋局 第一海洋研究所, 山东 青岛 266061;

3. 中国科学院 海洋研究所, 山东 青岛 266071)

摘要: 激光诱导荧光技术可广泛应用于油污染的监测中, 然而普通的油荧光光谱技术只能实现油污染监测的粗分类, 无法区分原油与燃料油的荧光特征。本文基于主成分分析方法(PCA)的时间分辨油荧光分类方法, 实验测量了 20 种油样本的时间分辨荧光光谱特征, 给出了对应的荧光寿命和时间分辨油荧光光谱的时序特征。在此基础上, 利用前三个主成分构成的三维特征矢量空间, 通过分析不同采集时刻下油样本矢量间相关距离的变化, 对油样本的时间分辨荧光光谱进行聚类分析。为了体现油荧光变化的时序性, 引入矢量距离的离散度参量, 提出基于 PCA 进行时间分辨油荧光光谱分析的优化方法。实验结果表明, 基于时间分辨油荧光光谱识别可实现原油与燃料油的光谱时序特征区分, 具备良好的油荧光分类效果。

关键词: 时间分辨荧光光谱; 油荧光分类; 主成分分析(PCA); 荧光寿命

中图分类号: O433.4; TE58 **文献标识码:** A **doi:** 10.3788/OPE.20172504.0884

Spectral analysis and optimization of time-resolved oil fluorescence based on PCA

LI Jie^{1,2}, LI Xiao-long^{3*}, TANG Qiu-hua², ZHAO Chao-fang¹, WANG Jiong-jiong¹

(1. College of Information Science & Engineering, Ocean University of China, Qingdao 266100, China;

2. The First Institute of Oceanography, State Oceanic Administration, Qingdao 266061, China;

3. Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China)

* Corresponding author, E-mail: lixiaolong@qdio.ac.cn

Abstract: Laser-induced Fluorescence (LIF) technique can be widely used in oil pollution monitoring. However, ordinary oil fluorescence spectra can only achieve cursory oil classification, which was disabled to distinguish crude and fuel oils. Herein, time-resolved fluorescence spectra classification method based on Principle Component Analysis (PCA) was investigated and employed to analyze the spectral features of 20 kinds of oils, of which the fluorescence lifetimes and the spectral timing characteristics were obtained. Then referring to fluorescence lifetimes of oils (less than 10 ns commonly), three-dimensional spectra of samples within this time range were used for obtaining a vector space which was composed of first three principal components and was considered as a three-

收稿日期: 2016-08-22; 修订日期: 2016-09-22.

基金项目: 国家自然科学基金资助项目(No. 61505221); 国家海洋局国际海洋合作与履约项目(No. QY0516014); 中央级公益性科研院所基本科研业务费专项资金资助项目(No. 0215G20)

dimensional coordinate system. In this coordinate system, correlation distances of position vectors at difference delay time of fluorescence acquisition were analyzed for spectral clustering of time-resolved oil fluorescence. To reflect timing characteristics of correlation distances, dispersion parameters were introduced into the PCA optimization method. The experimental result indicates that the method based on time-resolved fluorescence spectroscopy can discriminate between crude oils and fuel oils with a higher recognition rate.

Key words: time-resolved fluorescence spectrum; classification of oil fluorescence; Principal Component Analysis (PCA); fluorescence lifetime

