

# 一种甲藻显微图像顶刺区域分割及特征提取方法

乔小燕<sup>1</sup> 姬光荣<sup>2</sup>

(山东工商学院数学与信息科学学院 烟台 264005)<sup>1</sup> (中国海洋大学信息科学与工程学院 青岛 266100)<sup>2</sup>

**摘要** 为了提高藻种显微图像识别率,提出了一种基于自适应形态学的快速藻种顶刺分割和特征提取方法。该方法引入像素宽度的概念,以藻种目标像素宽度直方图和面积分布图为判别依据,自动判定最佳结构元尺寸,以联合滤波、区域归并等形态学操作分割顶刺区域;最后构造了两类局部生物形态学特征提取参数。实验证明,该算法针对不同甲藻细胞目标可自动计算最佳结构元素大小,具有精度高、速度快的优点,且两类特征提取参数具有视觉不变性。

**关键词** 甲藻,顶刺分割,自适应结构元素

中图分类号 TP391.4 文献标识码 A

## Method for Spine Segmentation and Feature Extraction of Pyrrophyta Microscopic Image

QIAO Xiao-yan<sup>1</sup> JI Guang-rong<sup>2</sup>

(College of Mathematic and Information Science, Shandong Institute of Business and Technology, Yantai 264005, China)<sup>1</sup>

(Department of Electronic Engineering, Ocean University of China, Qingdao 266100, China)<sup>2</sup>

**Abstract** A method for spine segmentation based on adaptive mathematical morphology was proposed in order to increase the correct rate of Pyrrophyta recognition. At first, the pixel-width was introduced, and the optimal structure element was computed by pixel-width histogram and area distribution automatically, and then the spine was extracted by mixed mathematical morphology operations. Finally, two kinds of local biological morphology feature parameters were constructed and their visual invariance was proved. The experiments results show that the optimal structure element can be computed by different Pyrrophyta cell and this method is of high precision and fast speed.

**Keywords** Pyrrophyta, Spine segmentation, Adaptive structure element

## 1 引言

作为21世纪的重要学科,计算机技术和生物技术已经渗透到国民经济的方方面面。藻种鉴定作为生命科学和计算机视觉技术的交叉研究领域,已经引起了自然科学界的广泛关注。其中,甲藻是重要的浮游植物类群<sup>[1]</sup>,在许多领域具有广泛的应用,如赤潮监测、水质检测、环境调查、石油勘探等,而这些应用都离不开对甲藻的鉴定工作。

目前针对藻类的自动鉴定和定量方法主要有吸收光谱法<sup>[2,3]</sup>、荧光光谱法<sup>[4-6]</sup>、液相色谱法<sup>[7-9]</sup>、流式细胞仪<sup>[10-12]</sup>和分子技术<sup>[13-15]</sup>等,但是这些方法大多过程繁琐,严重依赖于藻类的生理状态,且分辨率不高(除分子技术外,大多只能分类到门或纲级别)。近年来,依据藻类形态结构特征所进行的显微图像识别方法逐步成为一种发展潜力大而且很实用的方法<sup>[16,17]</sup>。

藻种细胞的形状和纹理是藻种分类的基本特征。国际上,欧共体“硅藻识别和分类的自动系统(ADIAC)”<sup>[18,19]</sup>、英国普利茅斯研究组“人工神经网络甲藻鉴定(DiCANN)”<sup>[20]</sup>、美国伍兹霍尔海洋研究所研究员Davis<sup>[21]</sup>与Tang等人<sup>[22,23]</sup>分别在淡水硅藻、甲藻、浮游动物细胞显微图像识别方面做了大量工作;国内,厦门大学王博亮等<sup>[24]</sup>、厦门大学高亚辉<sup>[16]</sup>、

