

形状特征的卫星云图检索方法研究

夏士明,李 骞,谢凯翔

(解放军理工大学 气象海洋学院,江苏 南京 211101)

摘要:卫星云图的相似性检索可以帮助预报人员较快地定位历史相似天气状况,提高天气预报的可靠性。文中尝试采用 Zernike 矩和网格法来提取云图中云的形状区域特征,并对卫星云图进行相似性检索。其过程为:首先对云图进行云地分离的预处理,然后提取云的形状区域特征,并在此基础上生成云图库所对应的特征向量库,接着进行相似度的距离计算和相似性匹配,最后根据距离的排序顺序输出最终的检索结果。通过比较两种方法的检索效果,提出了一种改进的形状特征提取方法—网格内切圆法。实验表明该方法具有以上两种方法的优点,能够得到较好的检索效果,可以有效检索出视觉上相似的云图。

关键词:卫星云图;特征向量;相似性检索;旋转不变性

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2014)04-0060-05

doi:10.3969/j.issn.1673-629X.2014.04.015

A Method of Satellite Cloud Image Retrieval of Shape Feature

XIA Shi-ming, LI Qian, XIE Kai-xiang

(Institute of Meteorological Oceanographic, PLA University of Science and Technology, Nanjing 211101, China)

Abstract: Satellite cloud image retrieval can help forecasters position similar weather in the past quickly and improve the reliability of the weather forecast. It attempts to use the Zernike moment and grid based method to extract the cloud shape regional characteristics, making similar retrieval for satellite cloud image. The retrieval process is: first of all, carry out the pretreatment of separation on the cloud image, extract the effective characteristics and generate characteristics of the cloud atlas library. Then carry out similarity measurement and similarity matching between a given cloud image and the cloud atlas. Finally, output the final result according to the order. By comparing retrieval results of the two methods, put forward an improved algorithm called grid based inscribed circle method. Experiments show that this method owns the advantages of the above two methods and can retrieve cloud images with similar visual characteristics from the history cloud image atlas.

Key words: satellite cloud image; feature vector; similarity retrieval; rotation invariance

0 引言

卫星云图在气象业务领域中发挥着重要作用,由于卫星云图具有高的时、空分辨率,能连续追踪云系的形成、天气系统发展加强与降水等的相互关系,可以发现大尺度云系分布的各种云型特征,为预报员准确预报天气提供了依据^[1]。目前,大量卫星云图堆积在数据库中没能得到有效利用^[2],可以采用图像检索的方法对历史卫星云图库中的云图进行相似性检索,这样可以借鉴过去的天气状况对当前的天气进行辅助预测,能提高预报的准确性。经过分析发现,气象卫星云图在形状方面表现出较强的区分度,因此可以用形状特征对卫星云图进行相似性检索。

在卫星云图上,云图的形状是指目标物对光的不同强弱的反射或其辐射的发射所形成的不同明暗程度物像点的分布式样,即表现为一定的形状结构^[3],云的形状结构有助于识别云的种类和了解云的形成过程。对于卫星云图,由于提取完整的轮廓比较困难,且轮廓使用的信息相比区域要少很多^[4],而基于区域的表示方法将区域当作一个整体看待,有效利用了区域内的所有像素,因而受噪声和形状变化的影响相对较小^[5],因此文中只采用区域的特征进行相似性检索。

1 系统框架

为了利用形状区域特征对卫星云图进行检索,文

收稿日期:2013-06-20

修回日期:2013-09-25

网络出版时间:2014-01-28

基金项目:国家自然科学基金面上项目(41174164)

作者简介:夏士明(1990-),男,硕士研究生,研究方向为图像处理、模式识别;李 骞,讲师,研究方向为计算机视觉。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140128.1201.061.html>

中方法提出的系统框架如图1所示,主要分为离线特征库生成和在线检索两个部分。在离线特征库生成阶段,首先对云图进行云地分离的预处理,即先进行直方图均衡化处理来改善云图的对比度,再进行半阈值化^[6](高于阈值的像素不变,低于阈值的像素变为0像素)的处理,对所分离的云图像进一步提取其形状区域特征,并结合云图文件的相应描述加入特征库,以供在线检索使用;在在线检索阶段,对用户输入的待检索云图进行云地分离的预处理,进一步提取该云图相应的形状区域特征值,用该特征值与特征库中存储的形状特征值进行距离计算,根据向量距离大小进行排序,将排序结果输出即可得到检索结果。

