



(12) 发明专利

(10) 授权公告号 CN 103063202 B

(45) 授权公告日 2015. 04. 15

(21) 申请号 201210588045. 6

(22) 申请日 2012. 12. 30

(73) 专利权人 同济大学

地址 200092 上海市杨浦区四平路 1239 号

专利权人 复旦大学

(72) 发明人 林怡 潘琛 王嘉楠 任文伟

叶勤 屈铭志 刘冰 陆渊

(74) 专利代理机构 上海科盛知识产权代理有限

公司 31225

代理人 赵继明

(51) Int. Cl.

G01C 11/00(2006. 01)

G01C 11/04(2006. 01)

(56) 对比文件

CN 102135531 A, 2011. 07. 27,

US 2005164333 A1, 2005. 07. 28,

潘琛. 基于小波核双重加权 SVM 模型的蓝藻水华识别与变化检测. 《上海国土资源》. 2013, 第 34 卷 (第 4 期),

林怡等. 基于遥感影像光谱分析的蓝藻水华识别方法. 《同济大学学报(自然科学版)》. 2011, 第 39 卷 (第 8 期),

审查员 张礅

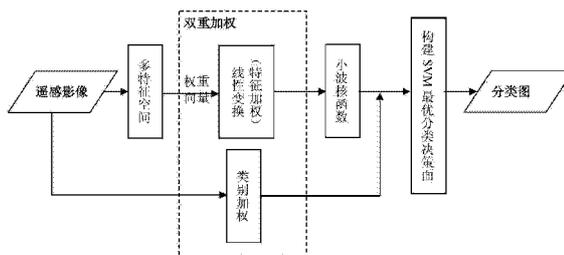
权利要求书3页 说明书12页 附图4页

(54) 发明名称

基于遥感影像的蓝藻生物量时空变化监测与可视化方法

(57) 摘要

本发明涉及一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法,该方法包括以下步骤:1) 对研究区的遥感影像进行预处理,并构建归一化蓝藻指数;2) 采用基于 VPRS_GID 的特征优化模型对遥感影像的特征进行优化,获得优化后的多特征空间;3) 根据多特征空间建立基于小波核的双重加权 SVM 分类模型,对蓝藻水华空间分布信息进行提取识别和变化检测,结合实地观测数据进行综合验证和精度分析;4) 将处理后的遥感影像、GIS 矢量数据与实地观测数据进行叠加显示,从而实现蓝藻水华爆发时空变化过程和规律的仿真模拟。与现有技术相比,本发明具有蓝藻识别精度和可靠性高等优点,有利于对蓝藻水华的成因和分布变化做出分析和判断。



1. 一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法,其特征在于,该方法包括以下步骤:

1) 获取蓝藻研究区域的遥感影像,并对其进行图像预处理,构建归一化蓝藻指数;

2) 采用基于 VPRS_GID 的特征优化模型对遥感影像的特征进行优化,获得优化后的多特征空间;

3) 根据多特征空间建立基于小波核的双重加权 SVM 分类模型,获取 SVM 最优分类决策面,并利用该分类模型提取蓝藻空间分布信息、检测蓝藻空间分布的变化,结合实地观测数据进行综合验证和精度分析;

4) 将处理后的遥感影像、研究区域的 GIS 矢量数据与该区域的实地观测数据进行叠加显示,从而实现蓝藻水华爆发时空变化过程和规律的仿真模拟;

所述的归一化蓝藻指数的构建具体包括:

11) 构建蓝藻的影像光谱曲线,分别对波段 4 与波段 3 以及波段 5 与波段 4 进行差值运算:

$$\begin{cases} a = B4 - B3 \\ b = B5 - B4 \end{cases}$$

式中, B3、B4 和 B5 分别表示波段 3、波段 4 和波段 5 的光谱值;

12) 对差值运算结果归一化处理:

$$\begin{cases} c = \min(a, b) \\ a' = a + |c| = a + |\min(a, b)| \\ b' = b + |c| = b + |\min(a, b)| \end{cases}$$

13) 通过比值运算,构建归一化蓝藻指数 NDI_CB:

$$NDI_CB = \frac{a'/b' - 1}{a'/b' + 1} = \frac{a' - b'}{a' + b'}$$

所述的双重加权包括对不同类别赋以权重和对每个特征属性赋以权重;

所述的建立基于小波核的双重加权 SVM 分类模型的具体步骤为:

31) 根据遥感影像的样本集,获得类别权重 λ_{y_i} :

$$\lambda_{y_i} = \begin{cases} 1 - \frac{l_1}{l}, y_i = 1 \\ \dots \\ 1 - \frac{l_n}{l}, y_i = n \end{cases}$$

式中, l_1 为类别 y_1 的样本数据, l 为样本总数, n 为类别数;

32) 构造特征权重向量 ω 和线性变换对角阵 P, 对多特征空间进行线性变换, 实现特征重要性加权;

33) 对特征加权后的多特征空间计算小波核函数, 构建基于小波核的双重加权 SVM 分类模型:

$$\left\{ \begin{array}{l} \max : L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i^T P, x_j^T P) \\ = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \prod_{n=1}^d \left[\frac{\cos 1.75 \left(\left((x_i^T P)_n - (x_j^T P)_n \right) \right)}{a_n} \right] \cdot \exp \left(- \frac{\left\| \left((x_i^T P)_n - (x_j^T P)_n \right) \right\|^2}{2a_n^2} \right) \\ S.T. \quad C \lambda_{y_i} \geq \alpha_i \geq 0, i=1, 2, \dots, l \\ \sum_{i=1}^l \alpha_i y_i = 0 \end{array} \right.$$

式中, α_i 为第 i 个拉格朗日乘子, α_j 为第 j 个拉格朗日乘子, x_i 表示第 i 个输入的度量指标分类特征向量, x_j 表示第 j 个输入的度量指标分类特征向量, y_i 表示第 i 个输出的学习结果, y_j 表示第 j 个输出的学习结果, C 为常量, 是对越界样本点施加的惩罚系数, $K(x_i^T P, x_j^T P)$ 为 Morlet 小波核函数:

$$K(x_i^T P, x_j^T P) = \prod_{n=1}^d \left[\frac{\cos 1.75 \left(\left((x_i^T P)_n - (x_j^T P)_n \right) \right)}{a_n} \right] \cdot \exp \left(- \frac{\left\| \left((x_i^T P)_n - (x_j^T P)_n \right) \right\|^2}{2a_n^2} \right);$$

所述的实地观测数据包括实地蓝藻生物量采样数据和气象观测数据;

所述的叠加显示具体为:

- 41) 读取实地观测数据, 并对遥感影像和矢量地图数据进行空间坐标配准;
- 42) 在遥感影像数据层上叠加具有空间坐标的实地观测数据;
- 43) 对离散的实地观测数据进行插值处理, 利用插值数据拟合出生物量等值线信息;
- 44) 以经过坐标空间配准的遥感影像数据为底地图, 在其上依次叠加矢量地图数据、生物量等值线信息, 然后对处理结果及其对应日期的气象观测数据进行可视化显示和仿真模拟。

2. 根据权利要求 1 所述的一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法, 其特征在于, 所述的图像预处理包括对遥感影像进行大气层顶表观反射率反演。

3. 根据权利要求 1 所述的一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法, 其特征在于, 所述的步骤 2) 具体包括:

21) 根据遥感影像获取遥感影像的特征信息, 采用变精度粗糙集模型对特征信息进行约简;

22) 对约简得到的特征集采用灰色关联决策进行最佳特征组合选取, 获得优化后的多特征空间。

4. 根据权利要求 1 或 3 所述的一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法, 其特征在于, 所述的多特征空间包括影像波段组合、归一化蓝藻指数、最佳指数特征组合和最佳 Gabor 滤波器纹理特征组合。

5. 根据权利要求 1 所述的一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法, 其特征在于, 所述的构造特征权重向量 ω 具体为:

$$\omega = GR = (\sqrt{\text{GainRatio}(\text{feature}_1)}, \dots, \sqrt{\text{GainRatio}(\text{feature}_n)})$$

