

荧光光谱结合主成分分析对赤潮藻的识别测定

苏荣国, 胡序朋, 张传松, 王修林

(中国海洋大学海洋化学理论与工程技术教育部重点实验室, 青岛 266003)

摘要: 利用主成分分析和非负最小二乘法, 建立了我国近海常见的 11 种赤潮藻在属水平上的活体荧光识别测定技术. 利用 Delaunay 三维插值法, 消除浮游藻三维荧光光谱的瑞利散射和拉曼散射峰, 根据 Fisher 判别分析方法对各主成分及其组合的分辨能力的判别结果, 确定赤潮藻三维荧光光谱的第一主成分得分谱和第二主成分得分谱作为赤潮藻识别特征谱, 建立赤潮藻荧光特征谱库, 在此基础上, 利用非负最小二乘法对 11 种赤潮藻在属水平上进行识别测定, 赤潮藻的识别正确率 $\geq 85\%$, 其中 9 种赤潮藻的识别正确率 $\geq 90\%$, 特别是对于近几年东海发生赤潮频率极高的东海原甲藻 (*Prorocentrum donghaiense*)、中肋骨条藻 (*Skeletonema costatum*) 和裸甲藻 (*Gymnodinium* sp.), 识别正确率 $\geq 95\%$. 研究表明, 利用主成分分析提取赤潮藻三维荧光光谱特征以实现赤潮藻在属水平上的识别测定是可行的.

关键词: 荧光光谱; 主成分分析; 赤潮藻; 识别

中图分类号: X55 文献标识码: A 文章编号: 0250-3301(2007)07-1529-05

Discrimination of Red Tide Algae by Fluorescence Spectra and Principle Component Analysis

SU Rong-guo, HU Xu-peng, ZHANG Chuan-song, WANG Xiu-lin

(Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, Ocean University of China, Qingdao 266003, China)

Abstract: Fluorescence discrimination technology for 11 species of the Red Tide algae at genus level was constructed by principle component analysis and non-negative least squares. Rayleigh and Raman scattering peaks of 3D fluorescence spectra were eliminated by Delaunay triangulation method. According to the results of Fisher linear discrimination, the first principle component score and the second component score of 3D fluorescence spectra were chosen as discriminant feature and the feature base was established. The 11 algae species were tested, and more than 85% samples were accurately determined, especially for *Prorocentrum donghaiense*, *Skeletonema costatum*, *Gymnodinium* sp., which have frequently brought Red tide in the East China Sea. More than 95% samples were right discriminated. The results showed that the genus discriminant feature of 3D fluorescence spectra of Red Tide algae given by principle component analysis could work well.

Key words: fluorescence spectra; principle component analysis; Red Tide algae; discrimination

赤潮已经成为我国和世界的一大海洋灾害, 引起了各国政府及科学界的高度重视^[1]. 准确、实时地对发生赤潮的浮游藻进行识别测定, 对于赤潮机理研究和赤潮灾害防治都具有重要意义. 常用的赤潮藻种类鉴别方法是图象识别技术^[2,3]、色素分析技术^[4,5]、荧光技术^[6,7]等. 荧光技术因具有灵敏度高、易于实现现场实时监测及对物种具有良好的鉴别性等优点而被广泛关注. Lee 等^[8]根据蓝藻所具有的藻蛋白发出的特征荧光光谱, 建立了现场监测蓝藻含量的荧光分析技术; Kolbowski 等^[9]使用发光二极管, 通过初始荧光产率 (F_0) 识别测定了 3 个主要的浮游藻类群; Boddy 等^[10]使用流式细胞仪, 基于浮游藻的荧光性质和散射光性质的差异, 借助神经网络技术实现了对 72 种浮游藻的识别测定. 目前物质的光谱鉴别技术主要是通过化学计量学方法提取光谱特征来实现的, 其中主成分分析是常用的技术之一^[11-14].

由于发生赤潮时, 基本上都是 1 种藻占绝对优势, 这样发生赤潮时对浮游植物群落组成的测定就转化为对引发赤潮的藻的识别测定. 本研究基于赤潮藻三维荧光光谱, 使用主成分分析和非负最小二乘法对 11 种我国近海常见的赤潮藻种进行分析, 旨在建立赤潮藻在属水平上的实时、快速的荧光识别测定技术.

