

赤潮藻类流式图像自动分析算法的研究

谢杰镇¹ 骆庭伟¹ 戴君伟¹ 王迪² 高艳³ 冉升³

(厦门大学海洋与地球学院 厦门 361005)¹ (厦门大学信息科学与技术学院 厦门 361005)²

(重庆西信天元数据资讯有限公司 重庆 401127)³

摘要 赤潮是常见海洋自然灾害之一。为了早期预测和预报赤潮的发生,开发了基于流式图像的赤潮监测仪器,它将流式细胞、显微成像以及图像处理技术结合起来,采用基于背景差的方法快速准确地分割出藻类图像;为了克服藻类细胞处于不同生长周期和环境带来的形态和个体差异的影响,提取具有平移、旋转和尺度不变性的几何形状特征与基于灰度共生矩阵的纹理特征,采用一对一的多分类支持向量机进行分类识别。实验结果表明,该算法实现了对海洋原甲藻、红色裸甲藻、拟菱形藻和中肋骨条藻的自动分类,平均识别准确率高达 94.37%。

关键词 赤潮,图像分割,特征提取,藻类识别,支持向量机

中图分类号 TP391.41 文献标识码 A

Algorithm for Automatic Recognition of Red Tide Algal Images Captured by Flow Cytometry

XIE Jie-zhen¹ LUO Ting-wei¹ DAI Jun-wei¹ WANG Di² GAO Yan³ RAN Sheng³

(College of Ocean and Earth Sciences, Xiamen University, Xiamen 361005, China)¹

(School of Information Science and Technology, Xiamen University, Xiamen 361005, China)²

(Chongqing Xixintianyuan Data Information Limited Company, Chongqing 401127, China)³

Abstract The red tide is a global marine natural disaster. In order to predict and forecast the occurrence of red tide, with the technology of flow cytometry, micro-image and image processing, a real-time harmful algae monitoring system was developed. A method based on background subtraction was used to quickly and accurately segment the algae images. In order to overcome the influence of algal cells in morphology and individual difference brought by the different growth period and environment, geometry features which are invariant to translation, rotation and scale and texture features based on GLCM were extracted. And finally, one-to-one multi-class support vector machine was adopted for identification. The experimental results show that the average recognition accuracy rate is as high as 94.37%.

Keywords Red tide, Image segmentation, Feature extraction, Algae recognition, Support vector machines

1 引言

赤潮是在特定的环境条件下,海水中某些浮游植物、原生动物或细菌爆发性增殖或高度聚集而引起水体变色的一种有害生态现象。据国家海洋局发布的中国海洋环境质量及灾害公报,从2000年至今,每年赤潮灾害发生累计面积都超过10000平方公里^[1],给海产养殖业和渔业资源等带来了严重的影响。加强对赤潮藻的监测是预防赤潮的主要手段之一,传统方法大多采用显微镜对海水样品进行人工分类和分析来完成对赤潮的监测。而此方法劳动强度大、效率低、精度低,必须由专业的人员才能完成。因此,建立赤潮生物的有效监测方法,研制一套赤潮藻类图像自动识别系统,是亟待解决的课题。

赤潮生物流式图像现场监测技术将流式细胞技术、显微

成像技术、图像处理技术结合起来,实现对目标赤潮生物的现场快速监测。该技术通过流式细胞技术提供定量检测的平台,其快速准确的定量对于“赤潮”这一与细胞数量直接相关的生态学现象的监测具有特别重要的意义;显微成像技术具有分类学形象、直观的优点;利用图像处理技术,自动分析和识别藻类。通过对这3个技术的结合,达到对赤潮生物进行识别和快速定量的目的。

2 赤潮藻识别原理与方法

目前,国内外不少科研人员对赤潮藻类图像分类识别算法进行了研究。最有代表的是美国 Fluid Imaging Technology 公司研发研制的 FlowCAM,这套系统比显微镜简便且耗时少,图像识别软件获得了80%~90%的分类正确率^[2],但图像仍必须由专业的技术员确认。当前国际上对水下流式细胞仪及

到稿日期:2012-09-17 返修日期:2012-12-25 本文受国家863计划(2012AA092002)资助。

谢杰镇(1977-),男,博士,工程师,主要研究方向为数字图像处理、仪器研发,E-mail: xjz@xmu.edu.cn;骆庭伟(1979-),男,硕士,工程师,主要研究方向为微型生物生态学;戴君伟(1982-),男,硕士,工程师,主要研究方向为嵌入式系统开发;高艳(1987-),女,硕士,工程师,主要研究方向为数字图像处理;冉升(1977-),男,硕士,工程师,主要研究方向为数据建模和文本挖掘。

