浮游植物群落组成的离散三维荧光光谱识别测定技术

张珊珊1,苏荣国1*,段亚丽1,宋志杰2,王修林1

- 1. 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, 山东 青岛 266100
- 2. 中国海洋大学信息科学与工程学院, 山东 青岛 266100

摘 要 从荧光分光光度计所测浮游植物三维荧光光谱中,选取 12 个激发波长点(400,430,450,460,470,490,500,510,525,550,570,590 nm)组成离散三维荧光光谱,采用 db7 小波技术及 Bayes 判别法对光谱进行特征提取,并利用多元线性回归辅以非负最小二乘建立识别技术,实现了浮游植物群落组成门、属水平上的荧光识别测定。所测样品中,32 种赤潮藻:模拟混合样品(优势种比例分别为 70%,80%,90%)及单种藻,属水平平均识别正确率分别为 67.5%,75.8%,81.4%,79.4%。43 种浮游植物:模拟混合样品(优势门比例分别为 50%,75%,100%)门水平平均识别率分别为 95.2%,99.7%,91.9%;平均定量测定百分比分别为 38.1%,63.2%,90.5%。

关键词 浮游植物群落; 离散三维荧光光谱; 小波分析; 特征提取; 识别中图分类号: O657. 3; Q949. 2 文献标识码: A DOI: 10. 3964/i, issn. 1000-0593(2011)03-0732-05

引言

常用的浮游植物鉴别技术主要有显微镜技术[1]、图像识 别技术[2,3]、色素分析技术[4,5]、分子探针技术[6,7]、荧光分 析技术[8,9]等。其中荧光技术因灵敏度高、易于实现实时现 场检测及对物质具有良好的鉴别性等优点而广受关注。1985 年 Yentsch & Phinney^[8]根据叶绿素与辅助色素的比率利用 荧光光谱将海洋浮游植物分为四类进行分析。1993年, Lee^[10]等使用系列单色发光二极管为光源,根据蓝藻含有的 藻蛋白的特征荧光谱,建立了现场监测活体蓝藻含量的荧光 分析技术。2002年, Beutler[11] 等利用五个激发波长(450, 525, 570, 590, 610 nm)下的叶绿素荧光激发光谱将浮游植 物分为四类[绿藻,蓝藻,隐藻,混合藻(含甲藻和硅藻)], 基于此技术,采用发光二极管作光源,研制了便携式仪器 BBE 藻类分析仪。国内研究者张前前等[12,13] 用三维荧光光 谱辅以主成分分析等方法,对甲藻和硅藻(11种浮游植物)的 荧光特征进行了研究、分类识别,得到了76%的门水平上正 确识别比率。张芳[14,15]利用小波分解手段结合系统聚类建 立了东海主要浮游植物(7门38种)的标准谱库,分别得到了 门、属水平上 96%以上及 86%以上的正确识别比率。刘 宝[16]选择两个正交小波(coif2, symlet7),一个双正交小波

(biorl. 1)基于荧光分光光度计所测连续三维荧光光谱分别 建立了我国近海常见浮游植物群落的门、属水平上的识别分 析技术。

综上所述,小波技术是提取荧光特征的有效手段,本文在实验室已有研究的基础上,受BBE 启发,以研制使用系列单色发光二极管为光源的适于现场应用的浮游植物荧光自动分析仪为目标,根据单色发光二极管的技术现状及保留最佳荧光特征的原则,选取了12个激发波长点组成浮游植物的离散三维荧光光谱,利用小波分析、Bayes 判别法、系统聚类和非负最小二乘法等技术,建立了能够对浮游植物群落组成进行门、属水平上识别测定的荧光技术,为浮游植物荧光自动分析仪的研制提供了理论支撑。

1 实验部分

1.1 仪器与分析条件

Jobin Yvon 公司生产的 Fluorolog3-11 荧光分光光度计及配套的工作站软件, 1 cm 石英比色池。

分析条件:激发波长 $400\sim650$ nm,发射波长 $600\sim750$ nm,步长为 5 nm,狭缝宽度为 5 nm,扫描速度为 80 nm·s⁻¹,扫描信号积分时间为 0.05 s。样品每隔 1 天取样,每份样品平行测定两次,进行三维荧光光谱测量,获得光谱。