1 材料与方法

1.1 仪器与分析条件

日本日立公司 F4500 荧光分光光度计及其配套的工作站软件, 1 cm 石英比色池.

仪器分析条件: 激发和发射狭缝为 5 nm, 扫描速度为 12 000 nm/min; 激发波长的扫描范围为 400 ~

收稿日期: 2006-08-25; 修订日期: 2007-01-15

基金项目: 国家高技术研究发展计划(863)项目(2004AA639620)

作者简介: 苏荣国(1973 ~), 男, 博士, 讲师, 主要研究方向为海洋监测技术, E-mail: surongguo@mail.ouc.edu.cn

600 nm,扫描间隔为 5 nm;发射波长的扫描范围为 500~800 nm,扫描间隔为 5 nm;每个样平行测定 3 次,取平均值作为该样品的荧光光谱.

1.2 藻种培养

选用的 11 种东海常见赤潮藻(表 1)均来自中国海洋大学海洋污染生态化学实验室,藻种的培养实验均在室内进行,培养容器经 1:5 的 HCl 浸泡处理后,再用蒸馏水和二次蒸馏水冲洗,洗净后蒸汽灭菌(121℃, 0.56 kg•cm⁻², 30 min)并烘干后待用.海水经孔径为 0.45μm 的滤膜过滤后,高压锅煮沸灭

菌,静置平衡 24 h,按 f/2 配方配制营养液,作为培养基质(硅藻门以外的藻种均采用不加硅的 f/2 营养盐配方).实验藻种于光照强度分别为 10 000 lx 和 15 000 lx(400~700 nm)的光照条件下培养,明暗周期 12 h:12 h,温度为(20±1)℃.培养期 15 d,每 3 d 取 1 次藻液进行分析测定,各光照条件下每个藻种平行培养 3 份,其中每个光照取 1 份样品所得荧光光谱作为训练集,用于建立识别方法,其余样品的荧光光谱作为测试集,用于所建方法的验证.

1.3 浮游藻三维荧光光谱的主成分分析

表 1 所选用的浮游藻种类
Table 1 Species of phytoplankton for the experiments

门类	属	藻种	缩写代号
甲藻门 Pyrrophyta	原甲藻属 <i>Prorocentrum</i>	东海原甲藻 <i>Prorocentrum donghaiense</i>	pr
		微小原甲藻 <i>Prorocentrum minimum</i>	pm
	裸甲藻属 <i>Gymnodinium</i>	裸甲藻 <i>Gymnodinium</i> sp.	gy
	亚历山大藻属 <i>Alexandrium</i>	塔玛亚历山大藻 <i>Alexandrium tamarense</i>	al
	斯克里普藻属 <i>Scripsiella</i>	锥状斯克里普藻 <i>Scripsiella trochoidea</i>	sc
硅藻门 Bacillariophyta	骨条藻属 <i>Skeletonema</i>	中肋骨条藻 <i>Skeletonema costatum</i>	sk
	角毛藻属 <i>Chaetoceros</i>	旋链角毛藻 <i>Chaetoceros curvisetus</i>	cu
		柔弱角毛藻 <i>Chaetoceros delicatula</i>	de
	拟菱形藻属 <i>Pseudonitzschia</i>	尖刺拟菱形藻 <i>Pseudo-nitzschia pungens</i>	ps
黄藻门 Xanthophyta	卡盾藻属 <i>Chattonella</i>	海洋卡盾藻 <i>Chattonella marina</i>	cm
金藻门 Chrysophyta	棕囊藻属 <i>Phaeocystis</i>	球形棕囊藻 <i>Phaeocystis globosa</i>	cg

根据 Delaunay 三角插值法^[15],消除浮游藻三维荧光光谱的瑞利散射和拉曼散射;然后将各浮游藻三维荧光光谱归一化(以发射波长为 680 nm 的激发光谱的平均荧光强度为基准);对各浮游藻的三维荧光光谱进行主成分分析,得到各主成分得分谱;利用 Fisher 判别分析方法对各主成分得分谱及其组合的分辨能力进行判别,选取分辨能力最佳的主成分得分谱或其组合作为赤潮藻的特征谱.以上工作利用 Matlab 6.5 完成.