成像系统也做了不少研究,如原位成像流式细胞仪 FlowCytobot。FlowCytobot 是以激光为基础的系统,它可以“看到”激光所激发产生荧光的单细胞,成像主要是荧光信号。姬光荣等^[3]通过光学显微镜构建了一个基于 Web 的有害赤潮显微图像采集与诊断系统。王锐、于新生等^[4]采用 SVM 和神经网络两种方案对浮游生物进行识别,基于显微镜成像建立一套监测系统。利用显微镜采集图像,在现场及走航的数据分析应用中存在一定缺陷。

为了能够对赤潮藻类长期监测,并能直观地观察藻类的形态,本文改进了厦门大学的海洋赤潮生物实时采集系统,并提出一种赤潮藻类流式图像自动分类识别算法。算法处理流程如图 1 所示。核心步骤主要分为:基于背景差法进行图像分割、特征提取、支持向量机分类器设计。

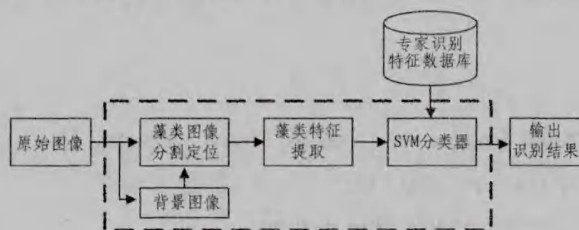


图 1 图像处理算法流程图

3 藻类图像采集、分割

将海水样品从上到下高速流经流动室,流动室是一个狭窄的通道,横截面大小为 $2\text{mm} \times 100\mu\text{m}$ 。经显微系统放大后利用 CCD 采集图像。

据统计,常见的海洋赤潮藻类细胞的体长和宽度分布范围在 $10\mu\text{m} \sim 200\mu\text{m}$ 之间,改进后的系统采用千兆网 CCD 摄像机 DMK 23G445,分辨率为 1280×960 ,在流动室和 CCD 之间添加一个 10 倍的显微物镜,通过 $10\mu\text{m}$ 的标准荧光小球校准,CCD 在流动室上实际有效检测区域大约为 $400\mu\text{m} \times 300\mu\text{m}$,可以有效地采集到清晰的赤潮藻类图像。

采集子系统控制 LED 以 3 次/秒的速度闪烁,触发 CCD 以 3 帧/秒进行图像采集。类似于相机照相,通过闪烁来控制 CCD 曝光,当曝光时间为 $10\mu\text{s}$ 时,藻类细胞高速运动造成的拖影轨迹小于 $1\mu\text{m}$,从而可以分辨出体长和宽度在 $10\mu\text{m}$ 的小藻类细胞。

海水中的杂质会慢慢在流动室沉积,为了不使这些沉积的杂质影响藻类分析和识别,需对采集的图像进行处理,即去除背景信息^[5]。在开始采集图像时,统计 20 帧图像的特点,将其作为起始背景信息;在采集过程中,定期定量修改背景信息。

对处理后的图像进行分割,通过设定阈值,获取藻类图像所在位置。分割的依据主要有灰度信息、纹理信息、尺度信息等。由于流式细胞仪是通过背景光源投射过成像区域成像的,细胞比纯净的环境液体对光线的通过能力差,故细胞区域的灰度亮度低。藻类细胞内部是细胞质等结构,在图像上反映为纹理较为丰富的形式,而气泡等杂质内部为单纯的结构,在图像上纹理细节少是区别细胞与杂质的重要线索。赤潮藻类细胞的尺寸一般大于 $10\mu\text{m}$,通过这个特征可以区别一些

小的杂质。

图 2(a)为改进后的厦门大学的海洋赤潮生物实时采集系统采取的藻类原始图像,图像大小 1280×960 ;图 2(b)为用本文方法提取的背景图;图 2(c)为对图 2(a)与图 2(b)的差值图进行二值化并形态学处理之后的效果图;图 2(d)为分割出的藻类位置图。从实验效果可以看出,本文方法可以很好地提取出背景图,实现藻类的准确快速分割,获取的典型赤潮藻类图像如图 3 所示。