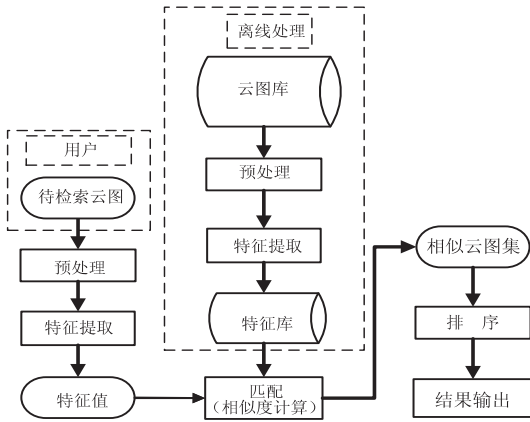


图1 检索框架图

2 形状区域特征的提取

形状的区域特征主要有区域的面积、欧拉数、离散度、偏心率、几何不变矩、Zernike 矩、Legendre 矩、网格法、角半径变换等^[7]。文中主要尝试采用 Zernike 矩和网格法对卫星云图进行相似性检索。

2.1 区域 Zernike 矩特征的提取

Zernike 矩是基于 Zernike 多项式的正交化函数^[8],所利用的正交多项式集是一个在单位圆内的完备正交集。当计算一幅图像的 Zernike 矩时,以该图像的形心(也称作重心)为原点^[9],把像素坐标映射到单位圆内,Zernike 矩是复数矩,一般把 Zernike 矩的模作为特征来描述物体形状,Zernike 矩的定义^[10]为

$$Z_{nm} = \frac{n+1}{\pi} \sum_x \sum_y V_{nm}^*(x,y) f(x,y), x^2 + y^2 \leq 1 \quad (1)$$

式中, $V_{nm}^*(x,y) = V_{nm}(\rho \cos \theta, \rho \sin \theta) = R_{nm}(\rho) \exp(jm\theta)$, $R_{nm}(\rho)$ 的定义为

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \times \rho^{n-2s} \quad (2)$$

式中, n 和 m 为非负整数,并且必须满足 $n - |m|$ 为偶数且 $n \geq |m|$ 。取满足条件的不同 n 和 m 的值,就可以构成一个特征向量。

2.2 网格特征的提取

网格法的定义为,将给定的云图等分成一定数目的规整网格^[11],文中分成 8×8 的网格,求出每个网格的均值 μ_i 和全局的均值 μ ,当网格的均值 μ_i 大于全局的均值 μ 时,这个网格为 1,否则为 0,这样每幅云图就对应一个 64 位的 0、1 字符串,作为特征值。用网格法,不具有旋转不变性,但具有尺度不变性,且受噪声的影响较小。不难理解,同样一幅图,旋转 90 度之后的网格特征值一般都会发生很大变动,因此网格法不具有旋转不变性。尺度不变性是指图像的放大缩小都不会影响特征值,只要图像还能分成 8×8 的网格,特征值基本不会发生改变。另外,当云图整体变亮或变暗时,也不会对特征值产生太大影响。

通过 4.2 节的实验可以发现 Zernike 矩具有较好的旋转不变性,但检索的效果不理想;相比网格法的检索效果较好,具有尺度不变性,但不具备旋转不变性^[12]。在实验过程中尝试了几种将 Zernike 矩和网格法融合的方法,但融合后的检索结果较差,因此文中设计网格内切换法使其具有旋转不变性。在此,仅对所截取云图的内切圆部分进行特征提取,使特征具有旋转不变性,文中对于截取的内切圆云图每隔 30 度提取一次网格特征值,由此每幅云图具有对应 12 个方向的网格特征值。在进行相似性比较时,对于给定的待检索云图,先提取其网格特征,然后用这个网格特征值跟特征库中云图的 12 个方向的网格值进行比较,找出最相似的方向,最后再来计算这两幅云图之间的相似度。图 2(a)为一幅图像的内切圆,图 2(b)为云图的内切圆部分对应的内容,图 2(c)为对内切圆所进行的网格划分。

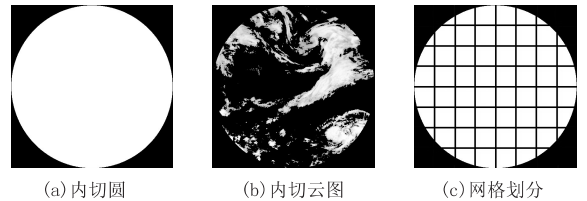


图2 网格内切圆

仔细分析一下,就可以发现,如果事先求出云图库中每幅云图对应的 12 个方向的网格特征,显然会使得离线生成特征库的工作量至少变为网格法的 12 倍。事实上完全不需要生成 12 个方向的网格特征,只需要生成 0 度方向的网格特征就可以了,但在相似性检索时,需要生成待检索云图的 12 个方向的网格特征,用这 12 个网格特征分别与每幅云图的 0 度网格特征进行比较,找出 12 个中最相似的作为这两幅云图之间的