式中, $\text{GainRatio}()$ 为信息增益率函数, $\text{GainRatio}() = \frac{\text{Gain}()}{\text{SplitInfo}()}$, $\text{SplitInfo}()$ 为某一特征所包含的信息熵, $\text{Gain}()$ 为该特征对一个数据样本集合划分的信息增益, feature_i 为特征标号, $i = 1, 2, \dots, n$ 。

6. 根据权利要求 1 所述的一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法, 其特征在于, 所述的坐标配准是指将坐标转换为归一化的坐标系统, 具体为:

a) 获得待转换数据的坐标范围, 即该数据在 x 轴方向的最小值 x_{\min} 、最大值 x_{\max} 和 y 轴方向的最小值 y_{\min} 、最大值 y_{\max} ;

b) 通过该坐标范围计算转换所需的尺度因子 Scale , 以 x 轴方向为例, 其计算公式如下:

$$\text{XScale} = \left[-\frac{x_{\min}}{x_{\max} - x_{\min}}, \frac{1}{x_{\max} - x_{\min}} \right];$$

c) 根据所得尺度因子将空间坐标转换为归一化坐标, 以 x 轴方向为例, 计算公式如下:

$$X_{\text{Conv}} = \text{XScale}_0 + \text{XScale}_1 * x_0$$

其中, x_0 为原始空间坐标, X_{Conv} 为转换后的归一化坐标; XScale_0 和 XScale_1 分别为尺度因子 XScale 的第一和第二个元素。

基于遥感影像的蓝藻生物量时空变化监测与可视化方法

技术领域

[0001] 本发明涉及一种水环境监测技术,尤其是涉及一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法。

背景技术

[0002] 从国内外富营养化及蓝藻水华的监测工作来看,由于受到自然条件和时空等因素限制,传统的监测方法具有一定局限性。成本高、耗时长,并很难对大面积湖泊的藻类的时空动态分布和变化做出全面的调查。随着遥感技术的不断进步,在广袤水域的污染识别和检测方面,使用卫星遥感数据已经成为大面积、迅速评估水质的有效方法。特别是利用高时空分辨率的多时相卫星影像数据结合已有的地理信息和实测数据,能全面、快速、可靠地识别和提取出蓝藻污染区域,进而获得蓝藻水华的污染情况。从而快速、全面掌握蓝藻时空分布变化信息,能实现控制蓝藻水华、评价蓝藻生态环境风险、研究蓝藻异常生长的原因以及建立水质的预警系统的作用。

[0003] 对水体中蓝藻水华的识别,国内外的研究者们使用的算法众多,包括使用归一化差值植被指数 (NDVI)、使用藻类的波谱反射和吸收特性、使用水体生物-光学模型等方法。使用这些方法的本质,都是对所获得的遥感影像进行分类提取,根据蓝藻污染水域本身的光谱特性以及其他生物学特性,从其中提取和识别出蓝藻水华污染的区域,从而对受污染水域进行分析。因此,从蓝藻在遥感影像上显示的光谱特性和生物学特性角度出发,蓝藻的识别方法可以分为两大类,一类是利用蓝藻本身的光谱特性,利用蓝藻污染水体和正常水体在不同波段吸收率和反射率的差异,使用归一化差值植被指数或构造其他评价指数的方法,对遥感影像进行解译,识别和提取出蓝藻污染区域;另一类则是利用蓝藻污染水体的生物学特性,利用蓝藻中所含的叶绿素 a、液泡结构等生物特性,通过遥感影像的光谱分析,对水体的叶绿素浓度含量等参数进行判断,进而获得蓝藻水华的污染情况。

[0004] 对于国内外目前使用的这两类蓝藻识别算法,其也有各自的优势和劣势。使用归一化差值植被指数或其他构造的植被指数,通过遥感数据反映的光谱信息,具有识别准确,易于处理的优点,在水体蓝藻浓度较高时,会在水面上产生类似陆上植被的光谱特征,识别处理效果较好,但是当水体蓝藻浓度较低时,蓝藻的光谱曲线与背景水体相混合,呈现高悬浮物水体的特性,给水华面积的提取带来了困难,还需要与其他地物光谱曲线相结合,进一步分析。使用含蓝藻水体的叶绿素 a 浓度和蓝藻液泡结构的生物特性的分类方法,具有适应性好,数据源广泛的特点。需要考虑水面其他水生植物,例如水草、浮萍等高叶绿素植物的影响因素,对于水深水质等环境背景因素也有要求。同时,高分辨率的影像对水体中叶绿素 a 浓度的判别有较好的效果,但是考虑到国内遥感研究的实际情况,价格高昂的高分辨率影像获取途径不易,还多使用中等分辨率的 MODIS 数据,影响了相应研究的深入。

[0005] 对于蓝藻水华变化情况的检测,目前国内外使用的主要方法是基于多时相的遥感影像,通过对一段时间序列内的遥感影像进行分析,对不同时相的遥感影像分别进行蓝藻的识别和提取,然后进行对比分析,获取蓝藻的变化情况,对蓝藻水华进行检测。针对我国

目前淡水水体污染严重的现状,使用遥感数据进行水污染的识别和检测有着广阔的应用前景。同时,遥感数据的应用也有其局限性,因此在使用遥感数据进行水污染识别和变化检测处理时,应当加强与环境、海洋、生物等学科的联系,使用多种方法同时进行,运用多学科的知识来弥补遥感数据识别与提取上的不足。

[0006] 在实际应用中,大面积浑浊水体和陆生植被将对蓝藻覆盖信息的分类和提取产生明显地干扰作用。传统的遥感影像分类技术利用单一的指数特征,如归一化植被指数、归一化差异水体指数 (NDWI) 等,其对蓝藻分布信息的提取较多受到自然条件和环境因素的影响。同时,目前类似的蓝藻生物量监测系统都缺少流域和水网分析,主要是单纯依赖遥感影像进行蓝藻爆发监测。无法在充分利用已有数字化地理信息的基础上对蓝藻水华的爆发机理、流域影响、风力和温度等环境和气象因素的影响进行分析和直观表达。从而难以对蓝藻水华的成因分析给以可靠支持和指导,也因此无法有效地实现蓝藻水华的时空分布和变化的全面分析。

发明内容

[0007] 本发明的目的就是为了解决上述现有技术存在的缺陷而提供一种蓝藻识别精度和可靠性高的基于遥感影像的蓝藻生物量时空变化监测与可视化方法。

[0008] 本发明的目的可以通过以下技术方案来实现:

[0009] 一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法,该方法包括以下步骤:

[0010] 1) 获取蓝藻研究区域的遥感影像,并对其进行图像预处理,构建归一化蓝藻指数;

[0011] 2) 采用基于 VPRS_GID 的特征优化模型对遥感影像的特征进行优化,获得优化后的多特征空间;

[0012] 3) 根据多特征空间建立基于小波核的双重加权 SVM 分类模型,获取 SVM 最优分类决策面,并利用该分类模型提取蓝藻空间分布信息、检测蓝藻空间分布的变化,结合实地观测数据进行综合验证和精度分析;

[0013] 4) 将处理后的遥感影像、研究区域的 GIS 矢量数据与该区域的实地观测数据进行叠加显示,从而实现蓝藻水华爆发时空变化过程和规律的仿真模拟。

[0014] 所述的图像预处理包括对遥感影像进行大气层顶表观反射率反演,所述的归一化蓝藻指数的构建具体包括:

[0015] 11) 构建蓝藻的影像光谱曲线,分别对波段 4 与波段 3 以及波段 5 与波段 4 进行差值运算:

$$[0016] \quad \begin{cases} a = B4 - B3 \\ b = B5 - B4 \end{cases}$$

[0017] 式中, B3、B4 和 B5 分别表示波段 3、波段 4 和波段 5 的光谱值;