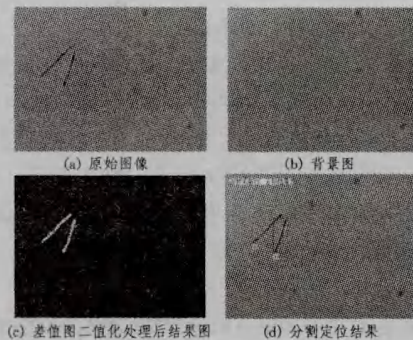


图 2 基于背景差的藻类图像分割定位图



图 3 获取的赤潮藻流式图像示例

4 特征提取

对图 3 的藻类进行了分析,发现藻类的几何形状特征存在较大的差别。为了克服藻类细胞不同生长周期和环境带来的形态和个体差异的影响,以及图像采集时视觉不同带来的影响,结合藻类本身的特点,提取藻类细胞的具有旋转、平移、和尺度不变性的 7 个不变矩、最小外接矩形的长宽度、矩形度、圆度、体态比作为藻类的几何形状特征。

赤潮生物为活体目标,不同种类的藻也可能具有相似的几何形状。因此,仅仅提取几何形状特征是不够的,还需要提取局部细节特征,即提取藻类的纹理特征。

特征提取是实现后续正确分类识别藻类的基础。在整个识别算法中占着举足轻重的作用。因此,在准确获得分割出的藻类图像之后,如何对藻类图像提取出可靠性高、可区别性强、独立性好、维数低的特征成为一个重要的问题。本文提取的特征主要分为两个方面:几何形状特征和纹理特征。

4.1 几何形状特征

在进行人工识别藻类时,专家主要是根据藻类外形上的区别来辨识^[6-8]。可以看出,几何形状特征含有大量的信息。

a) 7 个不变矩

不变矩最初是由 M. K. Hu 在 1962 年提出的,常用于描述图像的形状特征。图像发生平移、旋转或者尺度缩放时,其值基本保持不变。

通过对归一化的二阶和三阶中心矩进行非线性组合,即可得到下列的 7 个不变矩。它们对平移、旋转以及尺度变换具有不变性^[9]:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (1)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (2)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{33})^2 \quad (3)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{33})^2 \quad (4)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{33})^2] + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] \quad (5)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{33}) \quad (6)$$

$$\phi_7 = (3\eta_{21} - \eta_{33})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{33})^2] + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] \quad (7)$$

式中, η_{pq} 为归一化的 $(p+q)$ 阶中心矩。

b) 矩形度

$$R = \frac{A}{W \cdot H} \quad (8)$$

式中, A 表示物体区域的面积; W, H 分别表示藻类的宽度和长度。矩形度用于描述藻类与矩形的偏离程度。

c) 圆度

$$T = \frac{4\pi A}{C^2} \quad (9)$$

式中, C 表示藻类的周长,其主要量化与圆形的符合程度。

d) 体态比

$$C_1 = \frac{W}{H} \quad (10)$$

体态比即为藻类最小外接矩形的宽度与高度的比值。比重越接近于 1, 表明该物体越紧凑。

4.2 基于灰度共生矩阵的纹理特征

纹理是由图像灰度在空间位置上交替变化形成的,反映了图像灰度级的空间分布情况,是图像分析的重要指标之一。灰度共生矩阵是一种通过研究灰度的空间相关特性来描述纹理的常用方法。其数学公式表示为^[10]:

$$p(i, j, \delta, \theta) = \{[(x, y), (x+\Delta x, y+\Delta y)] | f(x, y) = i, f(x+\Delta x, y+\Delta y) = j; x=0, 1, \dots, N_x-1; y=0, 1, \dots, N_y-1\} \quad (11)$$

式中, $i, j=0, 1, \dots, L-1$; L 为图像的灰度级数; (x, y) 为图像中的像素坐标; N_x, N_y 表示图像对应的行列数。可以看出,灰度共生矩阵统计了图像中 (x, y) 处灰度为 i 、经过距离为 δ 、方向为 θ 的平移到 $(x+\Delta x, y+\Delta y)$ 处其灰度为 j 的像素点对概率。

本文试验中选取了 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 共 4 个方向距离为 1 的灰度共生矩阵。

具体实验中往往不是直接应用计算的灰度共生矩阵作为纹理特征进行分析,而是通过共生矩阵导出一些能够更直观

地描述纹理状况的参数。典型的有以下几种:

a) 能量

$$ASM = \sum_{i=1}^N \sum_{j=1}^N (p(i, j))^2 \quad (12)$$

b) 对比度

$$CON = \sum_{i=1}^N \sum_{j=1}^N (i-j)^2 p(i, j) \quad (13)$$

c) 相关

$$COR = \sum_{i=1}^N \sum_{j=1}^N \frac{ijp(i, j) - u_i u_j}{s_i s_j} \quad (14)$$