1 引言

石油产品对现代社会的发展和科技进步发挥着重要作用,随着石油污染问题的日趋严重,能否快速、有效地监测油污染成为国内外致力解决的科学问题^[1]。激光诱导荧光技术作为重要的海洋溢油监测手段,已广泛应用于油污染的监测和油种识别中。目前,成熟的海洋激光雷达溢油监测系统主要有美国国家海洋和大气局(National Oceanic and Atmospheric Administration, NOAA)和美国国家航空航天局(NASA)研制的 AOL (Airborne Oceanographic Lidar) 系统^[2],加拿大运输技术发展中心和环境局研制的 SLEAF (Scanning Laser Environmental Airborne Fluoresensor) 系统^[2]等。这些系统均采用普通的二维荧光光谱进行溢油识别,只能进行溢油种类的粗分类,油荧光“异类同谱”现象并未得到有效解决,溢油识别精度也有待提高。相对于二维油荧光光谱,三维油荧光光谱能反映更丰富的油组分信息^[3]。虽然基于激发-发射矩阵(Excitation-Emission Matrix, EEM)的三维荧光光谱可用于油种分类,但无法实现油污染的现场快速鉴别。时间分辨三维油荧光光谱探测结合海洋激光雷达技术,可用于现场油污染的快速监测,不过目前针对整个三维谱的聚类分析手段^[4-5]仍无法获得细致的油荧光分类。

时间分辨的三维油荧光光谱能够反映不同油组分的荧光衰变信息。在荧光寿命范围内,不同采集时刻的油荧光光谱会发生变化,如原油荧光光谱峰位出现的红移现象^[6-7]。为了分析时间分辨油荧光光谱的时序变化特征,本文针对20种

石油产品样本(包括原油、重油、燃料油等),基于主成分分析方法(Principal Component Analysis, PCA)提出了一种时间分辨三维油荧光光谱的优化识别方法,从而实现了更为准确的油样本种类区分。

2 时间分辨油荧光光谱测量实验

2.1 实验仪器

时间分辨三维油荧光光谱测量实验系统如图1所示,其中,激发光源采用德国 CryLaS GmbH 公司的3倍频 Nd:YAG 激光器,其单脉冲能量 $\geq 70 \mu\text{J}$,脉冲宽度(Full Width at Half Maximum, FWHM)为0.9~1.3 ns,激光能量波动 $< 2\%$,脉冲频率为100 Hz。

光谱仪内置3种闪耀光栅,光栅参数依次为150,600和2400 l/mm。由于ICCD探测像元宽度的限制,3种光栅对应的光谱宽度依次为260,63和13 nm,光谱分辨率分别为0.5,0.124和0.025 nm。本次实验采用低光谱分辨的光栅,测量荧光光谱为350~580 nm,包括激发波长回波和油荧光信号的主要波长。

荧光信号时序采集器包括时序采集控制器、ICCD和计算机三部分。为了精确地控制采样时间,采用外部光电触发方式,在激光发射窗处设置PIN管触发电路。依据触发电路反馈的信号和样本探测距离,时序采集控制器可调节ICCD的初始采集时刻和时间间隔。光电转换采用Princeton Instruments公司的PI-MAX 512RB型ICCD,其最短触发延迟时间为26 ns,Shutter模式下的最小采集间隔为2 ns,光谱灵敏范围为200~900 nm,通道数为512。

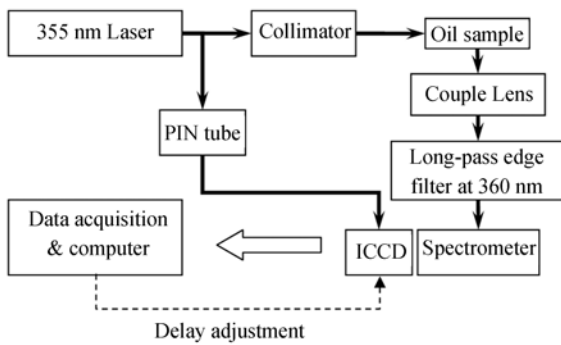


图 1 时间分辨油荧光测量实验系统示意图

Fig. 1 Sketch of experimental system of time-resolved oil fluorescence detection

2.2 油样本及其光谱测量

实验中,为了分析不同油种的时间分辨荧光光谱特征,采用 20 种油样本^[8](如表 1 所示),包含轻质油、原油、重油、燃料油等^[9]。另外,由于油样品的名称较长,为了简化表达,文中采用样品编号进行标注。所有测量样本未进行任何溶剂处理,均是采用石英玻璃制备的 0.2 mm 厚的油膜样本,保持原有的物理和化学特性。