[0018] 12) 对差值运算结果归一化处理:

$$[0019] \quad \begin{cases} c = \min(a, b) \\ a' = a + |c| = a + |\min(a, b)|; \\ b' = b + |c| = b + |\min(a, b)| \end{cases}$$

[0020] 13) 通过比值运算, 构建归一化蓝藻指数 NDI_CB :

$$[0021] \quad NDI_CB = \frac{a'/b'-1}{a'/b'+1} = \frac{a'-b'}{a'+b'}。$$

[0022] 所述的步骤 2) 具体包括 :

[0023] 21) 根据遥感影像获取遥感影像的特征信息, 采用变精度粗糙集模型对特征信息进行约简 ;

[0024] 22) 对约简得到的特征集采用灰色关联决策进行最佳特征组合选取, 获得优化后的多特征空间。

[0025] 所述的多特征空间包括影像波段组合、归一化蓝藻指数、最佳指数特征组合和最佳 Gabor 滤波器纹理特征组合。

[0026] 所述的双重加权包括对不同类别赋以权重和对每个特征属性赋以权重。

[0027] 所述的建立基于小波核的双重加权 SVM 分类模型的具体步骤为 :

[0028] 31) 根据遥感影像的样本集, 获得类别权重 λ_{y_i} :

$$[0029] \quad \lambda_{y_i} = \begin{cases} 1 - \frac{l_1}{l}, y_i = 1 \\ \dots \\ 1 - \frac{l_n}{l}, y_i = n \end{cases}$$

[0030] 式中, l_1 为类别 y_i 的样本数据, l 为样本总数, n 为类别数 ;

[0031] 32) 构造特征权重向量 ω 和线性变换对角阵 P , 对多特征空间进行线性变换, 实现特征重要性加权 ;

[0032] 33) 对特征加权后的多特征空间计算小波核函数, 构建基于小波核的双重加权 SVM 分类模型 :

$$[0033] \quad \begin{cases} \max : L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i^T P, x_j^T P) \\ = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \prod_{n=1}^d \left[\left(\frac{\cos 1.75 \left((x_i^T P)_n - (x_j^T P)_n \right)}{a_n} \right) \cdot \exp \left(- \frac{\| (x_i^T P)_n - (x_j^T P)_n \|^2}{2a_n^2} \right) \right] \\ S.T. \quad C \lambda_{y_i} \geq \alpha_i \geq 0, i = 1, 2, \dots, l \\ \sum_{i=1}^l \alpha_i y_i = 0 \end{cases}$$

[0034] 式中, α_i 为拉格朗日乘子, x_i 表示输入的度量指标分类特征向量, y_i 表示输出的学习结果, C 为常量, 是对越界样本点施加的惩罚系数, K 为特征加权核函数。

[0035] 所述的构造特征权重向量 ω 具体为 :

$$[0036] \quad \omega = GR = (\sqrt{GainRatio(feature_1)}, \dots, \sqrt{GainRatio(feature_n)})$$

[0037] 式中, $GainRatio()$ 为信息增益率函数, $GainRatio() = \frac{Gain()}{SplitInfo()}$, $SplitInfo()$ 为某一特征所包含的信息熵, $Gain()$ 为该特征对一个数据样本集合划分的信息增益, $feature_i$ 为特征标号, $i = 1, 2, \dots, n$ 。

[0038] 所述的实地观测数据包括实地蓝藻生物量采样数据和气象观测数据。

[0039] 所述的可视化显示具体为:

[0040] 41) 读取实地观测数据, 并对遥感影像和矢量地图数据进行空间坐标配准;

[0041] 42) 在遥感影像数据层上叠加具有空间坐标的实地观测数据;

[0042] 43) 对离散的实地观测数据进行插值处理, 利用插值数据拟合出生物量等值线信息;

[0043] 44) 以经过坐标空间配准的遥感影像数据为底地图, 在其上依次叠加矢量地图数据、生物量等值线信息, 然后对处理结果及其对应日期的气象观测数据进行可视化显示和仿真模拟。

[0044] 所述的坐标配准是指将坐标转换为归一化的坐标系统, 具体为:

[0045] a) 获得待转换数据的坐标范围, 即该数据在 x 轴方向的最小值 x_{min} 、最大值 x_{max} 和 y 轴方向的最小值 y_{min} 、最大值 y_{max} ;

[0046] b) 通过该坐标范围计算转换所需的尺度因子 $Scale$, 以 x 轴方向为例, 其计算公式如下:

$$[0047] \quad XScale = \left[-\frac{x_{min}}{x_{max} - x_{min}}, \frac{1}{x_{max} - x_{min}} \right];$$

[0048] c) 根据所得尺度因子将空间坐标转换为归一化坐标, 以 x 轴方向为例, 计算公式如下:

$$[0049] \quad X_{Conv} = XScale_0 + XScale_1 * x_0$$

[0050] 其中, x_0 为原始空间坐标, X_{Conv} 为转换后的归一化坐标; $XScale_0$ 和 $XScale_1$ 分别为尺度因子 $XScale$ 的第一和第二个元素。

[0051] 与现有技术相比, 本发明具有以下优点:

[0052] (1) 本发明针对蓝藻的特征, 构建了新的归一化蓝藻指数 (NDI_CB), 更加适用于提取低浓度蓝藻空间分布信息, 提高了蓝藻识别的可靠性和识别精度;

[0053] (2) 本发明利用基于小波核的双重加权 SVM 分类模型对蓝藻 (微囊藻) 空间分布信息进行提取和识别, 既能有效地识别各种密度的蓝藻空间分布信息, 又可以减少非监督分类误分的水体信息, 能够准确识别蓝藻覆盖区, 包括较低浓度蓝藻覆盖区, 满足环境调查的要求, 有效地解决了浑浊水体背景对蓝藻 (微囊藻) 提取的影响;

[0054] (3) 本发明将校正后的遥感影像数据、GIS 数据与处理分析后的各类实地观测数据叠加后进行可视化显示, 能够更加直观地观察蓝藻爆发过程生物量的动态迁移过程, 有利于对蓝藻水华的成因和分布变化做出分析和判断, 从而为蓝藻水华预报和控制提供了可靠的依据。

附图说明

[0055] 图 1 为蓝藻区域典型地物的影像光谱曲线示意图;

- [0056] 图 2 为本发明的 NDI_CB 与传统的 NDVI、RVI 的 k-均值非监督分类处理结果示意图；
- [0057] 图 3 为本发明基于 VPRS_GID 的特征优化模型流程图；
- [0058] 图 4 为基于小波核的双重加权 SVM 模型流程图；
- [0059] 图 5 为蓝藻生物量可视化信息层组织结构示意图；
- [0060] 图 6 为蓝藻生物量动态可视化仿真结果示意图。
- [0061] 图 2 中，(2a) 为 NDI_CB 的分类图；(2b) 为 NDVI 的分类图；(2c) 为 RVI 的分类图；
- [0062] 图 6 中，(6a) 为 2009 年 6 月 13 日的蓝藻生物量仿真示意图；(6b) 为 2009 年 6 月 29 日的蓝藻生物量仿真示意图；(6c) 为 2009 年 7 月 13 日的蓝藻生物量仿真示意图。

具体实施方式

[0063] 下面结合附图和具体实施例对本发明进行详细说明。本实施例以本发明技术方案为前提进行实施，给出了详细的实施方式和具体的操作过程，但本发明的保护范围不限于下述的实施例。

[0064] 实施例 1

[0065] 一种基于遥感影像的蓝藻生物量时空变化监测与可视化方法，该方法包括以下步骤：

[0066] 1) 获取蓝藻区域的遥感影像，并对其进行图像预处理，构建归一化蓝藻指数；

[0067] 2) 采用基于 VPRS_GID 的特征优化模型对遥感影像的特征进行优化，获得优化后的多特征空间；

[0068] 3) 根据多特征空间建立基于小波核的双重加权 SVM 分类模型，获取 SVM 最优分类决策面，并利用该分类模型对蓝藻（微囊藻）空间分布信息进行提取识别和变化检测，并结合实地调查数据进行综合验证和精度分析；