1.4 浮游藻识别方法的建立

利用选取的主成分得分谱或其组合,基于非负最小二乘法建立我国东海常见的 11 种赤潮藻的识别测定方法,并在实验室进行方法有效性验证.以上工作利用 Matlab 6.5 完成.

2 结果与讨论

2.1 浮游藻三维荧光光谱的主成分分析

水的散射峰(包括瑞利散射和拉曼散射)的存在显著地降低了三维荧光光谱的定性和定量测定能力,特别是对于低含量样品的分析干扰尤为严重.本研究利用 Delaunay 三角插值法消除散射的影响,即

将散射峰切除,以散射峰邻近区域保留的数据为基准进行三维插补,这种消除散射的方法要优于传统的差减法.对训练集的三维荧光光谱进行归一化处理,以消除荧光光强的影响.对归一化后的三维荧光光谱做主成分分析.主成分分析方法是寻求从高维空间到低维空间映射的方法,以便于用几个较少的指标来综合所研究目标总体各方面的信息,且这几个指标所代表的信息不重叠.主成分分析所得的载荷矩阵是投影的线性空间,而得分矩阵是原始荧光光谱在线性空间的投影,因此笔者使用主成分得分谱作为特征谱.由于前 3 个主成分的累积贡献已达 85%以上,故仅用前 3 个主成分就可表示原荧光光谱的主要信息.通过对 11 种藻 3 个主成分得分谱图的比对分析,发现前 2 个主成分得分谱具有一定的分类识别能力(图 1,甲藻为实线,硅藻为短虚线,黄藻和金藻为长虚线),并且特征具有较好的稳定性,而第 3 个主成分得分谱虽然也具有一定的区分识别能力,即不同藻种的第三主成分得分谱具有差异性,但是同 1 种藻在不同条件下测得的谱其第三主成分得分谱也具有显著差异,并且这种差异往往超过了种间差异,故第三主成分得分谱已不宜作为识别特

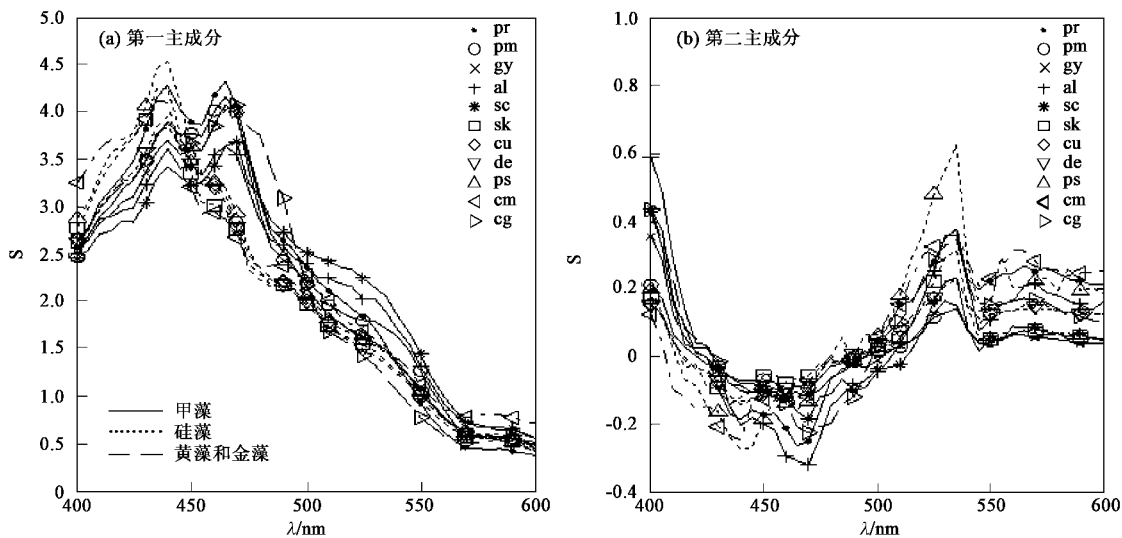


图 1 赤潮藻三维荧光光谱主成分得分谱图

Fig.1 Principle component score of 3D fluorescence spectra of the Red Tide algae

征谱.本研究将在第一和第二主成分得分谱中选择识别特征谱.