表 1 油样本及对应样本序号^[8]Tab. 1 Oils and serial numbers of samples^[8]

样品编号	样品名称	样品编号	样品名称
1#	马拉拉哥原油	11#	俄罗斯 2 号原油
2#	新加坡燃料油	12#	大庆乙烯原油
3#	委内瑞拉原油	13#	1C 原油
4#	安哥拉原油	14#	4F 燃料油
5#	中原油田原油	15#	克拉玛依原油
6#	3 号成品油	16#	韩国仁川燃料油
7#	韩国重油	17#	大连润滑油
8#	韩国加农柴油	18#	0 号柴油
9#	韩国柴油	19#	93 号汽油
10#	阿拉伯原油	20#	97 号汽油

实验采用短脉宽激光(355 nm 激光脉宽在 1 ns 左右)诱导油荧光,和 ICCD(时间分辨率为 2 ns)测量荧光光谱的时域变化,分析时间分辨油荧光的光谱特征,并实现对油样本时序荧光指纹的识别。图 2 为 1# 油样(马拉拉哥原油)的时间分辨三维荧光光谱,光谱数据用于分析不同采集时刻油荧光的荧光寿命和三维光谱的变化。

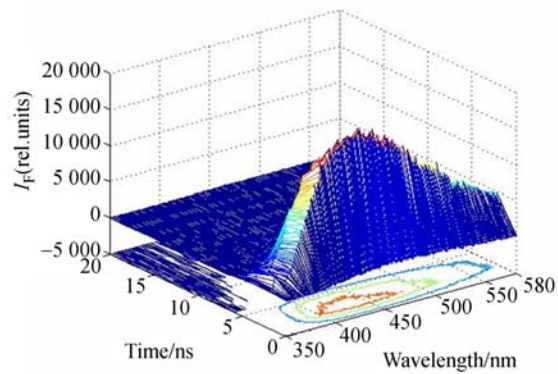


图 2 1# 油样本(N1 原油)的时间分辨荧光光谱

Fig. 2 Time-resolved fluorescence spectrum of 1# oil (N1 crude oil)

3 实验数据分析与讨论

3.1 数据预处理

针对典型的油荧光波长,根据不同时刻荧光强度的衰减比例进行荧光寿命估计。为了区分不同波长处油荧光寿命的差异,首先对所测时间分辨荧光光谱进行去除背景噪声和激光能量归一化处理,以消除单样本测量时激光强度差异的影响。原油及其它提炼油的荧光寿命一般小于 10 ns^[10],这里对 10 ns 内采集的油荧光进行处理,得到油荧光等高线谱,如图 3 所示。

由图 3 可以直观地看到,原油样本的荧光峰较宽,覆盖 390~500 nm 波段;汽油、柴油、润滑油等成品油样本的荧光峰较窄,多集中在 380~450 nm 波段。另外,原油样本间的荧光时序变化特征也存在差异。如 1# 和 5# 原油的荧光在采集时刻 2 ns 时最强,之后荧光强度减弱;而 4# 和 12# 原油的荧光峰则出现在 4 ns 处。

3.2 油荧光寿命

以油荧光的典型峰值波长 450 nm 为例,不同油种样本的荧光变化存在差异,荧光强度变化如图 4 所示。

通过油荧光的归一化数据(图 3 和图 4)判断实验所采用 20 种油样本的荧光寿命特征,结果如下:

(1) 1#、2#、3#、7#、14#、15#、16# 样本油的荧光寿命小于 4 ns;

(2) 4#、5#、10#、11#、12#、13# 样本油的荧光寿命约为 4 ns;