相似度。这样做,在离线生成特征库时的压力大大减少了,而在线检索时的压力几乎没有增加(因为只生成一幅云图的12个方向的网格特征跟只生成一幅云图一个方向的特征的时间差别是很小的,可以忽略不计)。这样既达到较好的检索效果,又不会花费太多时间代价。

3 距离度量

在进行在线检索时,需设定合理的距离度量准则反映两幅不同云图的特征向量相似性。文中采用 χ^2 距离来计算待检索云图的特征向量与特征库中特征向量的距离, χ^2 距离的定义^[13]如下:

$$d_{\chi^2}(A, B) = \sum_{i=1}^n \frac{(a_i - m_i)^2}{m_i} \quad (3)$$

式中, A, B 是两个 n 维的特征向量; m_i 的定义为

$$m_i = \frac{a_i + b_i}{2} \quad (4)$$

式中, a_i, b_i 分别对应特征向量 A, B 的第 i 个分量。

由 χ^2 距离的定义可以看出, χ^2 距离可以有效地减少量纲的影响,不需要对每个特征分量都进行归一化处理。

4 实验结果与分析

为了验证方法,文中在主频为2.26 GHz,内存为2 GB的PC机上,采用.Net3.0和OpenCV的混合编程进

行了实验。实验数据为风云二号卫星2010年和2011年的数据。在比较形状特征时只对2011年红外云图中的900幅兰勃托投影的云图进行实验,主要验证旋转不变性。在验证旋转不变性的实验过程中,将这900幅云图旋转180度后加入到这个云图库中,共1800幅云图。

4.1 预处理实验结果与分析

在云图预处理阶段,对2011年5月6日3时文件名为“EILY0630.AWX”的红外卫星云图实现云地分离,处理结果如图3所示,其中阈值设定为160,可以看出采用文中方法可较好地分离云与地面背景。

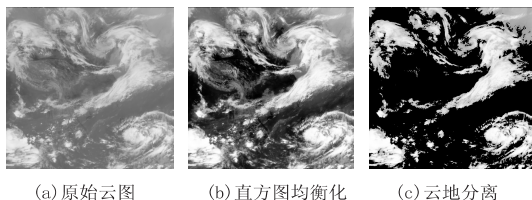


图3 云图的预处理

4.2 检索结果与分析

在进行库内卫星云图的相似性检索时,文中用2011年5月6日3时文件名为“EILY0630.AWX”的红外卫星云图,分别采用Zernike矩、网格法以及网格内切圆法对2011年的红外云图库进行检索。