[0069] 4) 将处理后的遥感影像、研究区域的 GIS 矢量数据与该区域的实地观测数据进行叠加显示，从而实现蓝藻水华爆发时空变化过程和规律的仿真模拟。

[0070] 1、遥感影像预处理

[0071] 图像预处理包括对遥感影像进行大气层顶表观反射率反演，即将数字图像中的 DN(Digital Number) 值转化为反射率，以近似取代需要大气校正之后得到的地表反射率。

[0072] 2、基于影像光谱分析的归一化蓝藻指数 (NDI_CB) 构建

[0073] 归一化蓝藻指数(normalized difference cyanobacteria bloom index, NDI_CB) 的构建是基于影像光谱分析，根据实地情况，选取微囊藻、浑浊水体和陆生植被训练样本，影像光谱曲线（第 1-5 和 7 波段），如图 1 所示。

[0074] 综合考虑波段 3、4 和 5 三个波段发现，从波段 3 到波段 4 蓝藻水华的影像亮度值逐渐提高，浑浊水体则恰好相反；从波段 4 到波段 5 蓝藻水华的亮度值的下降幅度则明显高于浑浊水体。如果将波段 4 与波段 3，以及波段 5 与波段 4 分别进行差值处理，那么发现蓝藻水华和浑浊水体的趋势截然相反，并且当蓝藻浓度很低时，这种趋势表现得依然显著。

[0075] 根据以上对微囊藻覆盖区和其它典型地物的光谱特征分析，综合考虑波段 3、4 和 5 三个波段的谱间关系，构造了归一化蓝藻指数，其方法如下：

[0076] ①分别对波段 4 与波段 3，以及波段 5 与波段 4 进行差值运算。

$$[0077] \quad \begin{cases} a = B4 - B3 \\ b = B5 - B4 \end{cases} \quad (1)$$

[0078] 式中, B3、B4 和 B5 分别表示波段 3、4 和 5 的光谱值 (亮度值或是反射率)。

[0079] 由于式 (1) 得到的差值 (a 和 b) 往往跨越正负值域, 为了保证下面的归一化处理能够得到 $[-1, 1]$ 之间的分布, 并且不改变原来直方图的形状, 取 a 和 b 中的最小值, 分别将 a 和 b 向上平移最小值个单位, 使 a 和 b 全部分布在正值区域。

$$[0080] \quad \begin{cases} c = \min(a, b) \\ a' = a + |c| = a + |\min(a, b)| \\ b' = b + |c| = b + |\min(a, b)| \end{cases} \quad (2)$$

[0081] ②通过比值运算, 构建 NDI_CB :

$$[0082] \quad NDI_CB = \frac{a'/b' - 1}{a'/b' + 1} = \frac{a' - b'}{a' + b'} \quad (3)$$

[0083] 对传统的 NDVI、RVI 和 NDI_CB 三种植被指数分别进行 k- 均值非监督分类处理, 如图 2 所示, 经过实验可以发现, 构建的 NDI_CB 较传统的 NDVI 和 RVI 能更加适用于提取低浓度微囊藻空间分布信息。

[0084] 3、多特征空间与优化

[0085] 如图 3 所示, 所述的步骤 2) 具体包括:

[0086] 21) 根据遥感影像获取遥感影像的特征信息, 采用变精度粗糙集 (VPRS) 模型对特征信息进行约简;

[0087] 22) 对约简得到的特征集采用灰色关联决策 (GID) 进行最佳特征组合选取, 获得优化后的多特征空间, 包括影像波段组合、归一化蓝藻指数、最佳指数特征组合和最佳 Gabor 滤波器纹理特征组合, 具体为:

[0088] ①影像波段组合: 对影像光谱分析, 发现 3、4、5 波段有利于微囊藻提取, 因此选择这三个波段的组合;

[0089] ②归一化蓝藻指数 (NDI_CB);

[0090] ③最佳指数特征组合: 建立适于微囊藻或水体提取的归一化差值植被指数 (NDVI)、比值植被指数 (RVI)、归一化差异水体指数 (NDWI)、改进型归一化差异水体指数 (MNDWI) 和新型水体指数 (NWI) 等 5 种指数特征。然后利用灰色关联决策 (GID) 方法选择最佳特征组合 (组合数为 3);

[0091] ④最佳 Gabor 滤波器纹理特征组合: 选用 2、3 两个尺度, 大小为 3×3 的窗口和 0° 、 45° 、 90° 、 135° 四个角度, 分别对影像 3、4、5 三个波段进行 Gabor 滤波器响应统计, 共 24 个频谱纹理特征。然后基于 VPRS_GID 特征优化模型选择出最佳特征组合 (组合数为 3)。

[0092] 4、基于小波核的双重加权支持向量机分类模型

[0093] 由于 Morlet 小波核函数计算简单、构造方便, 并经过证明得到其符合平移不变核函数的充分必要条件, 因此本实施例选择 Morlet 小波核函数, 如下式 (4) 所示来构建分类模型中的支持向量机:

$$[0094] \quad K(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y}) = \prod_{i=1}^d \left[\left(\frac{\cos 1.75(x_i - y_i)}{a_i} \right) \cdot \exp \left(-\frac{\|x_i - y_i\|^2}{2a_i^2} \right) \right] \quad (4)$$

[0095] 加权能进一步提高 SVM 算法的泛化能力和分类性能,目前对传统 SVM 进行加权的方式主要有三种:①对不同类别赋以权重;②对每个样本赋以权值;③对每个特征属性赋以权重。本实施例选择第①和第③两种方式对 SVM 进行加权,提出了一种类别样本数差异与特征重要性差异相结合的双重加权方法。

[0096] 基于类别样本数差异加权的支持向量机优化问题可描述为:

[0097]

$$\begin{cases} \min \Phi(\mathbf{w}) = \frac{1}{2}(\mathbf{w} \cdot \mathbf{w}) + C\lambda_{y_i} \left(\sum_{i=1}^l \xi_i \right) \\ \text{S.T.} \quad y_i[(\mathbf{w} \cdot \varphi(x_i)) + b] \geq 1 - \xi_i, i = 1, 2, \dots, l \end{cases} \quad (5)$$

[0098] 上式中, λ_{y_i} 为类别权重,由下式 (6) 给定:

$$[0099] \quad \lambda_{y_i} = \begin{cases} 1 - \frac{l_1}{l}, y_i = 1 \\ \dots \\ 1 - \frac{l_n}{l}, y_i = n \end{cases} \quad (6)$$

[0100] 式中, l_1 为类别 y_1 的样本数, l 为样本总数, n 为类别数。

[0101] 对于数量大的类别,通过赋予小的权值,使其错分率上界增大;而对于数量小的类别,通过赋予大的权值,使其错分率上界下降,从而起到平衡作用。

[0102] 在特征加权中权重向量 ω 的求取是关键。特征权重的计算通常是根据某种度量方法来度量特征的重要性从而量化特征与给定类别之间的相关性。本发明采用基于信息增益(比)率的评价函数进行特征相关分析。

[0103] 设 T 为一个包含 t 个数据样本的集合,有 m 个不同的类别 $C_i, i \in \{1, 2, 3, \dots, m\}$ 。一个特征 A 具有 v 个不同的值 $\{a_1, a_2, \dots, a_v\}$,则利用 A 可以将集合 T 划分为 v 个子集 $\{T_1, T_2, \dots, T_v\}$,其中 T_j 包含了 T 集合中特征 A 取 a_j 值的数据样本; T_{ij} 为子集 T_j 中属于 C_i 类别的样本集。根据 A 划分 T 集合所需要的信息熵为:

$$[0104] \quad E(A) = \sum_{j=1}^v \frac{T_{1j} + T_{2j} + \dots + T_{mj}}{T} I(T_{1j}, T_{2j}, \dots, T_{mj}) \quad (7)$$

$$[0105] \quad = - \sum_{j=1}^v \sum_{i=1}^m \frac{T_{ij} + T_{2j} + \dots + T_{mj}}{T} p_{ij} \log(p_{ij})$$