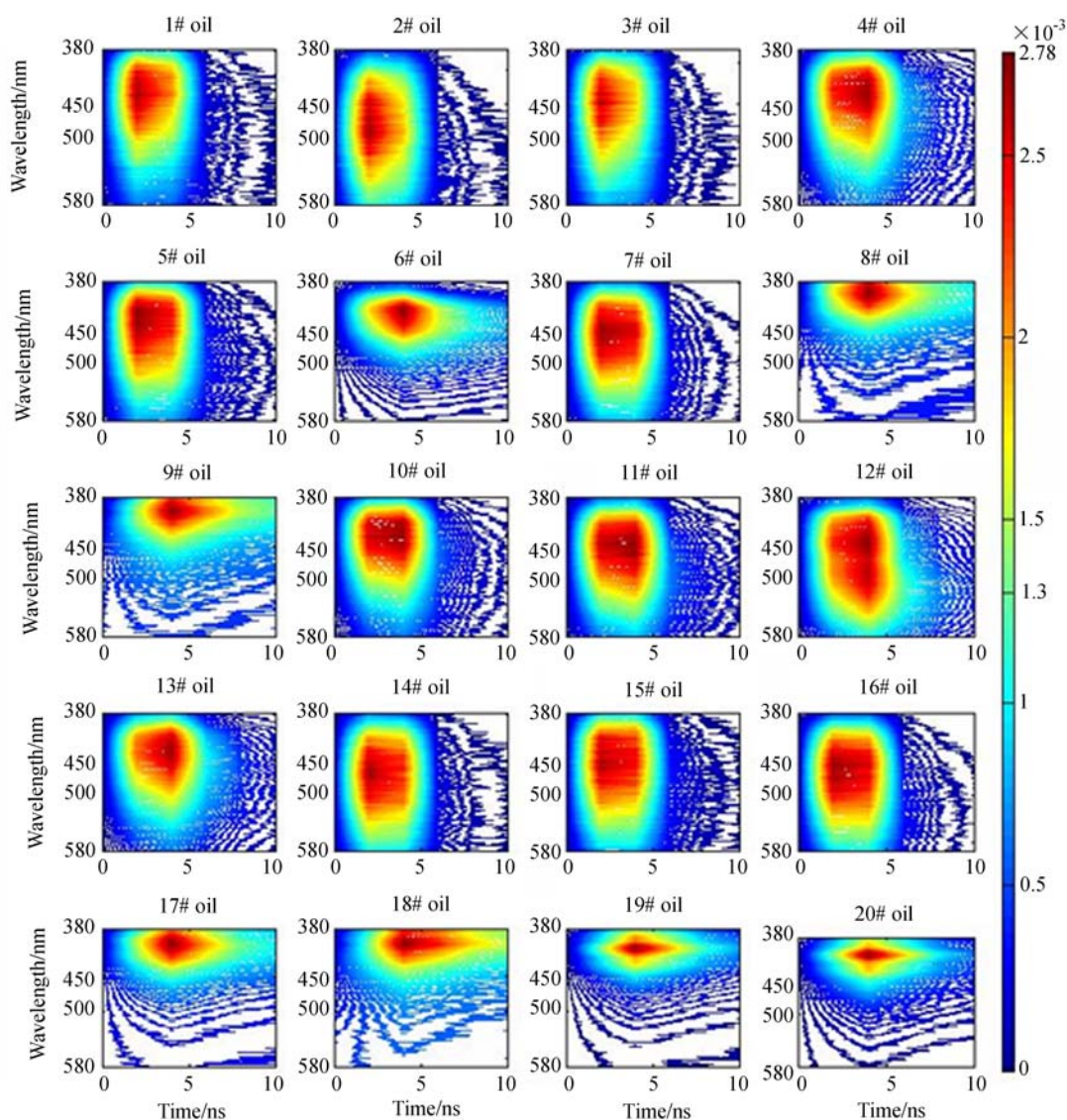


图 3 油样本时间分辨荧光光谱的等高线谱

Fig. 3 Contour spectra of time-resolved oil fluorescence of 20 samples

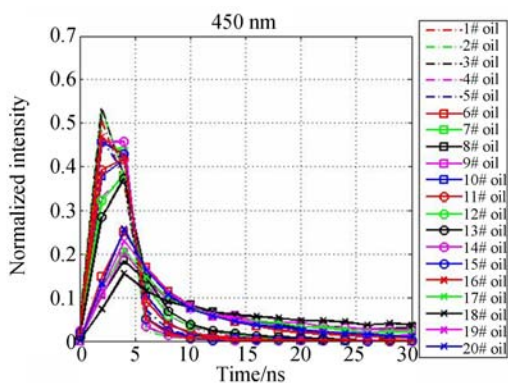


图 4 波长 450 nm 处的油荧光寿命

Fig. 4 Lifetime of oil fluorescence at 450 nm

(3) 6#、8#、9#、17#、19#、20# 样本油的荧光寿命大于 5 ns, 小于 8 ns;

(4) 18# 样本油的荧光寿命大于 8 ns。

依据油样本的种类可得:原油、重油和燃料油的荧光寿命 ≤ 4 ns; 成品油的荧光寿命 > 5 ns, 且 < 8 ns; 0 号柴油的荧光寿命 > 8 ns, 该结论与国外的研究结果^[9]一致。

3.3 时间分辨油荧光光谱的时序特征

由以上油荧光寿命特征可知, 8 ns 的采集时间内不同种类油样本的荧光强度随时间的变化趋势有明显差异, 这里利用 2, 4, 6 和 8 ns 时刻采集的油荧光光谱分析样本时间分辨光谱的时序变化特征。首先, 针对这 4 个时刻采集的样本荧光光