4.2.1 网格内切圆法的检索结果与分析

采用Zernike矩得到的前16幅最相似的检索结果如图4所示,检索结果按照相似度的大小从高到低排

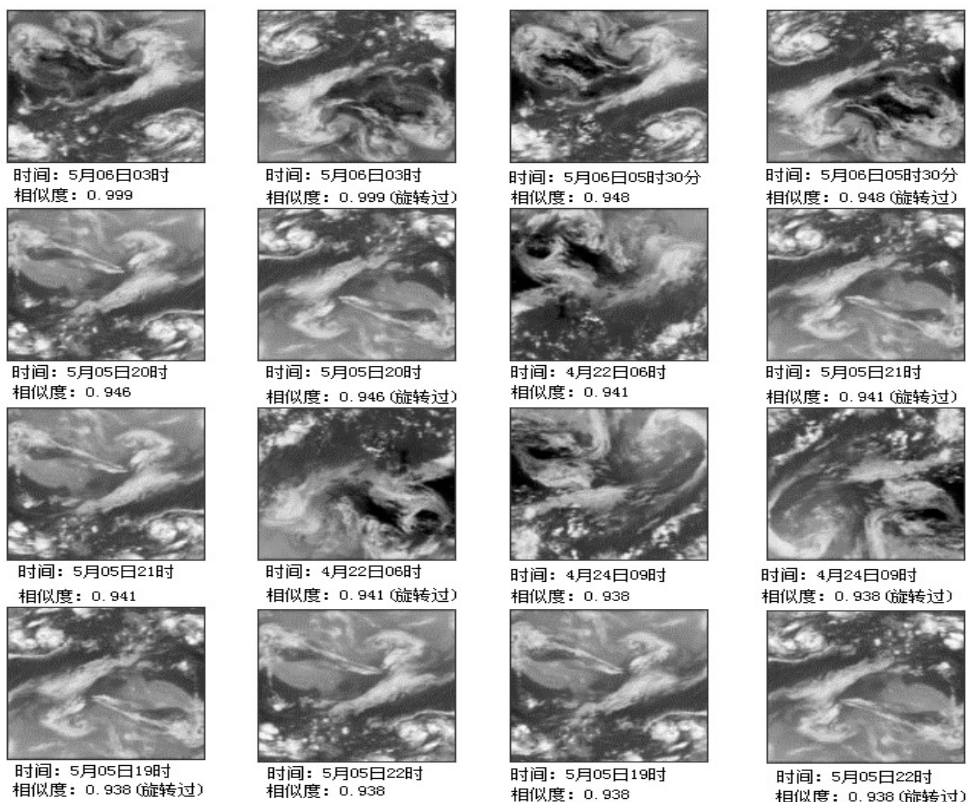


图4 Zernike矩的库内检索结果

列。图下的标注分别为检索到的云图时间以及有待检索云图的相似度。对检索结果分析可知,云图及其旋转的图像能够被同时检索出来,这说明 Zernike 矩具有极好的旋转不变性,但是检索的可靠性不高,部分检索结果在视觉上的相似度并不高,比如检索结果中排序