收稿日期: 2010-05-12**,修订日期:** 2010-09-27

基金项目: 国家自然科学基金项目(40706036) 和山东省自然科学基金项目(ZR2009EM001)资助

作者简介: 张珊珊, 女, 1987 年生, 中国海洋大学化学化工学院硕士研究生 e-mail: shanshan66395680@sina.com

*通讯联系人 e-mail: surongguo@mail ouc edu cn

1.2 浮游植物在实验室中的培养

选用的 43 种浮游植物(见表 1)均来自中国海洋大学海洋污染生态化学实验室。藻种光照强度分别为 $20\,000$, $12\,000$, $7\,000$, $4\,000\,lux$; 培养温度分别为 $20\,000$, $25\,^{\circ}$; 光暗比为 $12:12\,h$, 培养周期为 $12\,d$ 。每种藻液平行培养两份。

2 数据处理与分析

2.1 离散三维荧光光谱的获得及散射去除

本文对组成离散三维荧光光谱激发波长点的选择原则如下:第一,尽量使分解得到的谱图最大程度地反映原三维荧光光谱的谱形,使特征信息尽可能多的保留;第二,尽量减小非特征区域和特征区域的比例,使得不同门类的藻特征谱差异显著;第三,所选的特征片段线性独立,避免识别时的共线性问题。在满足以上条件的基础上,再结合目前单色发

光二极管的技术水平,选择了最佳波长点,即: 400, 430, 450, 460, 470, 490, 500, 510, 525, 550, 570, 590 nm。

浮游植物三维荧光光谱中散射峰的存在严重影响三维荧光光谱特征的表达,本文选择 Delaunay^[17]三角形插值法来消除三维荧光光谱的瑞利散射和拉曼散射,该技术是以散射峰邻近区域的数据为基准进行三维插值,能够有效消除散射峰的影响并保留散射区域的荧光信号。

2.2 浮游植物荧光特征谱的提取

将去除散射后的三维光谱数据进行二维形式的转化,然后作最大值归一化处理,再进行小波分解。小波分解是一个通过提取、凸显光谱和图像信号特征来对目标进行自动识别的非常有效的工具,本文选择紧支撑的、正交的 db7 小波函数对荧光光谱进行分解,得到相互独立的尺度分量 Ca(n)和小波分量 Cd(n),即小波特征谱(图 1)。

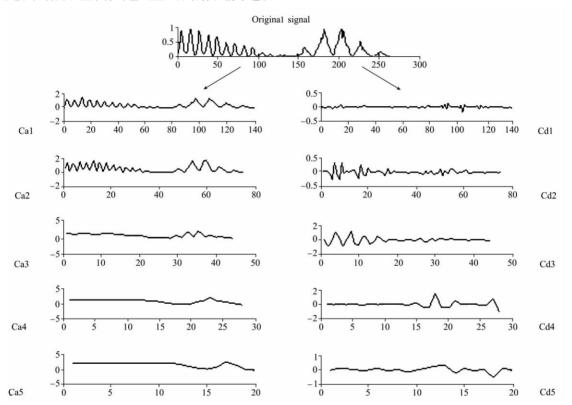


Fig. 1 Decomposition of the 3D discrete fluorescence spectra by db7: Ca1~Ca5 stands for the first to fifth scale vectors, and Cd1~Cd5 stands for the first to fifth wavelet vectors

对以 db7 小波函数分解光谱信号所得的各分量进行观察比对后发现:第一、二层特征谱受噪声影响仍较为明显,不能很好地体现荧光特征;相对而言,第三层之后的特征谱能较好的避免噪声干扰,因此本文中选择了第三~五层尺度分量和小波分量作为备选特征谱,对其进行 Bayes 判别分析,进而决定特征谱的选择。

对 43 种藻的三~五层尺度分量和小波分量分别进行 Bayes 判别分析发现:对于单种藻,判别正确率,Ca3 为:24 种藻大于 80%(其中有 20 个藻>90%),三种藻在 70%~ 80%之间,15 种藻小于 70%;Ca4 为:23 种藻大于 80%(其 中有 18 个藻>90%),五种藻在 70%~80%之间,15 种藻小于 70%;Ca5 为:24 种藻大于 80%(其中有 16 个藻>90%),三种藻在 70%~80%之间,16 种藻小于 70%;Cd3 为:22 种藻大于 80%(其中有 17 个藻>90%),三种藻在 70%~80%之间,18 种藻小于 70%;Cd4 为:22 种藻大于 80%(其中有 16 个藻>90%),三种藻在 70%~80%之间,18 种藻小于 70%;Cd5 为:22 种藻大于 80%(其中有 19 个藻>90%),五种藻在 70%~80%之间,16 种藻小于 70%。