[0106] 其中, $p_{ij} = \frac{T_{ij}}{|T_j|}$ 为子集 T_j 中任一个数据样本属于类别 C_i 的概率。

[0107] 特征 A 中包含的信息熵为:

[0108]

$$SplitInfo(A) = - \sum_{i=1}^v p(a_i) \log(p(a_i)) \quad (8)$$

[0109] 利用 A 对 T 集合划分所获得的信息增益是:

$$[0110] \quad \text{Gain}(A) = I(t_1, t_2, \dots, t_m) - E(A) \quad (9)$$

[0111] 信息增益（比）率函数为：

[0112]

$$\text{GainRatio}(A) = \frac{\text{Gain}(A)}{\text{SplitInfo}(A)} \quad (10)$$

[0113] 从而采用向量 $\text{GR} = (\sqrt{\text{GainRatio}(\text{feature}_1)}, \dots, \sqrt{\text{GainRatio}(\text{feature}_n)})$ 来构造特征权重向量 ω , feature_i 为特征标号, $i = 1, 2, \dots, n$.

[0114] 本发明提出的模型将小波核函数与双重加权方式相结合, 既利用了小波分析在影像细节方面具有的良好表现能力, 又平衡了不同样本类别分类的正确率, 并能有效避免被弱相关或不相关特征所支配。基于小波核的双重加权 SVM 分类模型可描述为：

$$[0115] \quad \left\{ \begin{array}{l} \max : L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i^T P, x_j^T P) \\ = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \prod_{n=1}^d \left[\frac{\cos 1.75 \left(\frac{(x_i^T P)_n - (x_j^T P)_n}{a_n} \right)}{a_n} \right] \cdot \exp \left(- \frac{\| (x_i^T P)_n - (x_j^T P)_n \|^2}{2a_i^2} \right) \\ \text{S.T.} \quad C\lambda_{y_i} \geq \alpha_i \geq 0, i=1, 2, \dots, l \\ \sum_{i=1}^l \alpha_i y_i = 0 \end{array} \right.$$

[0116] 式中, α_i 为拉格朗日乘子, x_i 表示输入的度量指标分类特征向量, y_i 表示输出的学习结果, C 为常量, 是对越界样本点施加的惩罚系数, K 为特征加权核函数。

[0117] 如图 4 所示, 其实现步骤具体为：

[0118] 31) 根据遥感影像的样本集, 获得类别权重 λ_{y_i} ；

[0119] 32) 构造特征权重向量 ω 和线性变换对角阵 P , 对多特征空间进行线性变换, 实现特征重要性加权；

[0120] 33) 对特征加权后的多特征空间计算小波核函数, 构建基于小波核的双重加权 SVM 分类模型。

[0121] 5、蓝藻生物量变化的可视化仿真

[0122] 以校正后的遥感影像为底层数据源, 对蓝藻实地观测数据进行插值并实现直观显示的数据处理和可视化仿真系统。可实现遥感影像和蓝藻观测数据的空间配准、对离散的观测点数据进行插值、利用插值数据生成等值线并进行可视化显示和交互操作。实地观测数据包括矢量地图数据和地面蓝藻生物量采样数据。

[0123] 可视化显示包括以下步骤：

[0124] 41) 读取实地观测数据, 并对遥感影像分类数据和矢量地图数据进行空间坐标配准；

[0125] 42) 在遥感影像数据层上叠加具有空间坐标的实地观测数据；

[0126] 43) 对离散的实地观测数据进行插值处理, 利用插值数据拟合出生物量等值线；

[0127] 44) 以遥感影像数据为底地图, 在其上依次叠加矢量地图数据、生物量等值线信息, 然后对处理结果及其对应日期的气象观测数据（风力、风向、温度等）进行可视化显示

和仿真模拟。

[0128] 所述的坐标配准是指将以 WGS-84 坐标系统的 3 度带为基准转换为归一化的坐标系统,具体为:

[0129] a) 获得待转换数据的坐标范围,即该数据在 x 轴方向的最小值 x_{\min} 、最大值 x_{\max} 和 y 轴方向的最小值 y_{\min} 、最大值 y_{\max} ;

[0130] b) 通过该坐标范围计算转换所需的尺度因子 Scale,以 x 轴方向为例,其计算公式如下:

$$[0131] \quad XScale = \left[-\frac{x_{\min}}{x_{\max} - x_{\min}}, \frac{1}{x_{\max} - x_{\min}} \right];$$

[0132] c) 根据所得尺度因子将空间坐标转换为归一化坐标,以 x 轴方向为例,计算公式如下:

$$[0133] \quad X_{\text{Conv}} = XScale_0 + XScale_1 * x_0$$

[0134] 其中, x_0 为原始空间坐标, X_{Conv} 为转换后的归一化坐标; $XScale_0$ 和 $XScale_1$ 分别为尺度因子 XScale 的第一和第二个元素。

[0135] 如图 5 所示,可视化处理的具体技术细节及实施步骤如下:

[0136] (1) 系统所需数据及其预处理

[0137] ① 矢量地图数据准备与预处理

[0138] 主要包括:与蓝藻实地观测区域相一致的 Landsat7 ETM 影像、研究区域及周边的最新矢量数据(Shape 文件格式,包括河网、道路、湖泊等图层及相应注记)。

[0139] 对于遥感影像数据,采用 UTM 投影,坐标基准为 WGS-84 坐标系统中的 3 度带第 49 区,并将所有数据转换为带地理坐标的 GeoTiff 格式,在显示时利用第 5、4、3 波段实现假彩色合成。

[0140] 对于矢量地图数据,同样采用 UTM 投影,坐标基准为 WGS-84 坐标系统中的 3 度带第 49 区,所有矢量数据均以 Shape 文件格式进行保存。其中,河网和道路数据采用线类型 Shape 文件进行表达,湖泊采用多边形类型 Shape 文件进行表达。

[0141] ② 地面采样数据准备与预处理

[0142] 将地面采样数据按日期分别保存于独立的 Excel 文件中,其中,表格行代表 22 个采样点名称,表格列依次表示生物量观测值、观测时间、水温、气温、主导风向、水流速度。

[0143] (2) 空间数据读取

[0144] 通过遥感栅格数据读取和矢量数据(Shape 格式)读取两大部分。

[0145] 栅格数据读取,将已经过辐射校正的遥感影像(带地理坐标的 GeoTiff)的 3 个波段:

[0146] 蓝光波段(Blue Band)、红光波段(Red Band)和近红外波段(NIR Band),保存在与图像大小相一致的矩阵中,并将该影像所对应的地理坐标和投影信息单独保存。矢量数据读取,将经过预处理的河网、道路和湖泊数据中的点位坐标和属性信息保存在相应的结构体中。

[0147] (3) 空间数据配准

[0148] 将遥感影像和矢量数据中的空间坐标统一由 WGS-84 坐标系统的 3 度带为基准转换为归一化的坐标系统(Normalized Coordinate System)。

[0149] (4) 生物量观测数据读取和管理

[0150] 将保存在 Excel 文件中的生物量观测数据导入到系统内部,并对不同数据类别进行分类存储和优化以提高读取和显示效率。

[0151] (5) 生物量等值线生成

[0152] 利用已导入的生物量观测数据,在数值分析的基础上,利用插值运算将二维空间中离散分布的观测点数据生成连续的生物量等值线数据。

[0153] (6) 等值线数据与空间数据的坐标配准

[0154] 将等值线数据中数据点的空间坐标由 WGS-84 坐标系统的 3 度带为基准转换为归一化的坐标系统 (Normalized Coordinate System)。