罗金飞<sup>[17]</sup>、天津大学王明时等<sup>[25]</sup>对赤潮藻、硅藻显微图像进行了识别研究,并提出了新的藻种图像分割组方法。

藻种细胞是变化的三维活体目标,其形态随目标游动发生变化,与静态图像相比在自动识别上难度较大,由于获取条件和视点的差异也会导致图像形态的变化。从模式识别角度来看,藻种鉴定是多类别、每种类别个体目标形态多变的复杂目标识别问题。虽然以上关于藻种鉴别的研究取得了丰富的研究成果,但如今普遍的藻种特征提取技术均是对藻种形状等参数的测量,并未针对藻种本身的生物学信息进行提取和度量,这是导致藻种识别率不高的主要原因。因此,必须提取同一类藻种在不同时期和不同形态之间保持较高稳定性的细节特征用于识别,即藻种生物形态学特征。尽管藻种细胞的生物特征区域在整幅图像中所占的面积并不大,但对这些区域错误的特征描述可能致使错误的分类,因此对藻种生物特征区域的准确而有效的提取将大大提高目标分类的效率和识别率,这是藻种生物学特征提取的重点。本文从数学形态学入手,研究一种能快速、准确分割甲藻细胞顶刺区域并建立特征度量的方法。

## 2 数学形态学

近年来,数学形态学发展成为一种有效的图像处理的非

到稿日期:2011-08-13 返修日期:2012-01-11 本文受国家自然科学基金(11171191),山东省高等学校科研计划项目(J11LG13)资助。

乔小燕(1982-),女,博士,讲师,主要研究方向为图像处理与模式识别,E-mail:mibaohao@126.com;姬光荣(1955-),男,教授,主要研究方向为模式识别、电子信息处理。

线性理论和研究方法,其基本思想是用具有一定形态的结构元素去度量和提取图像中对应的形状,以达到对图像分析和处理的目的<sup>[18]</sup>。数学形态学可用来解决平滑抑噪、边缘提取、纹理分析等图像处理问题,具有简洁、直观、速度快等优点。结构元素的选择是将形态学应用于某个具体处理问题的关键和难点。

数学形态学主要用于处理二值图像,也可以扩展到对灰度图像进行操作,其基本运算操作包括膨胀和腐蚀。设  $X$  为图像集合,  $B$  为结构元素,  $X$  被  $B$  腐蚀定义为:

$$X \ominus B = \{z | (B)_z \subseteq X\} \quad (1)$$

式中,  $(B)_z$  表示  $B$  将平移  $z$ 。  $X$  被  $B$  膨胀定义为:

$$X \oplus B = \{z | (B)_z \cap X \neq \emptyset\} \quad (2)$$

式中,  $(B)_z$  表示  $B$  的相对于它自身原点的映像。将腐蚀和膨胀运算进行复合以及集合操作,可以构造出最为重要的运算:开运算和闭运算。使用结构元素  $B$  对图像  $X$  进行开运算定义为:

$$X \circ B = (X \ominus B) \oplus B \quad (3)$$

使用结构元素  $B$  对图像  $X$  进行闭运算定义为:

$$X \cdot B = (X \oplus B) \ominus B \quad (4)$$

开、闭运算对图像具有独特的平滑滤波作用。开运算可抑制边界上的毛刺,截断细窄的连接带,分离接触区域,消除比结构元素小的孤点、碎线和斑块,并使图像的内边缘平滑。闭运算可填补边界上的凹陷部分,具有填充小洞和小孔、弥合裂缝等作用,并使外边缘更加光滑。在形态滤波中,开、闭运算在滤除噪声斑点或填充空洞的同时,能够完好地保护特定的局部结构特征。

### 3 甲藻顶刺区域分割

#### 3.1 甲藻生物特征介绍

甲藻门(Pyrrophyta)藻种广泛分布于海洋和淡水中,一般热带种类多、个体小,寒带种类少、个体大。其多为单细胞生物,也有单细胞连成的不同形状的群体,形状大多呈球形、针形或分枝形<sup>[1]</sup>。

甲藻生物形态学区域一般具有形状、亮度、形态等特征方面的视觉显著性。图 1 是甲藻生物形态学分类方法,横沟、顶刺、凹槽等均为藻种生物学特征。对这类藻种图像的生物特征区域的提取除了要用到一般的区域分割方法外,还需要结合生物学领域中的知识,才能做到合理提取。