对于模拟混合样品(由优势藻和次优势藻按照不同的比例运用程序模拟混合得到),当优势种比率为70%时,判别

正确率, Ca3 为: 19 种藻大于>80%(其中有 15 个藻>90%), 四种藻在 $70\% \sim 80\%$ 之间, 20 种藻小于 70%; Ca4 为: 18 种藻大于 80%(其中有 11 个藻>90%), 七种藻在 $70\% \sim 80\%$ 之间, 18 种藻小于 70%; Ca5 为: 15 种藻大于 80%(其中有 11 个藻>90%), 六种藻在 $70\% \sim 80\%$ 之间, 22 种藻小于 70%; Cd3 为: 18 种藻大于 80%(其中有 11 个藻>90%), 四种藻在 $70\% \sim 80\%$ 之间, 21 种藻小于 70%; Cd4 为: 16 种藻大于 80%, 两种藻在 $70\% \sim 80\%$ 之间, 其余 25 种藻小于 70%; Cd5 为: 17 种藻大于 80%, 三种藻在 $70\% \sim 80\%$ 之间, 23 种藻小于 70%。

当优势种比率为 80%时,判别正确率,Ca3 为:32 种藻大于 80%(其中有 28 个藻>90%),两种藻在 70% \sim 80%之间,九种藻小于 70%; Ca4 为:32 种藻大于 80%(其中有 26 个藻>90%),三种藻在 70% \sim 80%之间,八种藻小于 70%; Ca5 为:33 种藻大于 80%(其中有 25 个藻>90%),其余十种藻小于 70%; Cd3 为:32 种藻大于 80%(其中有 26 个藻>90%),一种藻在 70% \sim 80%之间,十种藻小于 70%; Cd4 为:25 种藻大于 80%,五种藻在 70% \sim 80%之间,13 种藻小于 70%; Cd5 为:32 种藻大于 80%(其中有 26 个藻>90%),三种藻在 70% \sim 80%之间,八种藻小于 70%。

由此可以看出,不论是单种藻还是混合样品均以第三层尺度分量的判别情况最好。在所判别的 43 个藻中,对单种藻, 仅 Chaetoceros(Ca), Odontella cf_sinensis(Oc), Coscinodiscus sp 312(Cf), Rhizosolenia setigera Brightwell(Rh), Gephyrocapsa oceanic(Ks), Prococentrum marinum(Ma), Gymnodinium sp(Gy) 这些藻 Ca3 的判别情况较差;对于模拟混合样品,当优势种比例为 80% 时,Ca3的 判别情况 较差的有 Ditylum brightwellii(Db),

Thalassiosira rotula (Tr), Ks, Gy, Synechococcus sp(Sy); 当优势种比例为 70% 时,由于优势种比例降低, Ca3 判别较 差的有 Chaetoceros curvisetus(Cu), Chaetoceros (Ca), Db, T_r . Thalassiosira nordenskioeldii Cleve(Tn), Leptocylindrus danicus (Ld), Rh, Ks, Prorocentrum minimum(Pm), Gy, Rhodomonas salina(Ra), Sy。对于Ca3 判别较差的这些藻, Ca4 对于 Cf; Ca5 对于 Db, Gy; Cd3 对 于 S_{V} ; Cd4 对于 Tr 和 Rh; 这些藻能够给予互补提高。因此 本文中选择第三层尺度分量作为一级识别特征谱,第4~5 层尺度分量及第3~5层小波分量作为二级识别特征谱,共 同组成互补标准谱库。以第三层尺度分量标准谱库作为识别 的第一阶段, 当识别结果落入上述 Cf, Db, Gy, Sy, Tr, Rh这些藻时,则进入相应的互补谱库进行第二次识别。

2.3 浮游植物标准谱库及相应识别测定技术的建立

本文获得了大量反映浮游植物不同生长条件下(光照,温度、生长期等)的荧光特征谱,这其中存在一些极为相似的特征谱,相似特征谱的存在会对识别产生不利影响,因此,需要提取出能够代表各门类和各属浮游植物荧光特征且尽可能少的特征谱——浮游植物荧光标准谱。