[0155] (7) 等值线数据、多媒体数据与空间信息的叠加显示

[0156] 利用 IDL 中面向对象的图像显示技术,以遥感影像为底地图,在其上依次叠加矢量数据 (河网、道路和湖泊数据)、生物量等值线。

[0157] (8) 多时相观测数据的选择和对比操作

[0158] 利用表格组件对已有观测数据文件进行管理和操作,将选定日期的生物量观测数据以等值线形式进行表达。利用滑动条组件动态依次显示多个时相的生物量等值线数据及其与之相对应的多媒体信息。

[0159] (9) 风力风向数据读取和显示

[0160] 按日期顺序依次读入某年全年的风力风向观测数据,其内容包括每日 4 个不同时间风力和风向观测值。其中,风力数据按蒲福风级进行表示。

[0161] 在进行显示时,将当前蓝藻观测数据和环境信息与风力风向观测信息相链接。在显示当日蓝藻观测数据和环境信息时,搜索观测前一日、当前日、以及后一日共三天的风力和风向观测数据。并以矢量箭头方向与横轴夹角表示风向,矢量长度表示风力等级。对于缺失的风力和风向观测数据,规定以长度为 0 且与横轴方向夹角为 0° 的矢量箭头符号予以表示。并将观测日当天出现频次最高的风向作为该日主导风向;主导风向的风力值取该风向对应风力的均值并进一步取整。

[0162] (10) 信息层的组织

[0163] 图 5 为该系统的信息层组织图。该图表达了遥感影像、水网矢量数据、蓝藻生物量观测数据、风力、风向、温度等气象观测信息与蓝藻生物量在不同时间动态变化的状态信息之间的组织关系。

[0164] (11) 蓝藻漂移分析的可视化显示

[0165] ①空间坐标与归一化的坐标系统 (Normalized Coordinate System) 的转换方法

[0166] 在数据准备和预处理阶段,所有的空间信息和多媒体信息都包含有地理空间坐标。在系统内部对输入数据进行分析处理后,会产生一系列与之对应的运算结果 (如蓝藻生物量等值线),为了使分析和原酸结果和正确叠加与空间信息之上使分析结果及原始信息能够正确叠加于与空间信息之上,并与多媒体信息协同显示,就必须保证各数据层之间具有相一致的坐标系统。为统一起见,本系统将所有信息层的地理空间坐标转化为归一化坐标 (Normal Coordinates)。在归一化坐标系统 (Normalized Coordinate System) 中, x, y, z 三轴的坐标范围均介于 0 和 1 之间。

[0167] ②蓝藻生物量观测数据插值方法

[0168] 蓝藻生物量观测数据插值是利用克里金 (Kriging) 插值法完成的。该插值法又称空间自协方差最佳插值法, 该算法的基本思想为:

[0169] 通过考虑已知数据点属性值在空间位置上的变异分布, 以此来确定已知点对待插值点的影响范围, 然后用此范围内的采样点来估计待插值的属性值。该方法是一种最佳线性无偏估计方法, 是在考虑了已知数据点相互间的空间位置等几何特征和空间结构之后, 为达到线性、无偏和最小估计方差的估计, 而对每一个已知点赋与一定的系数, 最后再进行加权平均来估计块段品位的方法。该方法可得到光滑的内插结果。其最大优势就在于能够根据已知点的数值及空间分布, 模拟出数据分布的趋势。这一点, 在对蓝藻水华受风力等气象环境因素影响下进行时空漂移模拟和可视化是非常重要的。

[0170] ③蓝藻生物量等值线生成方法

[0171] 通过数值插值, 可得到研究区域平面内所有点的蓝藻生物量数据。在规定的精度 (等值线间隔) 范围内搜索数值最为接近的格网点, 即可构成蓝藻生物量等值线。在此基础上, 依据蓝藻生物量与蓝藻水华表现现象之间的联系, 将蓝藻水华严重程度由弱至强分为 5 级; 利用灰度密度分割的方法, 将一个确定的生物量数值区间与代表某一蓝藻水华严重程度的固定颜色值之间建立对应关系, 生成生物量与色彩索引查找表。

[0172] ④蓝藻生物量等值线可视化方法

[0173] 在显示生物量等值线之前, 依据 JPEG 格式协议, 按照已建颜色查找表的对应关系, 以生物量数值为基准, 生成 JPG 格式的 RGB 真彩色图像文件。并且, 在此过程中, 必须同时生成 ALPHA 透明层, 以确保等值线图像超出研究区域外的部分保持透明。

[0174] 在显示时, 将该图像文件以设备归一化坐标 (即 x, y 坐标值区间均为 $[-1, 1]$) 为基准进行空间坐标转换, 与遥感影像实现多个图层的叠加显示。

[0175] 本发明通过构建归一化蓝藻指数 (NDI_CB), 在优化特征空间的基础上, 利用小波核双重加权 SVM 分类模型对蓝藻 (微囊藻) 空间分布信息进行提取和识别, 可有效提高基于遥感影像的微囊藻识别精度。另外, 本发明也利用了实地监测的流域风力风向与温度等气象数据, 综合验证与分析了遥感变化检测的结果和精度。并进一步将校正后的遥感影像、GIS 数据与处理分析后的地面实测数据、风力风向、温度等气象信息与多媒体信息结合, 实现了蓝藻爆发和迁移过程的定量、可视化分析与时空变化仿真。该过程更加直观而真实地反映了水体中的蓝藻生物量变化与流域风力风向等气象因素变化之间的关系。从而能够对蓝藻形成和爆发机理及过程进行更全面的观察和分析。

[0176] 实施例 2

[0177] 以下采用实施例 1 中的基于遥感影像的蓝藻生物量时空变化监测与可视化方法进行实际操作。

[0178] 1、蓝藻 (微囊藻) 空间分布提取和变化检测结果

[0179] 选取 2000 年 8 月 1 日和 9 月 2 日获取的 Landsat7ETM+ 影像。利用小波核双重加权 SVM 分类模型, 得到淀山湖区域微囊藻空间分布格局的识别与检测结果。

[0180] 在淀山湖沿岸设置 10 个采样区, 对蓝藻生物量进行实地采样。发现遥感影像的提取结果与实际调查数据相符合。并选择微囊藻较为明显的 2000 年 8 月 1 日影像进行误差矩阵统计。

[0181] 表 1 小波核双重加权 SVM 分类技术对蓝藻 (微囊藻) 信息提取的误差矩阵

	微囊藻覆盖区	湖泊水体	使用者精度
[0182] 微囊藻覆盖区	101	17	85.59%
湖泊水体	19	483	96.22%
生产者精度	84.17%	96.60%	
总体分类精度: 94.1935%		Kappa 系数: 0.8128	

[0183] 从上表可知,基于小波核双重加权 SVM 分类技术能够有效地利用多特征空间数据,准确地识别出蓝藻(微囊藻)爆发区。

[0184] 2、生态学调查与遥感解译分析对比

[0185] 调查的采样区要尽量覆盖到全湖大多数区域,但是考虑到采样的可实施性,最终沿湖选择了 19 个采样区。调查重点时间段为 2009 年淀山湖微囊藻水华暴发期(6-9 月)。由于微囊藻随着时间和温度上升在不断增殖,又随着风向等气象因子不断发生位置变化,因此选择风力较小且晴天的日子进行研究区微囊藻的生态学调查。

[0186] 从结果可知,遥感影像的提取结果与实际调查数据相符合,1、4、5 分布区域在影像提取结果中得到了很好的表达,而 2、3 区域是因为截取淀山湖时,所用的矢量图有误差,特别是在 3 区域处有面积缩减,所以未能得到表现。此外,由于实地调查的工作量大、周期长,不可能在短时间内走遍所有采样区,因而遥感影像的微囊藻提取能够弥补缺乏采样数据的区域,比如在西北处几个未能到达的采样区。

[0187] 3、蓝藻生物量变化可视化仿真结果

[0188] 将 2009 年在若干不同日期对淀山湖蓝藻实地观测的数据与相近时相的遥感数据和矢量数据相结合,对所观测的蓝藻分布情况和气象数据实现动态可视化仿真,结果如图 6 所示。