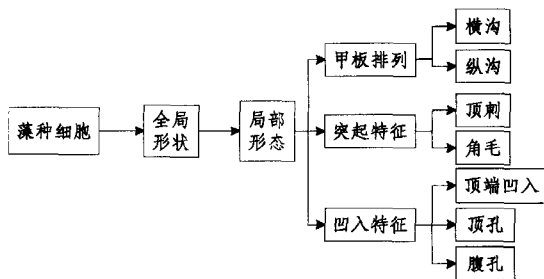
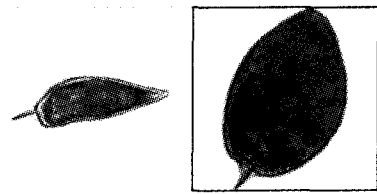


图 1 甲藻生物形态学分类

细胞的顶刺是甲藻的重要特征。是否有顶刺以及顶刺的长短、粗细都是甲藻门细胞分类的一个重要依据,因此,对顶刺的判别与提取对藻种细胞识别具有重要意义。若精确分割出顶刺区域,就可以计算其区域描述几何特征,将其与细胞形状参数一同作为特征向量用于藻种识别。图 2 给出了两种带顶刺特征的甲藻细胞显微图像。



(a) 反曲原甲藻 (b) 海洋原甲藻

图 2 有顶刺特征的甲藻示意图

鉴于藻细胞顶刺的形态学特点,目前可用来提取顶刺位置比较有代表性的方法有细化法<sup>[26]</sup>、梯度法<sup>[27]</sup>等。这两种算法计算量都较大,实时性较差,满足不了藻种显微图像实时监测的需要。

#### 3.2 算法原理

通过对大量含顶刺特征的甲藻显微图像分析可知,顶刺着生部位与细胞主体部分宽度的差异程度与藻种类别有密切关系。图 3 为对图 2(a)正放反曲原甲藻目标图像进行列扫描时,扫描线经过背景→顶刺→主体→背景时的目标像素个数变化图。该图显示当扫描线经过顶刺部位时有一个小的脉冲产生,脉冲的高度即为顶刺部位的横截面的像素点个数,脉冲的宽度即为顶刺部位的长度;由于细胞主体部位相对面积较大,且横截面宽度较大,因此当扫描线经过细胞主体时形成了一个大的凸形,凸起的高度即为细胞在该垂直方向上的横切面的直径,凸起的宽度即为细胞的主体长度。当扫描的结果中出现高度相差较大的两个脉冲时,可推断藻种细胞图像可能存有顶刺。

如前所言,形态开运算恰恰具有断开狭窄连接部位的作用,将原图像与形态开运算的结果进行比较可提取出图像中的峰值。因此,可采用开运算进行顶刺判别和目标区域提取。

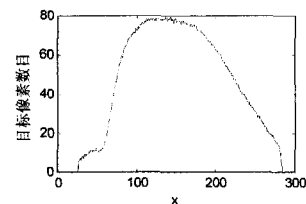


图 3 反曲原甲藻目标像素个数变化曲线

#### 3.3 位置归一化

因采集方法和角度不同,实验获取的藻种图像目标主轴与  $x$  轴正向存在一定夹角,首先需要通过计算目标惯性主轴的方向实现目标位置归一化,实现细胞的准确定位。如图 4 所示,在直角坐标系中,设坐标轴的原点通过细胞的质心,直线  $y = x \tan \theta$  是通过细胞质心的一条直线,反曲原甲藻目标像素  $(x, y)$  对直线  $y = x \tan \theta$  的惯性矩为  $r^2$ ,其中  $r = x \sin \theta - y \cos \theta$ ,若图形上所有点到直线的惯性矩最小,则该直线叫做图形的惯性主轴。图形的几何质心可表明物体的位置,而惯性主轴可表明其摆放角度,通过计算惯性主轴的方向来标识细胞旋转的角度,再做逆向旋转变换可以将图像的位置摆正。