本文采用系统聚类法对每种浮游植物在不同培养条件下得到的第一培养样的 3~5 层尺度分量和小波分量特征谱分别进行聚类分析。以欧氏距离作为特征谱相似程度的衡量,将特征谱聚为不同的类别,然后取每一类中所有特征谱的平均作为该类的代表,即为该浮游植物的一条标准谱。由此确定每一种浮游植物的荧光标准谱及其数目(图 2),形成 43 种浮游植物的荧光标准谱库。

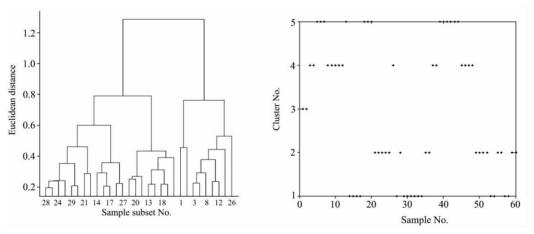


Fig. 2 Hierarchical clustering and samples belonging graphic of the pseudo-nitzschia pungens

在此基础上,以多元线性回归辅以非负最小二乘法建立相应的浮游植物识别测定技术。识别时,首先进入门水平上的标准谱库进行门水平上的识别,如果是非赤潮样品,不进入属水平上的识别,直接给出门水平上的结果;如果是赤潮样品,则按照门水平上的识别结果,进入到相应门的属水平上的标准谱库,进行属水平上的识别。以上工作利用 Mat-

lab6. 5 软件完成。

3 结果与讨论

用上述所得标准谱库及所建立的识别测定技术对模拟混合样品及单种藻样品分别进行识别测定,所得结果如下。

Table 1 Identifying results of 32 red tide algae at the level of genus

	门属	藻种	优势度 70%	优势度 80%	优势度 90%	优势度 100%
硅藻	拟菱形藻属	Ps	65. 5	81. 7	93. 7	98. 3
	骨条藻属	Sk	62. 7	81. 0	91. 5	85. 0
	角毛藻属	Cu	80. 7	90. 4	99. 3	100. 0
		De	84. 4	93. 3	100.0	100. 0
		Ca	78. 8	85. 2	94. 1	100. 0
	盒形藻属	Oc	38. 4	48. 6	48. 6	36. 7
	圆筛藻属	Cf	_	_	_	_
		Cs	_	_	_	_
	双尾藻属	Db	20. 4	31. 7	27. 5	68. 3
	海链藻属	Tr	5. 9	6. 7	2. 2	1. 7
	海链藻属	Tn	62. 2	57. 8	74. 1	45. 5
		Tw	87. 4	96. 3	100.0	100. 0
	细柱藻属	Ld	45. 8	75. 4	89. 4	75. 0
	根管藻属	Rh	94. 4	97. 2	100.0	67. 6
金藻	棕囊藻属	Cg	80. 6	99. 3	100.0	70.0
	等边金藻属	I_g	_	_	_	_
	颗石藻属	Ks	_	_	_	_
	定鞭金藻属	Pp	93. 0	95. 8	99. 3	95. 8
	硅鞭藻属	Se	6. 2	3. 4	0. 0	75. 0
	叉鞭金藻属	Di	_	_	_	_
绿藻	小球藻属	ch	_	_	_	_
	盐藻属	Ds	_	_	_	_
	四爿藻属	Pu	_	_	_	_
	微囊藻属	Mp	_	_	_	_
甲藻	亚历山大藻属	Al	91. 5	97. 9	100.0	98. 3
	前沟藻属	Am	59. 9	78. 2	97. 2	66. 7
	原甲藻属	Pr	91. 7	99. 2	100.0	100. 0
		Pm	87. 1	93. 9	100.0	100. 0
		Ma	81. 1	90. 9	99. 2	78. 3
		Pd	90. 2	97. 0	100.0	100.0
	裸甲藻属	Gs	64. 7	84. 2	99. 3	100.0
		Gy	12. 2	12. 2	7. 2	41. 7
	凯伦藻属	Kb	99. 3	100.0	100.0	33. 3
		Km	49. 6	66. 7	85. 1	50.0
	斯比藻属	Sc	62. 7	77. 5	100.0	98. 3
黄藻	异弯藻属	На	71. 8	85. 2	96. 5	78. 3
	卡盾藻属	Cm	90. 1	99. 3	100.0	100.0
隐藻	隐藻属	Rs	100.0	100.0	100.0	100.0
		Ra	100.0	100.0	100.0	98. 3
蓝藻	鱼腥藻属	Cy	0. 0	0. 0	0. 0	83. 3
	螺旋藻属	Lx	_	_	_	_
	聚球藻属	Sy	_	_	_	_
	束毛藻属	Te	100.0	100.0	100.0	91. 7