谱进行主成分分析(Principal Component Analysis, PCA),分别选取光谱主成分的累积方差贡献大于 90%的前 3 个主成分,构建矢量空间分布,如图 5~图 8 所示。

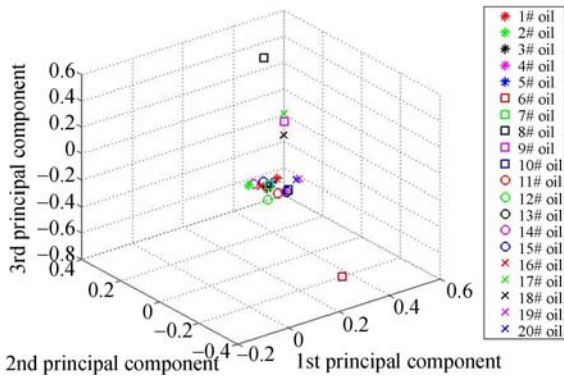


图 5 采集时刻为 2 ns 时油荧光光谱前三个主成分的空间分布

Fig. 5 Relationship between the first three principal components of fluorescence spectra at 2 ns

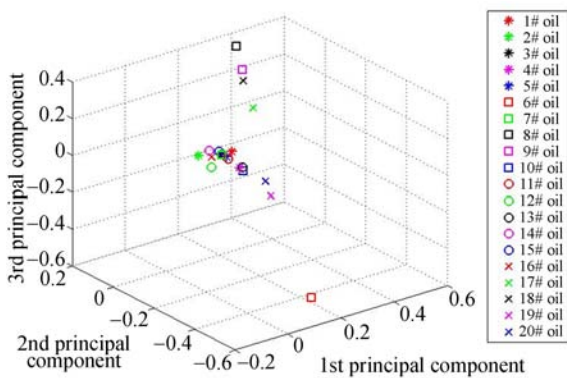


图 6 采集时刻为 4 ns 时油荧光光谱前三个主成分的空间分布

Fig. 6 Relationship between the first three principal components of fluorescence spectra at 4 ns