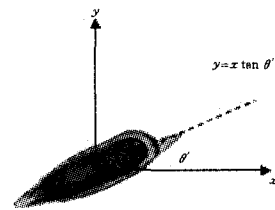


图 4 反曲原甲藻的惯性主轴

### 3.4 结构元素选取

为了估计顶刺横截面宽度,实现顶刺部位与细胞体完整地分离,首先引入像素宽度(pixel-width,  $\omega$ )的概念:将目标旋转至水平方向后,在  $x$  轴方向上进行等距采样,并在每个采样点上作  $x$  轴垂线与形状边界相交,计算并统计该垂线落入目标区域内的像素点数目,这个数目即为该采样点上的像素宽度<sup>[23]</sup>。假设一幅二值图像为:

$$f(i, j) = \begin{cases} g_f, & f(i, j) \geq T \\ g_b, & f(i, j) < T \end{cases} \quad (5)$$

式中,  $i, j \in [1, N], \forall i \in [1, N]$ 。

$$\omega_j = \sum_{i=1}^N k(i, j), k(i, j) = \begin{cases} 1, & f(i, j) = g_f \\ 0, & f(i, j) = g_b \end{cases} \quad (6)$$

即像素宽度  $\omega_j$  是对图像第  $j$  列中目标像素数目的度量值,其中  $g_f$  为二值化后前景像素灰度值。目标的像素宽度直方图(pixel-width histogram)  $L(\omega)$  横坐标值为图像中所有的像素宽度,纵坐标值为统计具有某像素宽度的采样点数目;目标的面积分布图纵坐标  $S(\omega)$  表示为  $S(\omega) = L(\omega) \times \omega$ , 即计算某采样点像素宽度与个数的乘积,  $S(\omega)$  中的最大值即为目标中显著组成部分的面积。

图5为图2(a)反曲原甲藻的像素宽度直方图与面积分布结果,(a)中用圆圈标明的脉冲处对应的横坐标(10与78)分别代表了目标中两个最显著的像素宽度,即尖刺部位平均宽度与细胞体平均宽度。(b)中用圆圈标明的脉冲处对应的横坐标(78)代表了目标中具有最大面积的像素宽度。

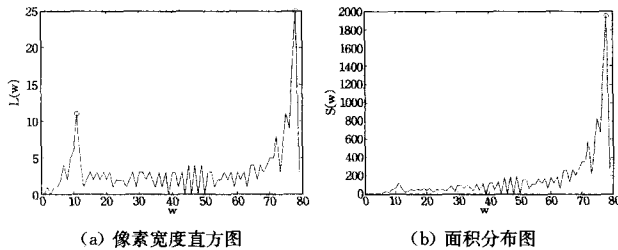


图5 反曲原甲藻像素宽度计算结果

以下将像素宽度面积分布中面积比例判别作为顶刺横截面宽度的估计依据,设面积分布图中像素宽度阈值  $\omega'$  将  $S(\omega)$  划分为两个部分,若两部分的面积积分值满足

$$T_1 \leq \frac{\int_0^{\omega'} S(\omega) d\omega}{\int_0^{\max(\omega)} S(\omega) d\omega} \leq T_2 \quad (7)$$

即面积比值在  $[T_1, T_2]$  时,认为顶刺区域存在,并将满足条件的阈值集合  $\{\omega'_1, \omega'_2, \dots, \omega'_k\}$  作为结构元素的估计尺寸集合,  $T_1, T_2$  为实验中确定的阈值。从中选择具有最多像素宽度个数的  $\omega$  值,即满足

$$L(\omega^*) = \max\{L(\omega'_1), L(\omega'_2), \dots, L(\omega'_k)\} \quad (8)$$

的  $\omega^*$  作为最佳结构元尺寸。

### 3.5 算法流程

1) 图像分割:藻种显微图像二值化及目标提取;  
2) 位置归一化:计算细胞目标区域的惯性主轴的方向,以惯性主轴方向表示细胞目标旋转的角度,逆向旋转变换可以将图像的位置摆正,实现细胞的精确定位;

3) 列扫描图像:将二值图像向水平方向投影,获取图像的像素宽度直方图和面积分布直方图作为选取阈值的一维信号,选取式(8)中  $\omega^*$  值作为最佳结构元尺寸;

4) 顶刺提取:利用原图像减去用所选取的结构元素大小进行开运算的结果可获得图像中的凸出部位;