对 32 种赤潮藻不同优势度模拟混合样品及单种藻属水平识别结果分析,当优势种比例分别为 70%, 80%和 90%时,32 种赤潮藻在属水平上识别的平均正确率分别为 67.5%, 75.8%, 81.4%; 即当优势度大于 90%时,能对混合样品达到较好的属水平上正确识别。对于单种藻样品,属水平识别的平均正确率为 79.4%,识别较差的主要为 Oc,Tr, Tn, Gy, Kb, Km。

Table 2 Identifying results of 43 phytoplankton at the level of division

	/D ## P= /0/	>= == = = = = = = = = = = = = = = = = =	定量百分比/%	
优势门	优势度/%	识别正确率/%-	范围	平均
	50	98. 2	15.8∼91.6	47. 2
硅藻	75	99. 9	20. 2%~99. 8	66 . 6
	100	91. 7	66.1~100	90.8
	50	97. 7	29. 6~47. 8	37. 7
金藻	75	100	46.9∼73.5	61. 7
	100	92. 6	72.0~100	88. 6
	50	95. 7	6. 2∼50. 4	34. 9
绿藻	75	99. 8	28. 2∼76. 0	60.4
	100	78. 3	35.8∼100	84. 0
	50	99. 0	27. 5~48. 5	40. 2
甲藻	75	99. 1	46. 9∼72. 3	64. 9
	100	97. 5	36.3∼100	90.0
	50	91. 7	20.4~29.2	24. 8
黄藻	75	98. 9	39. 3∼57. 8	48. 5
	100	89. 2	76. 4~100	86. 7
	50	100	44. 0%~47. 5	45. 8
隐藻	75	100	63. $5\sim$ 67. 1	65 . 3
	100	99. 2	99.1~100	99. 6
	50	83. 9	23. 2~46. 1	36. 3
蓝藻	75	100	62. 7~100	75 . 1
	100	94. 7	83. 3~100	93. 8

对 43 种近海常见浮游植物不同门类优势度样品的门水平识别结果分析,当优势度为 50%时,平均定量测定百分比为 38.1%;平均识别比率为 95.2%。优势度为 75%时,平均定量测定百分比为 63.2%;平均识别比率为 99.7%。门类优势度为 100%时,门水平的定量测定百分比为: $84.0\% \sim 99.6\%$,平均 90.5%,平均识别比率为 91.9%。识别结果较差的是绿藻门,其中绿藻门内的 Ch 易被误判为金藻,它与金藻门中 Ig 的特征谱同在第 35 点处有最大荧光值并在 38 点处有次大值,从而容易造成误判,推测可能是在培养时与 Ig 串了种。

对胶州湾采集的样品和麦岛围隔样品进行识别,胶州湾样品为 2007 年 8 月 21 日所取的 12 个站位的水样,硅藻为调查区域内的绝对优势门类,优势度范围为 94.64%~99.97%;麦岛围隔样品为 2007 年 7 月 27 日及 30 日的 12 个围隔水样,其中硅藻为 10 个水样的优势门,优势度为 76.19%~99.95%,甲藻为剩余 2 个水样的优势门,优势度为 56.67%~70.49%。对样品按照相同的处理方法,得到分解后的尺度分量和小波分量后,采用上述所建立的标准谱库进行识别,得到胶州湾现场样品的识别正确比率为门类水平可达 100%,定量识别比例 0.62~1,平均 96.8%;麦岛围隔样品的识别正确比率为门类水平 91.7%,定量识别比例 0~1,平均 85.1%。

4 结 论

综上所述,本文通过结合单色发光二极管的实际特点及

仪器节能、便携等方面的要求,从原始三维荧光光谱中选取了 12 个激发波长点组成浮游植物的离散三维荧光光谱,并对 43 种浮游植物(包括 32 种赤潮藻)不同比例的模拟混合样品及单种藻进行了门、属水平上的定性识别及门水平上的定