2.2 浮游藻三维荧光光谱各主成分得分谱的 Fisher 判别分析

Fisher 线性判别分析是 1 个快速的统计判别技术,它寻找由原变量线性组合成的新变量,这些新变量能够尽可能地将目标从其他类别中分离出来.通过一定的运算规则产生 n 个用于判别的标量,第 1

个判别量是最大判别标量,第 2 个判别量次之,并且第 3 个判别量与第 1 个判别量正交,即不相关,依次类推.本研究利用第 1(n_1)和第 2 判别量(n_2)对特征谱进行判别分析,以确定最佳的赤潮藻属水平上的识别特征谱.从图 2 可以看出,第一主成分得分谱无法分辨中肋骨条藻 (*Skeletonema costatum*)和尖刺拟菱形藻 (*Pseudo-nitzschia pungens*), 对角毛藻 (*Chaetoceros*)、微小原甲藻 (*Prorocentrum minimum*)和

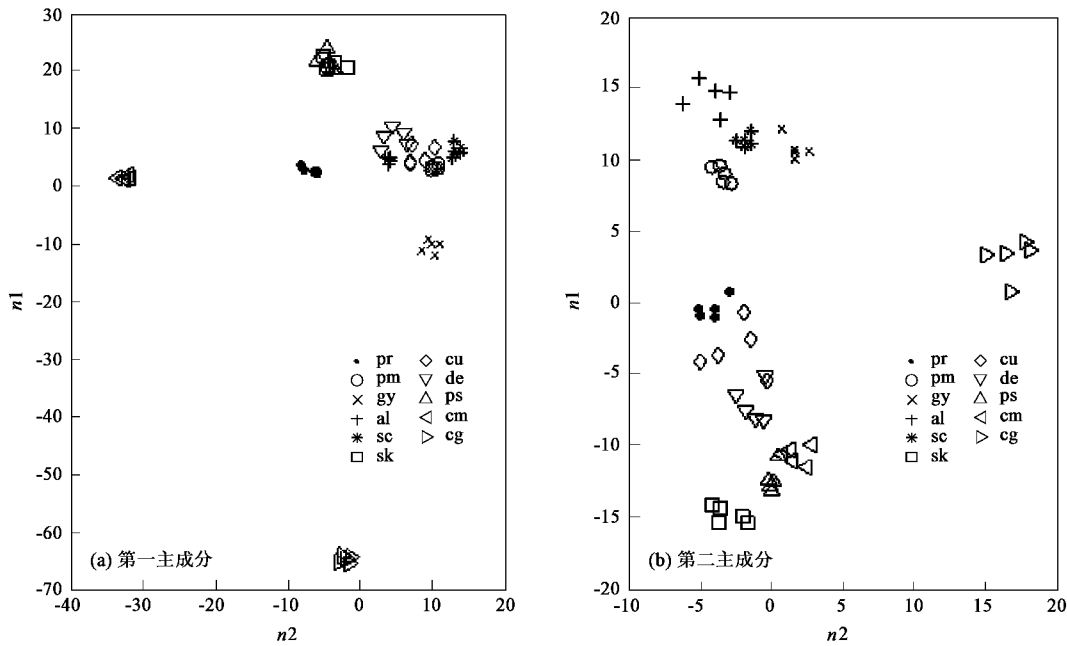


图 2 赤潮藻荧光光谱主成分得分谱 Fisher 判别分析

Fig.2 Fisher discrimination of the principle component score of the fluorescence spectra of Red Tide algae

塔玛亚历山大藻(*Alexandrium tamarense*)的分辨能力较弱,而第二主成分得分谱恰好能够弥补;第二主成分得分谱则无法分辨尖刺拟菱形藻(*Pseudo-nitzschia pungens*)和海洋卡盾藻(*Chattonella marina*),对角毛藻(*Chaetoceros*)和东海原甲藻(*Prorocentrum donghaiense*)的分辨能力有限,而第一主成分得分谱则能够很好地分辨这几种赤潮藻.据此,本研究使用第一主成分得分谱和第二主成分得分谱作为赤潮藻属水平上的识别特征.图3为赤潮藻识别特征谱的Fisher判别分析.就实验所用11种我国东海常见赤潮藻而言,所选特征谱具有稳定的可分辨差异.对比各藻不同生长期的特征谱,发现各生长期特征谱谱形基本相似,故将各藻不同生长期特征谱合并为1条谱.将各藻在不同光照条件下所得特征谱合并建立浮游藻荧光特征谱库,用于在属水平上对浮游藻的识别分析.