5) 形态开运算:以小结构元素做形态开运算,去除孤立点、不连续的碎线和虚假区域;

6) 区域标记:形态开运算后的结果包含了若干个小的连通部分,按照指定连通规则(四连通或八连通)找到图像中所有连通部分,为同一连通成分中所有像素点分配同样的标记,不同连通区域享有不同标记;计算各连通成分面积,将最大连通成分面积的区域提取出来,即为顶刺区域。

## 4 顶刺特征提取

几何参数和不变矩只描述了藻种细胞的全局形状信息,这对具有复杂生物特征的识别目标来说往往是不够的。因此,在顶刺区域分割的基础上,根据藻种形态特征与机器识别特征对应关系,同时考虑这些特征应具有视觉不变性,本文构造了两类局部形态特征参数(顶刺与细胞面积比、顶刺分长轴比)用于识别分类,此参数为度量性比值,理论上具有平移、缩放和旋转不变性。其中,顶刺与细胞面积比  $f_1$  定义为顶刺部位面积  $S_1$  与细胞目标区域面积  $S_2$  的比值,顶刺分长轴比  $f_2$  定义为顶刺区域长轴  $L_2$  与细胞区域长轴  $L_1$  的比值,如图6所示。

$$f_1 = \frac{S_2}{S_1} \quad (9)$$

$$f_2 = \frac{L_2}{L_1} \quad (10)$$

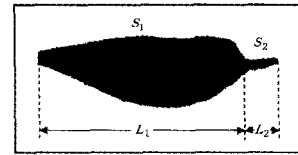


图6 顶刺特征计算示意图

## 5 实验结果

实验所用藻细胞图像是对藻种取样后通过 Olympus BX61 配置高灵敏高速显微照相系统 Retiga 4000R FAST 1394CCD 拍摄所得。图像尺寸统一为  $512 \times 512$  像素,每像素 8bit。上述算法采用 MATLAB 编程,运行在 2.4G CPU、512M 的 PC 机上。

### 5.1 顶刺分割结果

实验对图2中两类藻种显微图像运用自动确定最佳结构元的形态学方法分割细胞体顶刺区域,结果如图7所示。

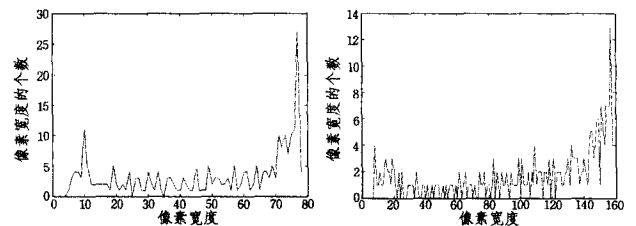


图7 细胞图像结构元素选取

如图7所示,(a)直方图横轴最大值为80,而(b)直方图横轴最大值为160,这反映了反曲原甲藻的显著宽度比海洋原甲藻大;(a)直方图的纵轴最大值为30,而(b)直方图纵轴

最大值为 14,这反映了反曲原甲藻目标中具有显著像素宽度的位置比图(b)中多;(a)直方图的密度比(b)直方图密度小,反映了反曲原甲藻的像素宽度比海洋原甲藻变化缓慢;从像素宽度直方图求得反曲原甲藻和海洋原甲藻目标选择的最佳结构元素尺度分别为 10 和 8,实际上这分别与它们在顶刺根部的横截面宽度大小是一致的。

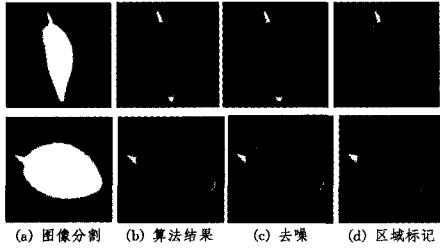


图 8 实验图像顶刺分割结果

图 8 给出了两类藻种顶刺部位分割结果,其中图(c)与图(d)分别为加入边缘检测后的结果。实验结果表明,利用像素宽度直方图和面积分布作为判别依据可精确获得最佳形态学结构元素大小,提取甲藻细胞顶刺区域,具有精度高的优势。