量测定分析,结果表明所建立的浮游植物荧光识别技术是可行的,为研制使用以单色发光二极管为光源的适于现场应用的浮游植物荧光自动监测仪提供了技术支撑。

References

- [1] Suzuki K, Handa N, Kiyosawa H. J Oceanogr, 1997, 53; 383.
- [2] Pech-Pacheco J L, Alvarez-Borrego J. Marine Biology, 1998, 132: 357.
- [3] Sieracki C K, Sieracki M E, Yentsch C S. Marine Ecology Progress Series, 1998, 168; 285.
- [4] Lewitus AJ, White DL, Tymowski RG. Estuaries, 2005, 28: 160.
- [5] Mackey M D, Mackey D J, Higgins H W. Marine Ecology Progress Series, 1996, 144: 265.
- [6] Mura Yama-ka Yano E, Yoshimatsy S, Kayano T. J. Ferment Bjioeng, 1998, 85(3): 343.
- [7] Walsh D, Reeves R A, Saul D J. Bioche System and Eco., 1998, 26(5): 495.
- [8] Yentsh C S, Phinney D A. Journal of Plankton Research, 1985, 7: 617.
- [9] Seppala J, Balode M. Hydrobiologia, 1998, 363: 207.
- [10] Lee T, Tsuzuki M, Takeuchiet T. Analytical Chimica Acta, 1995, 302(1): 81.
- [11] Beutler M, Wiltshire K H, Meyeret B. Photosynthesis Research, 2002, 72(1): 39.
- [12] ZHANG Qian-qian, WANG Xiu-lin, ZHU Chen-jian(张前前,王修林,祝陈坚). Marine Environmental Science(海洋环境科学), 2004, 23(1): 77.
- [13] ZHANG Qian-qian, LEI Shu-he, WANG Xiu-lin(张前前, 类淑河, 王修林). High Technology Letters(高技术通讯), 2005, 25(4): 75.
- [14] ZHANG Fang, SU Rong-guo, WANG Xiu-lin(张 芳, 苏荣国, 王修林). Chinese Journal of Lasers(中国激光), 2008, 35(12): 2052.
- [15] ZHANG Fang, WANG Liang, SU Rong-guo(张 芳,王 良,苏荣国). Chinese Journal of Sensors and Actuators(传感技术学报), 2007, 20(10); 2143.
- [16] LIU Bao, SU Rong-guo, SONG Zhi-jie(刘 宝, 苏荣国, 宋志杰). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2010, 30 (5); 1275.
- [17] Zepp R G, Sheldon W N, Moran M A. Marine Chemistry, 2004, 89: 15.

Research on the 3D Discrete Fluorescence Spectrum Technique for Differentiation of Phytoplankton Population

ZHANG Shan-shan¹, SU Rong-guo^{1*}, DUAN Ya-li¹, SONG Zhi-jie², WANG Xiu-lin¹

- Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, Ocean University of China, Qingdao 266100, China
- 2. College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China

Abstract The present research was targeted to develop a fluorescence analyser for phytoplankton population which uses a series of LEDs as the light source. So the 3D discrete fluorescence spectra with 12 excitation wavelengths (400, 430, 450, 460, 470, 490, 500, 510, 525, 550, 570 and 590 nm) were determined by fluorescence spectrophotometer for 43 phytoplankton species. Then, the wavelet, Daubechies-7 (Db7), and Bayes Classifier were applied to extract the characteristics for each classes from the 3D discrete fluorescence spectra. Lastly, the fluorescence differentiation method for phytoplankton populations was established by multivariate linear regression and non-negative least squares, which could differentiate phytoplankton populations at the levels of both divisions and genus. This method was tested; for simulatively mixed samples (the dominant species accounted for 70%, 80%, 90% and 100% of the gross biomass, respectively) from 32 red tide algal species, and the correct discrimination rates at the level of genus were 67. 5%, 75. 8%, 81. 4% and 79. 4%, respectively. For simulatively mixed samples (the dominant divisions algae accounted for 50%, 75% and 100% of the gross biomass, respectively) from 43 algal species, the discrimination rates at the level of division were 95. 2%, 99. 7% and 91. 9% with average relative content of 38. 1%, 63. 2% and 90. 5%, respectively.

Keywords Phytoplankton population; 3D discrete fluorescence spectrum; Wavelet analysis; Characteristics extraction; Differentiation

* Corresponding author

(Received May 12, 2010; accepted Sep. 27, 2010)