其中,

$$u_i = \sum_{i=1}^N \sum_{j=1}^N i p(i, j)$$

$$u_j = \sum_{i=1}^N \sum_{j=1}^N j p(i, j)$$

$$s_i^2 = \sum_{i=1}^N \sum_{j=1}^N p(i, j) (i - u_i)^2$$

$$s_j^2 = \sum_{i=1}^N \sum_{j=1}^N p(i, j) (j - u_j)^2$$

d) 熵

$$ENT = - \sum_{i=1}^N \sum_{j=1}^N p(i, j) \log p(i, j) \quad (15)$$

e) 逆差距

$$IDM = - \sum_{i=1}^N \sum_{j=1}^N \frac{p(i, j)}{1 + (i-j)^2} \quad (16)$$

5 分类器设计

本文采用支持向量机分类器对藻类进行识别。支持向量机(Support Vector Machine, SVM)是 Vapnik 等人提出的基于最小结构风险和 VC 维理论的二分类方法,它能较好地解决小样本、高维数、非线性和局部极小点等问题^[11],具有较好的推广性。

给定样本集 $(x_i, y_i), i=1, 2, \dots, n, x_i$ 为样本向量, y_i 为类别标识符, $y_i \in \{-1, 1\}$ 。支持向量机的核心是最终求得最优超平面,即可转化为求解如下所示的二次规划方程:

$$\begin{cases} \max Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \sum_{i=1}^n \alpha_i y_i = 0, i=1, 2, \dots, n \\ 0 \leq \alpha_i \leq C, i=1, 2, \dots, n \end{cases} \quad (17)$$

式中, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 为拉格朗日乘子 α_i 组成的向量; n 为样本的个数; $K(x_i, x_j)$ 为核函数; C 为惩罚因子。

采用径向基核函数 RBF 的支持向量机作为分类器需要确定的参数:核函数中的参数和惩罚系数 C 。常用的 RBF 核函数是如下形式:

$$K(x, x_i) = \exp(-\gamma \|x_i - x_j\|^2) \quad (18)$$

6 实验结果与分析

为了验证本文提出的识别方法,对现场海水中提取的海洋原甲藻、红色裸甲藻、拟菱形藻、中肋骨条藻 4 类藻种培养后进行实验。选取 476 个样本作为训练样本,260 个样本为测试样本。

本文算法中对分割出的藻类提取了几何形状特征和纹理特征,共 32 维的特征。采用一对一的多分类支持向量机对其

进行识别,并采用网格化搜索算法获取参数 $C=4.00, \gamma=0.0312$ 。其识别结果如表 1 所列。

表 1 分类识别结果

藻类名称	测试 样本数	识别 正确数	错误 识别数	据识数	识别 准确率
海洋原甲藻	79	69	6	4	87.34%
红色裸甲藻	67	64	0	3	95.52%
拟菱形藻	60	59	0	1	98.33%
中肋骨条藻	54	52	2	0	96.30%
平均识别率					94.37%

从表 1 的识别准确率可以看出,本文算法对海洋原甲藻、红色裸甲藻、拟菱形藻、中肋骨条藻的识别效果较好,平均识别率达到 94.37%。

实验结果表明 SVM 对于简单的分类问题显现出很好的性能,即种类少则边界数目少,边界划分简单,分类性能高。

实验中以顶刺作为重要特征来区分海洋原甲藻与其他的赤潮生物。由于拍摄角度的问题,有些海洋原甲藻的顶刺未被拍摄到,进而影响海洋原甲藻的识别准确率。红色裸甲藻因拍摄角度而存在 2 种形状特征,通过建立它们的特征关联来提高识别准确率。中肋骨条藻和拟菱形藻都是长条状,在尺寸上相似,但中肋骨条藻有明显的圆形细胞组成的链状结构,而拟菱形藻为梭形细胞,首尾相叠形成链状群体,以此特征可较好地对其区分。

结束语 本文介绍了一种基于流式图像的赤潮藻类监测方法,提出了用背景差法准确地分割定位出藻类细胞,以平移、旋转和尺度不变性的几何形状特征和基于灰度不变共生矩阵的纹理特征共同构成藻类的特征集,采用一对一的多分类支持向量机进行训练识别。利用本文所提出的算法对培养的藻类样品进行了试验,结果表明该算法能够克服藻类细胞