另外,实验选取了 100 幅甲藻细胞显微图像,分别采用梯度法、细化法和本文方法来分割藻种顶刺部位。图 9 比较了 3 种方法所用的时间,结果显示采用形态学分割顶刺部位比其他两种方法所用时间明显减少,且运算速度快。

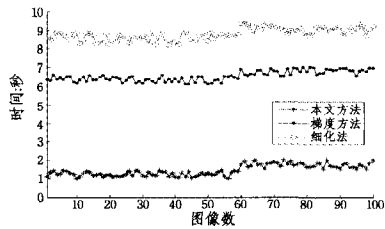


图 9 本文方法与梯度法、细化法所用时间比较

## 5.2 特征提取结果

如图 10 所示,对藻种原始图像分别做顺时针旋转 45 度(A1)、旋转 90 度(A2)、旋转 135 度(A3)、旋转 180 度(A4)、放大 1 倍(A5)和缩小 1 倍(A6)操作作为测试图像集合,按照式(9)和式(10)计算藻种局部生物特征参数,计算结果见表 1。实验结果表明,藻种在不同的尺度和旋转角度下,两类顶刺形态特征参数在同一种内稳定,且具有视觉不变性,因此这些特征可以选入识别分类的特征集合。



图 10 实验图像集

表 1 顶刺特征参数结果

原图	A1	A2	A3	A4	A5	A6
反曲 $f_1$	0.0247	0.0234	0.0234	0.0245	0.0245	0.0240
原甲藻 $f_2$	0.1765	0.1688	0.1688	0.1757	0.1757	0.1727
海洋 $f_1$	0.0113	0.0108	0.0108	0.0106	0.0106	0.0113
原甲藻 $f_2$	0.1243	0.1207	0.1207	0.1200	0.1200	0.1253

**结束语** 本文在分析甲藻细胞生物形态学特征的基础上,基于目标不同部位横切面像素个数的差异,将目标图像的

像素宽度直方图和面积分布作为判别依据,自动选取最佳结构元尺寸,结合有序的形态学运算精确分割顶刺区域,并最终构造了顶刺与细胞面积比、顶刺分长轴比两类局部生物学特征。经大量实验验证,该方法理论完备,原理直观,算法快捷,能快速、完整地分割原始藻种顶刺区域,并保留藻种细节特征,且特征提取参数具有平移、缩放和旋转不变性。