[0189] 本发明将高分辨率的多时相遥感影像数据、GIS 数据与相应的湖泊及其周边流域内水网的各种气象数据(风力、风向、温度等)相叠加,同时与处理分析后的地面实测数据相结合显示,能够对蓝藻形成和爆发的规律及过程进行更全面的观察和分析,更加直观而真实地反映水体中的蓝藻生物量的变化,从而实现在蓝藻水华爆发的预测预警的过程定量分析,准确地对蓝藻时空变化过程和规律进行仿真模拟。动态可视化结果显示,蓝藻生物量的时空格局变化,受到风力、风向以及周边水文环境的综合影响,尤其风力和风向对蓝藻漂移趋势产生了较大的决定作用。

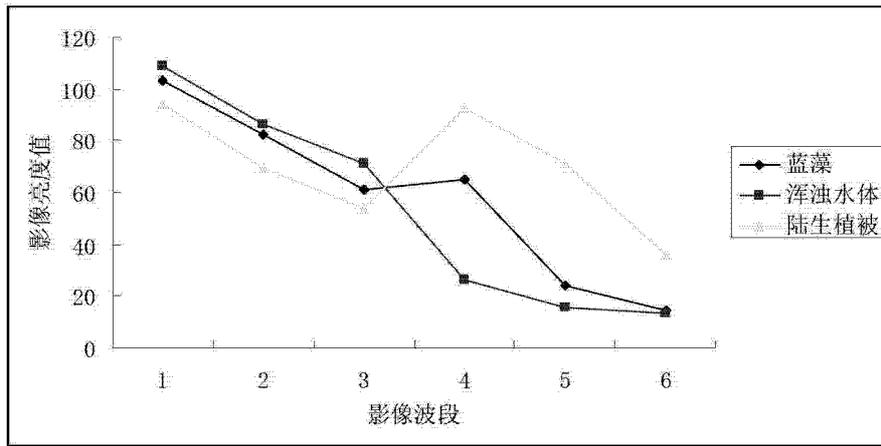


图 1