对比图 5~图 8 可以发现,不同时刻的矢量空间中,样本间的相关距离发生不同程度的变化。以 8#和 9#样本为例,如图 5 所示,2 ns 时刻油样本间的相关距离较大,而之后的采集时刻中样本间的相关距离逐渐减小。如图 5 和图 6 所示,2 ns 和 4 ns 时刻的样本荧光光谱的主成分分布基本一致,6#和 8#油样本相对于其他样本的相对距离较大,即荧光光谱的前三个主成分差异较大,9#、17#和 18#油样本的前三个主成分相似。相对于采集时刻 2 ns,在 4 ns 时 19#和 20#样本的主成分分布发生变化,但与剩余油样本的相

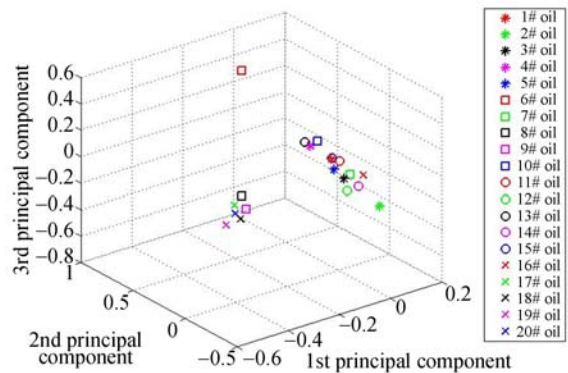


图 7 采集时刻为 6 ns 时油荧光光谱前三个主成分的空间分布

Fig. 7 Relationship between the first three principal components of fluorescence spectra at 6 ns

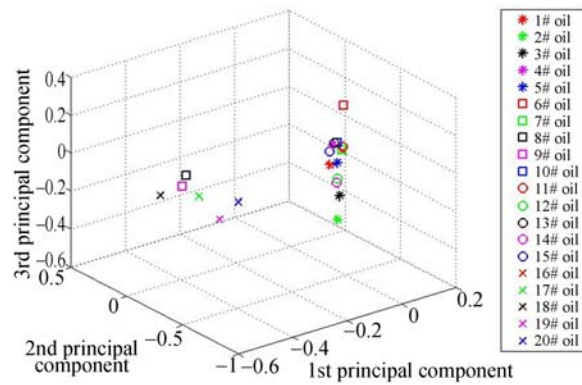


图 8 采集时刻为 8 ns 时油荧光光谱前三个主成分的空间分布

Fig. 8 Relationship between the first three principal components of fluorescence spectra at 8 ns

对距离较小。而由图 7 和图 8 可以看到,19#和 20#样本的主成分分布更接近 8#、9#、17#和 18#样本。因此,油样本相关距离随时间的变化特征可作为样本区分的依据。

4 时间分辨油荧光光谱的主成分分析

一般来说,油荧光光谱的主成分分析通常针对一组光谱数据进行正交变换来得到一组线性不相关的数据,并选择累积方差贡献率较大的几个主成分表征油荧光光谱的变化差异。

本文为了表征油荧光光谱的时序变化特征,基于不同采集时刻下油荧光光谱的主成分特征,采用前三个主成分特征构建矢量空间,提取样本间的相关距离参数^[11-12]作为样本差异的判断指

标。具体步骤如下:

(1)分别对不同采集时刻 380~580 nm 的油荧光光谱进行面积归一化;

(2)通过主成分分析选取前三个主成分特征,将油荧光光谱数据转换为三维特征空间矢量,即确定低维空间中各油样本的位置。在所构建的矢量空间中单个样本的位置表示为 (x, y, z) 。在 t 时刻的矢量空间中,计算油样本间的相关距离,即:

$$l_{ij}(t) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}, \quad (1)$$

其中 i 和 j 对应样本编号。

(3)将所有时刻下任意两个样本的相关距离 $l_{ij}(t)$ 组成数组:

$$\text{array}[l_{ij}(t)] = [l_{ij}(t_1), l_{ij}(t_2), l_{ij}(t_3), \dots, l_{ij}(t_n)]. \quad (2)$$