## 参考文献

- [1] 齐雨藻. 赤潮[M]. 广州: 广东科技出版社, 1999: 1-65
- [2] Kirkpatrick G J, Millie D E, Moline M A, et al. Optical discrimination of a phytoplankton species in natural mixed populations [J]. *Limnol Oceanogr*, 2000, 45(2): 467-471
- [3] 李继刚, 王世忠, 高志, 等. 藻类和悬浮物的光谱识别研究[J]. 高技术通讯, 2000(7): 68-71
- [4] Cowles T J, Desiderio R A, Neuer S. In situ characterization of phytoplankton from vertical profiles of fluorescence emission spectra [J]. *Marbiol*, 1993, 115: 217-222
- [5] Seppaelae J, Balode M. The use of spectral fluorescence methods to detect changes in the phytoplankton community [J]. *Hydrobiologia*, 1998, 363(1-3): 207-217
- [6] 张前前, 类淑河, 王修林, 等. 浮游植物活体三维荧光光谱分类识别方法研究[J]. 光谱学与光谱分析, 2004, 24(10): 1227-1229
- [7] Furuya K, Hayashi M, Yabushita Y, et al. Phytoplankton dynamics in the East China Sea in spring and summer as revealed by HPLC2 derived pigment signatures [J]. *Deep Sea Research Part II Topical Studies in Oceanography*, 2003, 50(2): 367-387
- [8] 王海黎, 洪华生, 徐立. 反相高效液相色谱法分离、测定海洋浮游植物的叶绿素和类胡萝卜素 [J]. 海洋科学, 1999, 4: 6-9
- [9] 姚鹏, 于志刚, 米铁柱. 海洋浮游藻类的化学分类法 [J]. 海洋环境科学, 2003, 22(1): 75-80
- [10] Balfout H W, Snoek J, Smiths J R M, et al. Automatic identification of algae: neural network analysis of flow cytometric data [J]. *J Plankton Res*, 1992, 14(4): 575-589
- [11] 焦念志, 杨燕辉. 四类海洋超微型浮游生物的同步监测 [J]. 海洋与湖沼, 1999, 30(5): 506-511
- [12] 张利华, 张经, 晁敏. 流式细胞术对超微型浮游植物识别初探 [J]. 海洋科学, 2002, 26(3): 60-65
- [13] Scholin C A, Marin R, Miller P E, et al. DNA probes and a receptor-binding assay for detection of *Pseudo-nitzschia* (Bacillariophyceae) species and domoic acid activity in cultured and natural samples [J]. *J Phycol*, 1999, 35: 1356-1367
- [14] Scholin C A, Anderson D M. Two distinct small subunit ribosomal RNA genes in the north American toxic dinoflagellate *Alexandrium fundyense* (Dinophyceae) [J]. *Journal of Phycology*, 1993, 29: 209-216
- [15] Scholin C A, Anderson D M. Identification of group and strain specific genetic markers for globally distributed *Alexandrium* (Dinophyceae) [J]. I. RFLP analysis of SSU rRNA genes. *Journal of Phycology*, 1994, 30(4): 744-754
- [16] 高亚辉, 杨军霞, 骆琦琦, 等. 海洋浮游植物自动分析和识别技术 [J]. 厦门大学学报: 自然科学版, 2006, 45(S2): 40-45
- [17] 高亚辉, 罗金飞, 骆琦琦, 等. 数学形态学在海洋浮游植物显微图像处理中的应用 [J]. 厦门大学学报: 自然科学版, 2008, 47(S2): 242-245
- [18] du Buf H, Bayer M M. Automatic Diatom Identification [J]. World Scientific, Series in Machine Perception and Artificial In-

[19] Jalba A C, Wilkinson M H F, et al. Automatic diatom identification using contour analysis by morphological curvature scale spaces[J]. Machine Vision and Applications, 2010, 16(4): 217-227

[20] 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

[21] Davis C S, Fredrik T, et al. A three-axis fast-tow digital Video Plankton Recorder for rapid surveys of plankton taxa and hydrography[J]. American Society of Limnology and Oceanography, Methods, 2005(3): 59-74

[22] Tang X, Stewart W K, Huang H, et al. Automatic plankton image

recognition[J]. Artificial Intelligence Review, 1998, 12: 177-199

[23] Tang X, Lin F, Remsen A. Binary plankton image classification [J]. IEEE Journal of Oceanic Engineering, 2006, 31(3): 728-735

[24] 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

[25] 王明时, 王学民, 张磊. 显微图像分析技术在赤潮生物识别中的应用[J]. 天津大学学报, 2004, 37(5): 392-395

[26] Ruiz L A, Molto E, et al. Location and characterization of the stem-calyx area on oranges by computer vision[J]. Journal of Agricultural Engineering, 1996, 64: 165-172

[27] Sarkar N, Wolfe R R. Feature extraction techniques for sorting tomatoes by computer vision[J]. Transactions of the ASAE, 1985, 28(3): 970-974

(上接第 244 页)

已存在的对齐关系就可以推导出更复杂的关系。



图 6 复合关系图

**结束语** 近年来, 范畴论被用于开发本体描述的新方法及解决视图更新问题的新途径<sup>[21]</sup>。范畴论的高度概括、抽象和更强、更直观的表达力为本体集成提供了一种框架, 这种特性可屏蔽本体间的异构。本文将范畴论引入本体对齐的研究中, 在范畴论的基础上, 结合地理本体特征, 对态射进行了重新定义, 构造了一个更复杂的范畴。其中范畴中的对象为地理本体, 而态射不仅能表达相等关系, 也可以描述地理本体间的多种语义关系。在态射定义的基础上, 根据态射的特性定义了本体对齐和本体对齐的复合, 而研究本体对齐的最终目的是实现本体的集成, 因此对范畴论提供的本体集成的方法之一即外推进行了阐述。