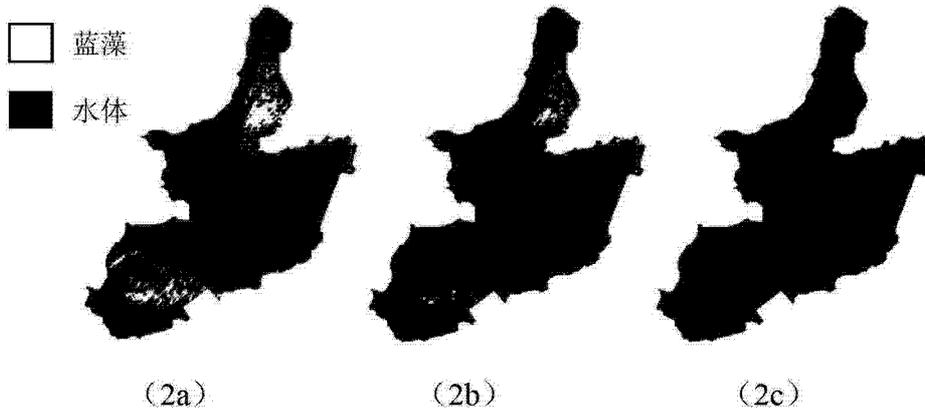


图 2

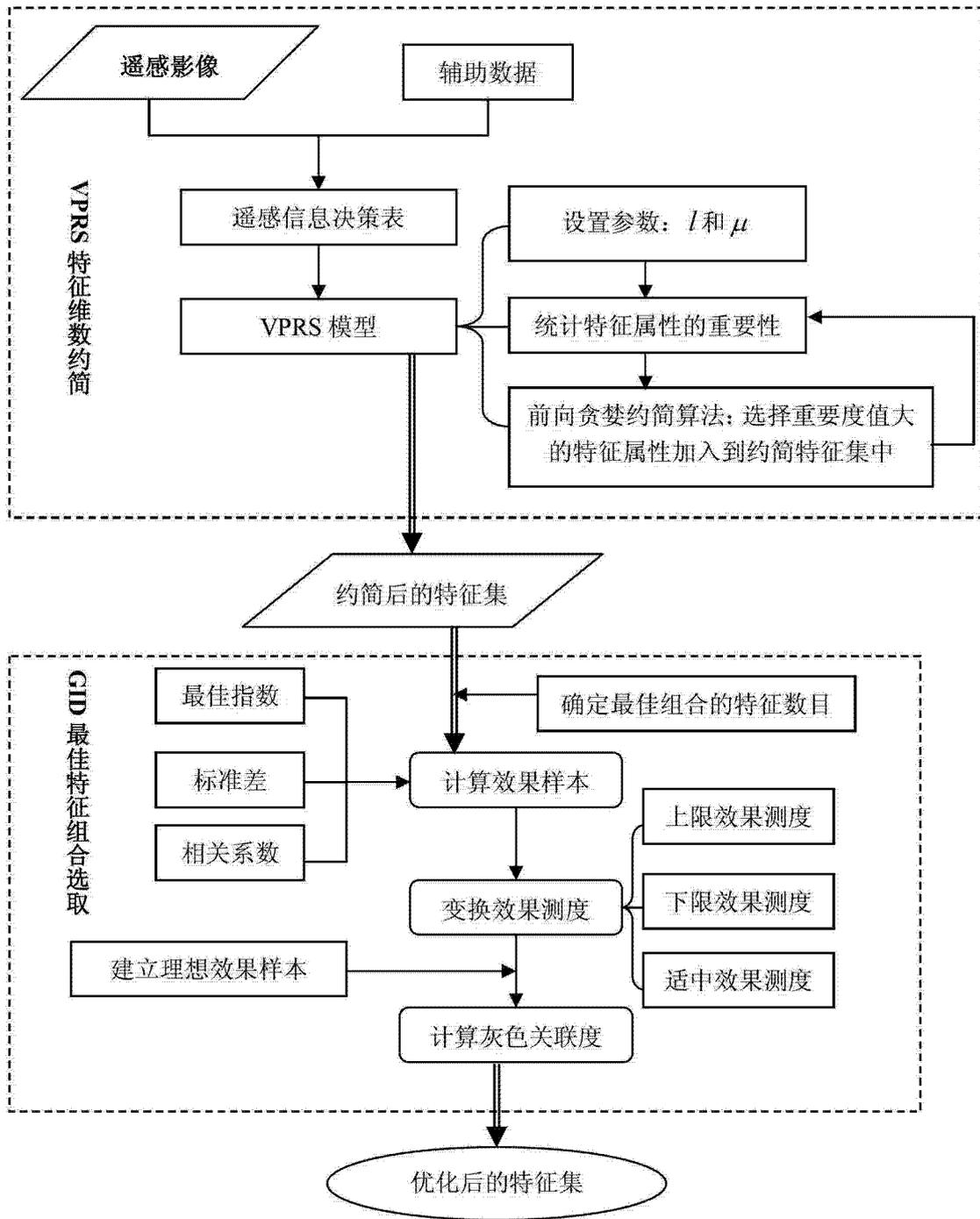


图 3

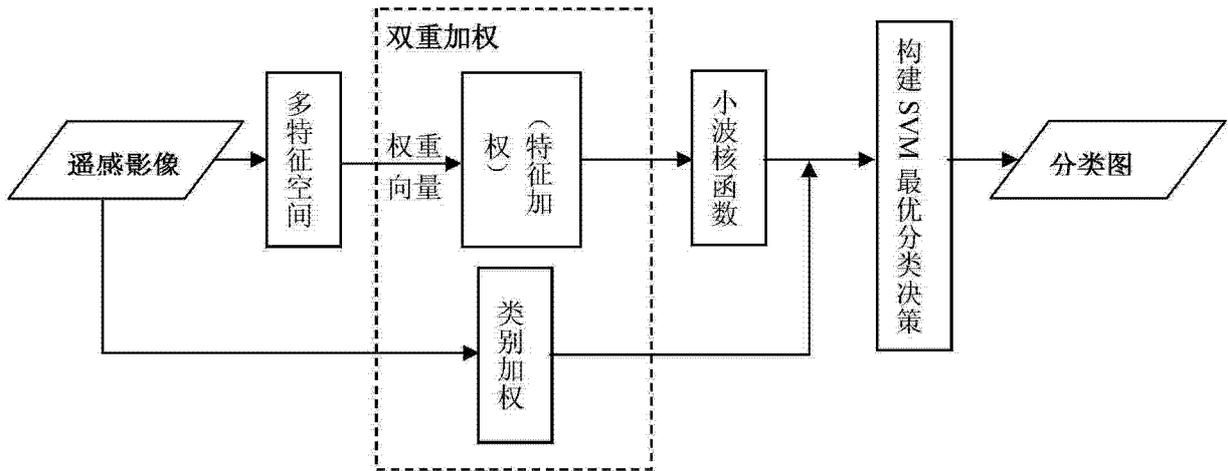


图 4

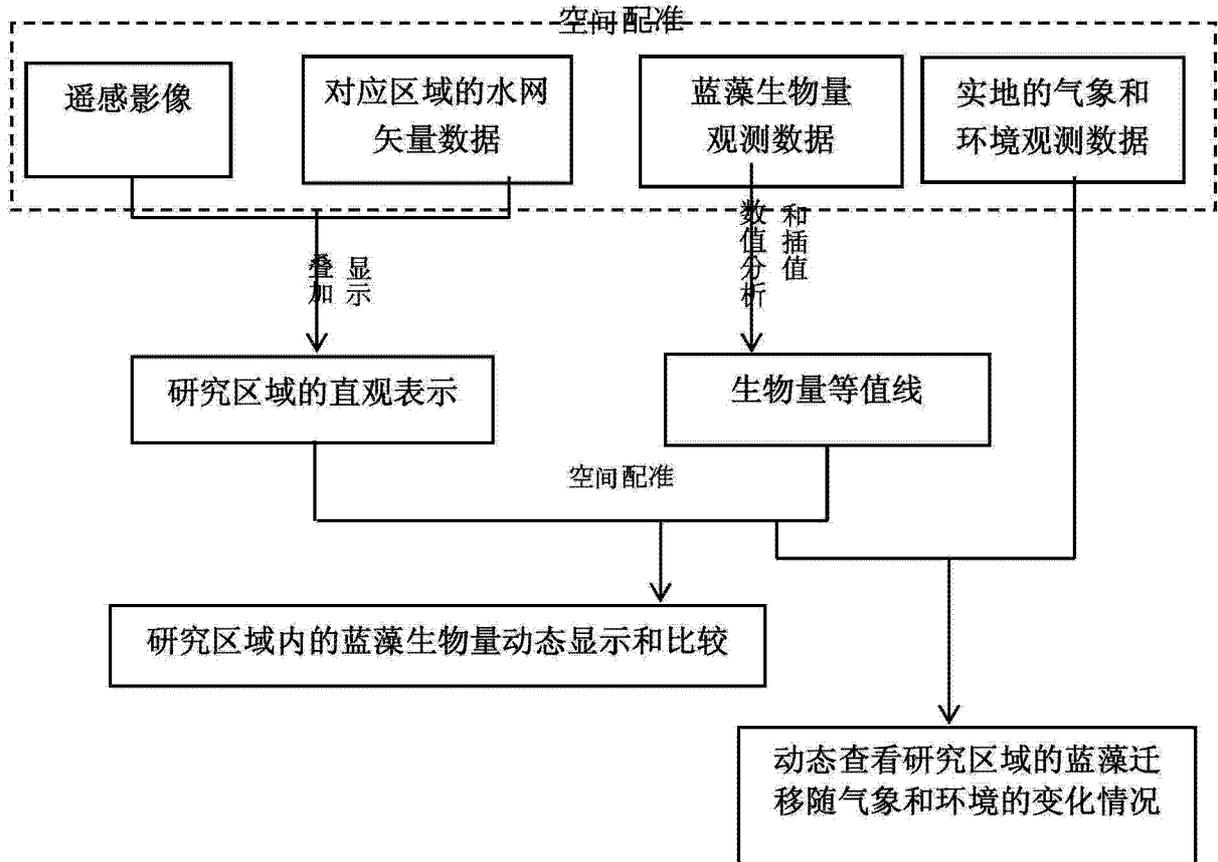
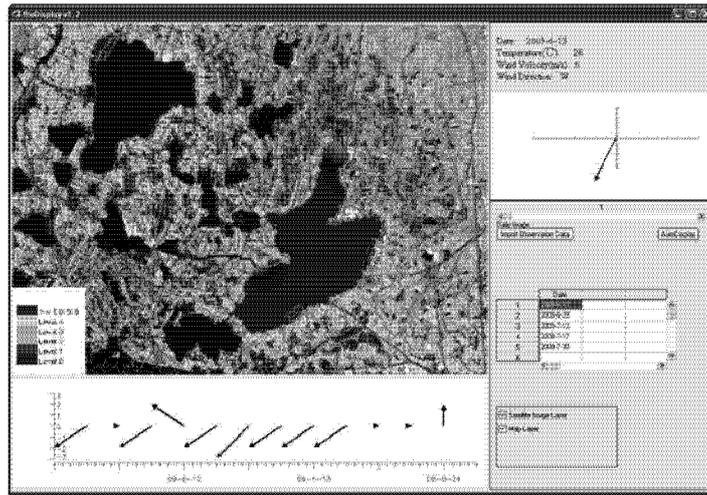
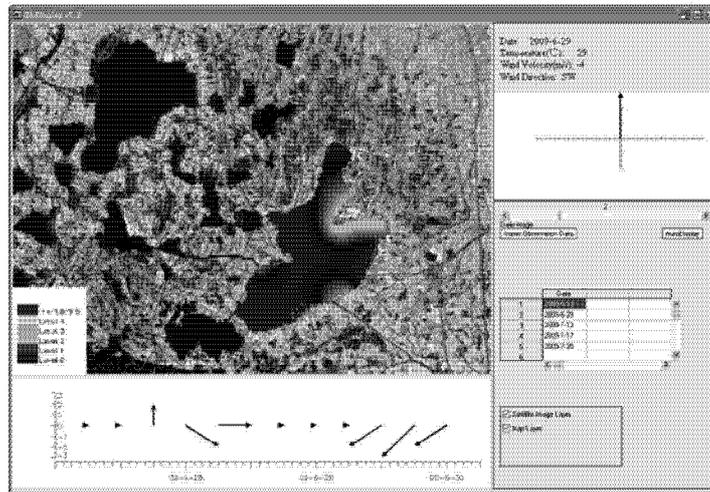


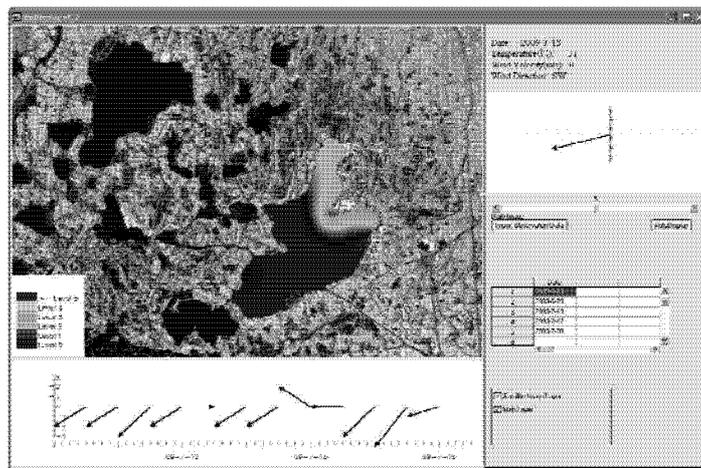
图 5



(6a)



(6b)



(6c)

图 6