当 $i=j$ 时, $l_{ij}(t)=0, l_{ij}(t)=l_{ji}(t)$ 。

(4)计算其最大值 $\max(\text{array}[l_{ij}(t)])$,以及最大值和最小值的差与该向量总和值 $\text{sum}(\text{array}[l_{ij}(t)])$ 的比值,即:

$$\text{ratio} = \frac{\{\max(\text{array}[l_{ij}(t)]) - \min(\text{array}[l_{ij}(t)])\}}{\text{sum}(\text{array}[l_{ij}(t)])}. \quad (3)$$

这里通过式(3)表征时间分辨荧光光谱的特征,以突出油荧光光谱主成分特征的时序变化差异。通过放大该时序差异来实现油样本的聚类分析。

(5)将 $\max(\text{array}[l_{ij}(t)])$ 和 ratio 作为并列判据,分别对油样本进行分层聚类分析,聚类结果如表2所示,利用Matlab中linkage、cophenetic、cluster函数创建聚类来进行数据处理。

表2 依据 ratio 和 $\max(\text{array}[l_{ij}(t)])$ 得到的油样本聚类结果

Tab. 2 Cluster analysis of oils by ratio and $\max(\text{array}[l_{ij}(t)])$

油样本编号	$\max(\text{array}[l_{ij}(t)])$ 聚类	参量 ratio 聚类	油样本编号	$\max(\text{array}[l_{ij}(t)])$ 聚类	参量 ratio 聚类
1#	4	16	11#	5	3
2#	5	16	12#	5	17
3#	5	6	13#	1	9
4#	1	5	14#	5	7
5#	5	8	15#	5	10
6#	2	17	16#	5	1
7#	5	2	17#	3	12
8#	3	13	18#	3	14
9#	3	15	19#	3	10
10#	4	4	20#	3	11

表2表明,参量 $\max(\text{array}[l_{ij}(t)])$ 作为聚类依据,可以将油样本识别为5类,这与通过荧光寿命的判断结果一致,如:第1类中4#、13#样本的荧光寿命约为4 ns;第2类中6#样本,以及第3类中8#、9#、17#、19#、20#样本的荧光寿命大于5 ns,且小于8 ns;第5类中2#、3#、7#、14#、15#、16#样本的荧光寿命小于4 ns,11#、12#样本的荧光寿命约为4 ns。由此说明参量 $\max(\text{array}[l_{ij}(t)])$ 作为油荧光时序变化的最值,与荧光寿命判断原理相符合。

采用另一突出光谱时序变化差异的参量 ratio 作为聚类依据,可对油样本进行更为细化的分类,通过表征油荧光光谱的时序光谱变化差异,20种样本聚类为17类,尤其是对原油和燃料油样本之间的识别效果远优于普通光谱识别方法。因为在原油和燃料油两大类中,油样本本身的物理特性

差异(包括组分、透明度和黏度差异等)也会影响时间分辨荧光光谱的特征,部分原油或燃料油样本间仍可以通过时间分辨荧光光谱的差异进行识别。

5 结 论

时间分辨的三维荧光光谱反映了油样本荧光的时序衰变信息,本文基于PCA提出优化方法,对荧光寿命范围内的油样本荧光光谱进行区分。为了进一步分析油荧光光谱变化的时序性,基于PCA方法构建特征矢量空间,引入矢量距离的离散度参量 $\max(\text{array}[l_{ij}(t)])$ 和 ratio ,以形成针对时间分辨油荧光光谱分析的优化方法。实验结果表明,针对时间分辨的油荧光光谱分类,该方法可用于原油与燃料油样本的光谱区分,解决应用激光诱导技术进行这两类油识别时出现的“异类同

谱”现象,具有较好的油样本区分效果。该方法所引入的参量 $\max(\text{array}[l_{ij}(t)])$ 作为聚类依据,与依据荧光寿命进行油样本区分的结果相吻合,而根据参量进行样本聚类,可更为准确地区分油

样本。下一步的研究工作将围绕该优化方法的可靠性和实用性展开,继续增加样本数量进行油种类区分,以进一步验证参量 $\max(\text{array}[l_{ij}(t)])$ 和 ratio 作为依据的聚类方法。