将范畴论引入本体对齐的研究领域中的最大特点在于, 其表达能力和可扩展性较强, 不仅能表达多种语义关系, 而且能通过本体对齐的复合从已存在的对齐关系中推导出新的对齐关系, 可用于在分布式系统中多个本体的联合查询等任务。因此, 范畴论为地理信息本体语义集成提供了一种适合的形式化框架。

### 参 考 文 献

[1] Lambrix P, Tan H. A framework for aligning ontologies[M]. Principles and Practice of Semantic Web Reasoning, 2005: 17-31

[2] Bouquet P, Euzenat J, Franconi E, et al. Specification of a common framework for characterizing alignment[J]. Kweb deliverable, D, 2004, 2

[3] Euzenat J. Algebras of ontology alignment relations [C]// The 7th International Semantic Web-ISWC. 2008: 387-402

[4] Shareha A, Rajeswari M, Ramachandram D. Multimodal Integration (Image and Text) Using Ontology Alignment[J]. American Journal of Applied Sciences, 2009, 6(6): 1217-1224

[5] Lambrix P, Tan H. Ontology alignment and merging [M]. Anatomy Ontologies for Bioinformatics: Principles and Practice. Springer, 2008: 133-149

[6] Cafezeiro I, Haeusler E. Semantic interoperability via category

theory [C]// ER'07 Tutorials, Posters, Panels and Industrial Contributions at the 26th International Conference on Conceptual modeling. New Zealand: Australian Computer Society, Inc., 2007, 83: 197-202

[7] Hitzler P. Integrated view and comparison of alignment semantics[J]. Deliverable, Knowledge Web NoE, 2005, 2

[8] Oyenon W, DeLoach S. Using Category Theory to Compose Multiagent Organization Design Models[R]. Manhattan, K. S. University, 2010: 1-34

[9] 王红. 国家基础地理信息本体研究 [D]. 武汉: 武汉大学, 2009: 1-124

[10] 李霖, 朱海红, 王红, 等. 基于形式本体的基础地理信息语义分析—以陆地水系要素类为例[J]. 测绘学报, 2008, 38(2): 230-242

[11] 胡玲, 李霖, 王红. 基于范畴论的地理本体集成研究[J]. 计算机工程与设计, 2010, 31(16): 3732-3735

[12] Zimmermann A, Krotzsch M, Euzenat J, et al. Formalizing ontology alignment and its operations with category theory[J]. Frontiers in Artificial Intelligence and Applications, 2006, 150: 277-289

[13] Pierce B. Basic category theory for computer scientists [M]. Cambridge: The MIT Press, 1991: 1-20

[14] Zimmermann A, Euzenat J. Three semantics for distributed systems and their relations with alignment composition [C]// The Semantic Web-ISWC 2006. 2006: 16-29

[15] 于娟, 党延忠. 本体集成研究综述[J]. 计算机科学, 2008, 35(7): 9-13

[16] de Bruijn J, Ehrig M, Feier C, et al. Ontology mediation, merging and aligning [M]// Davies J, Studer R, Warren P. Semantic Web Technologies: trends and research in ontology-based system. West Sussex (UK): John Wiley, 2006

[17] Kalfoglou Y, Schorlemmer M. Ontology mapping: the state of the art[J]. The knowledge engineering review, 2003, 18(1): 1-31

[18] Euzenat J, Valtchev P. Similarity-based ontology alignment in OWL-lite [C]// Ios Pr Inc. 2004: 333-337

[19] Euzenat J, Barrasa J, Bouquet P, et al. D2. 2. 3: State of the art on ontology alignment [OL]. <http://www.starlab.vub.ac.be/research/projects/knowledgeweb/kweb-223.pdf>, 2004

[20] Solomon A. Pushout: A Mathematical Model of Architectural Merger[J]. Lecture Notes in Computer Science, 2007, 4378: 389-399

[21] Johnson M, Rosebrugh R. Ontology Engineering, Universal Algebra, and Category Theory [M]. Theory and Applications of Ontology, Computer Applications, 2010: 565-576