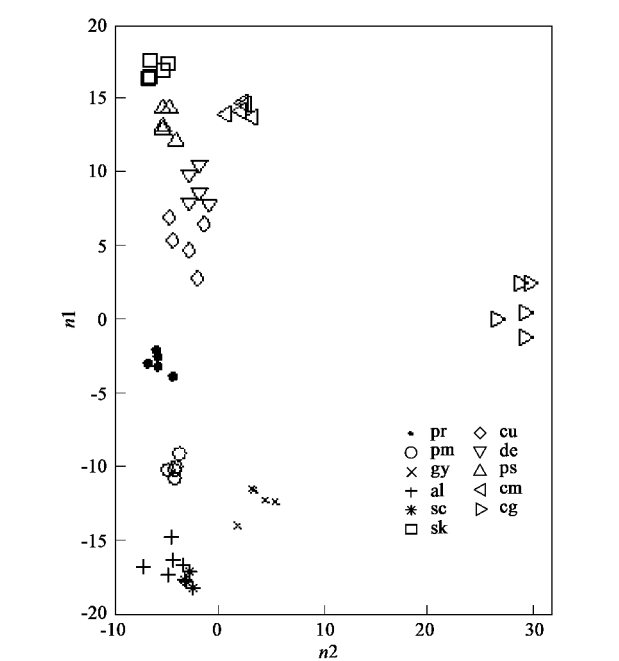


图3 赤潮藻特征谱 Fisher 判别分析

Fig.3 Fisher discrimination of the feature spectra of the Red Tide algae

2.3 赤潮藻的识别分析

通过浮游藻的特征谱识别赤潮藻的基本思想是找到与样品特征谱最为相似的特征谱库里的特征谱,可转化为线性回归问题,即以标准谱库中的每1条特征谱对样品特征谱进行拟合,误差最小者的特征谱对应的赤潮藻即为识别结果.为使误差最小,一般采用最小二乘法,因未知待估参数不可能是负值,因此在运算中加以非负限制,使结果更加符合实

际情况,即非负最小二乘法解析多元线性回归模型.表2为11种赤潮藻在属水平上的识别测定结果,赤潮藻的识别正确率 $\geq 85\%$,其中9种赤潮藻的识别正确率 $\geq 90\%$,特别是对于近几年东海发生赤潮频率极高的东海原甲藻(*Prorocentrum donghaiense*)、中肋骨条藻(*Skeletonema costatum*)和裸甲藻(*Gymnodinium* sp.),识别正确率 $\geq 95\%$.从识别结果可以看出,该技术利用三维荧光光谱能够较为准确地对11种赤潮藻在属水平上进行识别测定.通过对识别错误的样品进行分析发现,识别错误的浮游藻样品基本都是培养前期(第3 d)的样品,而在这期间,多数藻的细胞密度较低.由于所用荧光分光光度计的灵敏度限制,使得赤潮藻的荧光光谱谱形无法完全显现,这可能是造成识别错误的重要原因.

表2 11种赤潮藻识别测定结果
Table 2 Discrimination of the 11 Red Tide algae

赤潮藻种	光强 /lx	样品数 /个	正确识别数 /个	正确识别率 /%
Pr	10 000	10	9	95
	15 000	10	10	
Pm	10 000	10	8	85
	15 000	10	9	
Gy	10 000	10	9	95
	15 000	10	10	
Al	10 000	10	9	95
	15 000	10	10	
Sc	10 000	10	10	100
	15 000	10	10	
Sk	10 000	10	10	100
	15 000	10	10	
Cu	10 000	10	10	95
	15 000	10	9	
De	10 000	10	10	100
	15 000	10	10	
Ps	10 000	10	8	85
	15 000	10	9	
Cm	10 000	10	8	90
	15 000	10	10	
Cg	10 000	10	10	100
	15 000	10	10	