参考文献:

- [1] JHA M N, LEVY J, GAO Y. Advances in remote sensing for oil spill disaster management: state-of-the-art sensors technology for oil spill surveillance [J]. *Sensors*, 2008, 8(1): 236-255.
- [2] BROWN C E, FINGAS M F. Review of the development of laser fluorosensors for oil spill application [J]. *Marine Pollution Bulletin*, 2003, 47(9-12): 477-484.
- [3] BASZANOWSKA E, OTREMBA Z, ROHDE P, et al.. Adoption of the time resolved fluorescence to oil type identification[J]. *Journal of Kones Power-train and Transport*, 2011, 18(2): 25-29.
- [4] UTKIN A B, LAVROV A, VILAR R. Evaluation of water pollution by LIF LIDAR[C]. *5th Workshop on Remote Sensing of the Coastal Zone*, 2011: 104-113.
- [5] FINGAS M. *Handbook of Oil Spill Science and Technology*[M]. New Jersey: John Wiley & Sons, Inc., 2015: 53-73.
- [6] 刘德庆, 栾晓宁, 韩晓爽, 等. 原油样品激光诱导荧光的时间分辨光谱特性研究[J]. *光谱学与光谱分析*, 2015, 35(6): 1582-1586.
- LIU D Q, LUAN X N, HAN X SH, et al.. Characterization of time-resolved laser-induced fluorescence from crude oil samples[J]. *Spectroscopy and Spectral Analysis*, 2015, 35(6): 1582-1586. (in Chinese)
- [7] 韩晓爽, 刘德庆, 栾晓宁, 等. 基于激光诱导时间分辨荧光的原油识别方法研究[J]. *光谱学与光谱分析*, 2016, 36(2): 445-448.
- HAN X SH, LIU D Q, LUAN X N, et al.. Discrimination of crude oil samples using laser-induced time-resolved fluorescence spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2016, 36(2): 445-448. (in Chinese)
- [8] 李晓龙. 视场可调节海洋激光雷达实验系统研制与 ICCD 激光荧光实验研究[D]. 青岛: 中国海洋大学, 2013.
- LI X L. *The Experiment System of Ocean LIDAR with Variable Field-of-view and Laser Induced Fluorescence Study by ICCD*[D]. Qingdao: Ocean University of China, 2013. (in Chinese)
- [9] ALMHDI B. *Oil Classification with Fluorescence Spectroscopy*[D]. Master of Engineering, Engineering Physics, 2005.
- [10] PATSAYEVA S, YUZHAKOV V, VARLAMOV V, et al.. Laser spectroscopy of mineral oils on the water surface[C]. *Proceedings of EAR-SeL-SIG-Workshop LIDAR, Dresden/FRG*, 2000: 106-114.
- [11] 何元磊, 刘代志, 王静荔, 等. 利用独立成分分析的高光谱图像波段选择方法[J]. *红外与激光工程*, 2012, 41(3): 818-824.
- HE Y L, LIU D ZH, WANG J L, et al.. Independent component analysis-based band selection for hyperspectral imagery[J]. *Infrared and Laser Engineering*, 2012, 41(3): 818-824. (in Chinese)
- [12] ROSEN K, GROSSMAN J. *Student's Solutions Guide for Discrete Mathematics and its Applications*[M]. 7th ed. New York: McGraw-Hill Education, 2011: 121-125.

作者简介:



李杰(1985—),男,山东莒县人,工程师,博士研究生,2009年于国家海洋局第一海洋研究所获得硕士学位,主要从事海洋测绘、激光雷达的研究。E-mail: lijie@fio.org.cn

通讯作者:



李晓龙(1985—),男,山东临沂人,工程师,2013年于中国海洋大学获得博士学位,主要从事激光遥感、海洋探测技术方面的研究。E-mail: lixiaolong@qdio.ac.cn