不同生长周期和环境带来的形态和个体差异影响,对不同角度得到的藻类细胞完成识别,实现对海洋原甲藻、红色裸甲藻、拟菱形藻、中肋骨条藻的快速分类识别。

参考文献

- [1] 国家海洋局. 2000—2012 年中国海洋灾害公报[OL]. <http://www.coi.gov.cn/gongbao/zaihai/>
- [2] Buskey E J, Hyatt C J. Use of the FlowCAM for semi-automated recognition and enumeration of red tide cells (*Karenia brevis*) in natural plankton samples[J]. *Harmful Algae*, 2006, 5: 685-692
- [3] 乔小燕,姬光荣. 基于生物形态学的赤潮藻显微图像分割与特征提取研究[D]. 青岛:中国海洋大学,2010
- [4] 王锐,于新生,唐颖,等. 图像自动识别技术在海洋浮游生物分析中的应用[J]. *海洋科学*, 2007, 31(10): 61-66
- [5] 蔡丽. 基于背景差法的车辆停车检测算法研究[D]. 西安:长安大学,2011
- [6] 郑少平,谢杰镇,王博亮,等. 基于形状特征描述子的赤潮优势种识别研究[J]. *厦门大学学报:自然科学版*, 2009, 48(5): 676-679
- [7] 江涛,王程,王博亮,等. 基于 SVDD 和 SVM 的赤潮藻类识别[J]. *厦门大学学报:自然科学版*, 2010, 49(1): 47-51
- [8] 余肖翰,谢杰镇,郑少平. 傅里叶识别分析在赤潮生物自动识别中的应用研究[J]. *海洋湖沼通报*, 2011(4): 11-15
- [9] 郭晓宇. 基于 Hu 矩和 Zernike 矩的图像目标识别算法设计[D]. 沈阳:沈阳航空工业学院,2009
- [10] 徐小军,邵英,郭尚芬. 基于灰度共生矩阵的火焰图像纹理特征分析[J]. *计算技术与自动化*, 2007, 26(4): 64-67
- [11] 高艳. DR 图像铸件工件号识别算法的进一步研究[D]. 重庆:重庆大学,2012

(上接第 292 页)

- [3] Staib L H, Duncan J S. Boundary fitting with parametrically deformable models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(11): 1061-1075
- [4] Chakraborty A, Staib L H, Duncan J S. Deformable boundary finding in medical images by integrating gradient and region information [J]. *IEEE Transactions on Medical Imaging*, 1996, 15(6): 859-870
- [5] Caselles V, Kimmel R, Sapiro G. Geodesic active contours[A]// *Proceedings of International Conference on Computer Vision* [C]. Boston, MA, USA, 1995: 694-699
- [6] Malladi R, Sethian J A, Vemuri B C. Shape modeling with front propagation: A level set approach [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(2): 158-174
- [7] Sapiro G. *Geometric Partial Differential Equations and Image Analysis*[M]. Cambridge, UK: Cambridge University Press, 2001
- [8] 唐利明,何传江,申小娜. 几何活动轮廓模型的多尺度扩散分割算法[J]. *计算机辅助设计与图形学学报*, 2007, 19(5): 661-666
- [9] Caselles V, Kimmel R, Sapiro G. Geodesic active contours[J]. *International Journal of Computer Vision*, 1997, 22(1): 61-79
- [10] Chan T, Vese L. Active contours without edges [J]. *IEEE Transactions on Image Processing*, 2001, 10(2): 266-277
- [11] Li C M, Kao C Y, Gore J C, et al. Minimization of region-scalable fitting energy for image segmentation [J]. *IEEE Transactions on Image Processing*, 2008, 17(10): 1940-1949
- [12] Mishra A K, Fieguth P W, Clausi D A. Decoupled active contour (DAC) for boundary detection [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2011, 33(2): 310-324
- [13] Yu W, Franchetti F, Chang Y J, et al. Fast and robust active contours for image segmentation[C]// *Proceedings of the 17th IEEE International Conference on Image Processing*. Los Alamitos: IEEE Computer Society Press, 2010: 641-644
- [14] 胡同森,项方云. 使用标识分水岭算法确定 Snake 模型的初始轮廓[J]. *浙江工业大学学报*, 2008, 36(4): 374-377
- [15] 李丽勤,高焕文,周兴祥. Snake 模型初始轮廓选取的研究[J]. *计算机工程与应用*, 2004, 11: 43-45
- [16] 刘彩霞,范延滨,杨厚俊. GVF Snake 模型中一种新的初始轮廓设置方法[J]. *计算机应用*, 2006, 26(7): 1614-1619
- [17] Soille P. *Morphological Image Analysis Principles and Applications*[M]. Berlin, Germany: Springer Verlag, 1999: 123-140
- [18] Kass M, Witkin A, Terzopoulos D. Snake active contour models [J]. *International Journal of Computer Vision*, 1987, 1(4): 321-331
- [19] Lankto S, Tannenbaum A. Localizing Region-Based Active Contours [J]. *IEEE Transactions on Image Processing*, 2008, 11(17): 2029-2039