3 结论

本研究将三维荧光光谱技术与主成分分析方法相结合,实现了对我国东海常见的11种赤潮藻的在属水平上的快速识别测定,识别正确率 $\geq 85\%$,特别

是对于近几年东海发生赤潮频率极高的东海原甲藻(*Prorocentrum donghaiense*)、中肋骨条藻(*Skeletonema costatum*)和裸甲藻(*Gymnodinium* sp.), 识别正确率 $\geq 95\%$. 本实验中, 由于赤潮藻培养前期(第 3 d)细胞密度较低, 使得赤潮藻的荧光特征无法完全显现, 这可能是造成识别错误的重要原因. 随着仪器技术和光谱信号处理技术的发展, 利用荧光技术识别测定赤潮藻必然更加快捷和准确.

参考文献:

[1] 崔廷伟, 张杰, 马毅, 等. 赤潮光谱特征及其形成机制[J]. 光谱学与光谱分析, 2006, **26**: 884 ~ 886.

[2] Culverhouse P F, Simpson R G, Ellis R, *et al.* Automatic classification of field-collected dinoflagellates by artificial neural network [J]. Marine Ecology Progress Series, 1996, **139**: 281 ~ 287.

[3] Sieracki C K, Sieracki M E, Yentsch C S. An image-in-flow system for automated analysis of marine microplankton[J]. Marine Ecology Progress Series, 1998, **168**: 285 ~ 296.

[4] Lewitus A J, White D L, Tymowski R G, *et al.* Adapting the CHEMTAX method for assessing phytoplankton taxonomic composition in southeastern U. S. estuaries[J]. Estuaries, 2005, **28**: 160 ~ 172.

[5] Mackey M D, Mackey D J, Higgins H W, *et al.* CHEMTAX-a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton[J]. Marine Ecology Progress Series, 1996, **144**: 265 ~ 283.

[6] Beutler M, Wiltshire K H, Arp M, *et al.* A reduced model of the fluorescence from the cyanobacterial photosynthetic apparatus

designed for the in situ detection of cyanobacteria[J]. Biochimica et Biophysica Acta, 2003, **1604**: 33 ~ 46.

[7] Seppälä J, Balode M. The use of spectral fluorescence methods to detect changes in the phytoplankton community[J]. Hydrobiologia, 1998, **363**: 207 ~ 217.

[8] Lee T, Tsuzuki M, Takeuchi T, *et al.* Quantitative determination of cyanobacteria in mixed phytoplankton assemblages by an in vivo fluorimetric method [J]. Analytical Chimica Acta, 1995, **302**: 81 ~ 87.

[9] Kolbowski J, Schreiber U. Computer-controlled phytoplankton analyzer based on a 4-wavelength PAM Chl fluorometer[A]. In: Photosynthesis: From Light to Biosphere[C]. Dordrecht/Boston/ London: Kluwer Academic Publishers, 1995. 825 ~ 828.

[10] Boddy L, Morris C W, Wilkins M F, *et al.* Identification of 72 phytoplankton species by radial basis function neural network analysis of flow cytometric data[J]. Marine Ecology Progress Series, 2000, **195**: 47 ~ 59.

[11] 褚小立, 袁洪福, 陆婉珍. 光谱结合主成分分析和模糊聚类方法的样品聚类 and 识别[J]. 分析化学, 2000, **28**(4): 421 ~ 427.

[12] 徐永群, 孙素琴, 袁子民, 等. 红外光谱结合主成分分析鉴别道地山药[J]. 分析化学, 2002, **30**(10): 1231 ~ 1233.

[13] 张前前, 类淑河, 王修林, 等. 浮游藻活体三维荧光光谱分类判别方法研究[J]. 光谱学与光谱分析, 2004, **24**(10): 1227 ~ 1229.

[14] Zhang Q Q, Lei S H, Wang X L, *et al.* Discrimination of phytoplankton using characteristic spectra of 3D fluorescence spectra [J]. Spectrochimica Acta A, 2006, **63**: 361 ~ 369.

[15] Zepp R G, Sheldon W M, Moran M A. Dissolved organic fluorophores in southeastern US coastal waters: correction method for eliminating Rayleigh and Raman scattering peaks in excitation-emission matrices[J]. Marine Chemistry, 2004, **89**: 15 ~ 36.