第 7、10、11、12 的,视觉上看起来不是很相似。

4.2.2 网格法的检索结果与分析

用同样的待检索云图,采用网格法得到的前 16 幅最相似的检索结果如图 5 所示。对检索结果分析可知,旋转后的云图并没有被检索出来,这说明网格法不

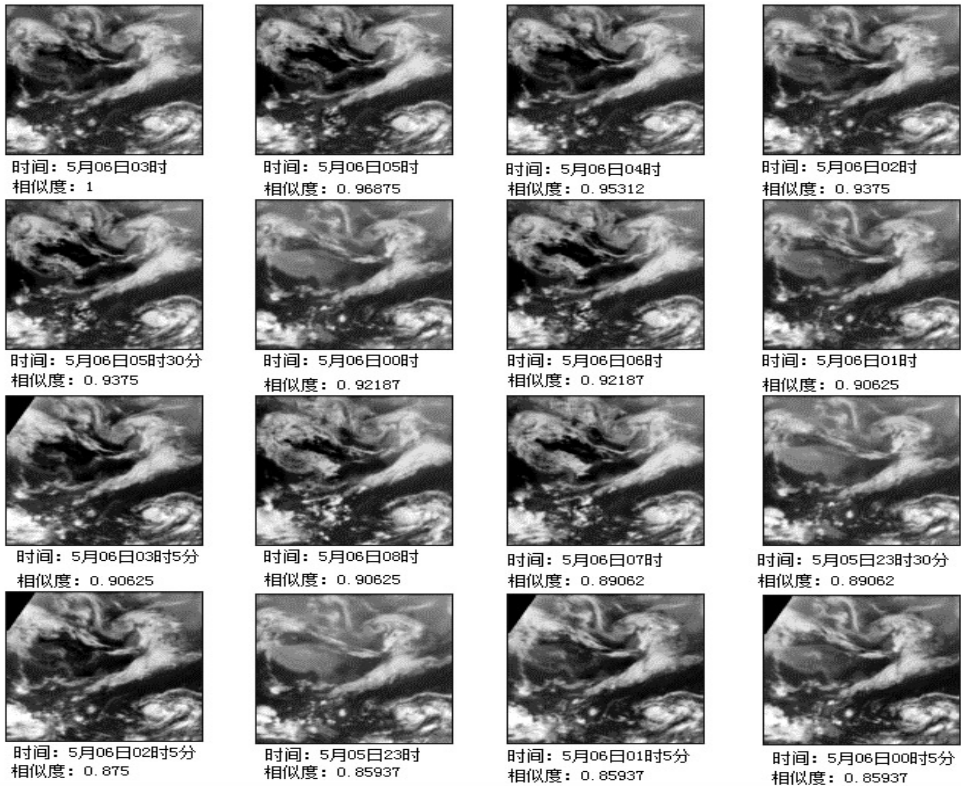


图 5 网格法的库内检索结果

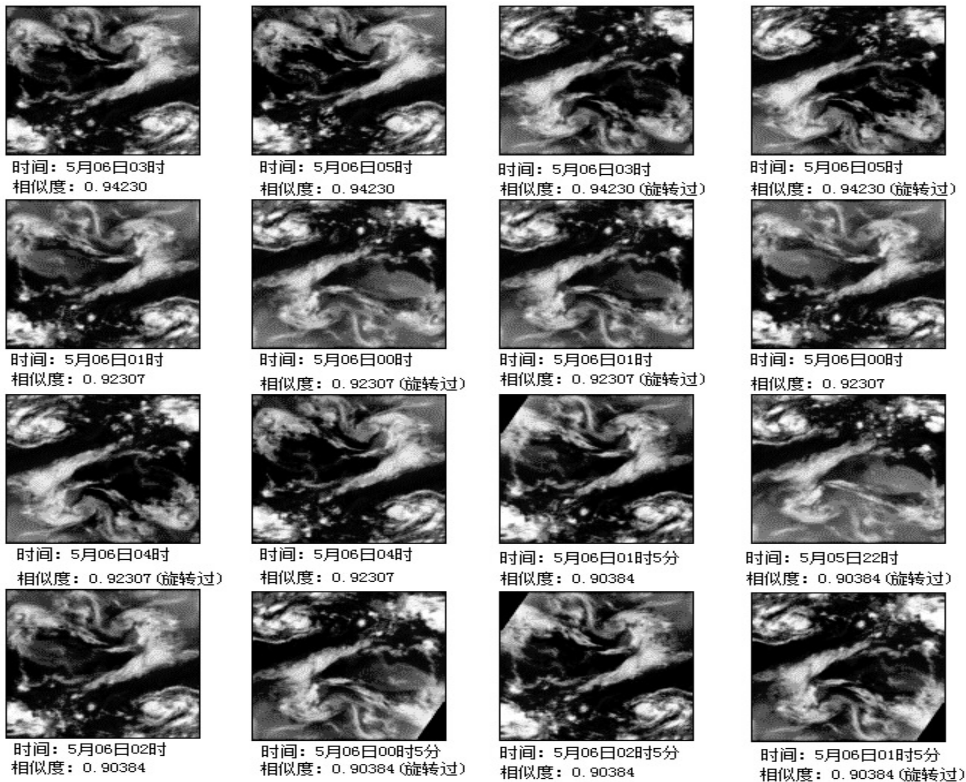


图 6 网格内切圆法的库内检索结果

具有旋转不变性,但是检索出来的未旋转的云图的效果较好,并且具有较高相似度的云图在成像时间上与待检索云图十分接近,符合云在时序范围内连续运动的规律。

4.2.3 网格内切圆法的检索结果与分析

文中采用网格内切圆法,继续用从 2011 年的云图库中选取的 5 月 6 日 3 时文件名为“EILY0630. AWX”的兰勃托投影下的红外卫星云图,对 2011 年的兰勃托投影下的红外云图库进行检索。其前 16 幅最相似的检索结果如图 6 所示,对检索结果分析可知,网格内切圆法不但具有一定的旋转不变性,而且检索的可靠性也较好,具有较高相似度的云图在成像时间上与待检索云图十分接近,符合云在时序范围内连续运动的规律,该方法明显比前两种方法的检索效果要好。

5 结束语

文中采用几种常用形状区域特征对卫星云图进行相似性检索,并介绍了详细的检索流程。文中值得注意的地方是:首先,在进行云地分离前先进行直方图均衡化,可较好地对整个云图库进行云地分离,即便是不同类别的云图,也可以取一个普遍适用的阈值;其次,对比了几种形状区域特征的优缺点,提出了一种改进的方法。实验结果表明,文中方法能有效提取云图的特征信息,能够较好地对历史卫星云图进行相似性检索,具有一定的应用前景。在未来的工作中,将重点研究如何将形状特征与纹理特征进行融合,提高检索的可靠性。

(上接第 59 页)

参考文献:

- [1] Harrison W H. Compiler analysis of the value ranges for variables[J]. IEEE transactions on software engineering, 1977, 3(3):243-250.
- [2] Nielson F, Nielson H R, Hankin C. Principle of program analysis[M]. Berlin: Springer Verlag, 1999:211-282.
- [3] 李梦君, 李舟军, 陈火旺. 基于抽象解释理论的程序验证技术[J]. 软件学报, 2008, 19(1):17-26.
- [4] 张幸儿. 计算机编译原理: 编译程序构造实践[M]. 北京: 科学出版社, 2009.
- [5] 苏青琴, 刘久富, 陈 魁, 等. 基于抽象解释的非函数依赖不变量的检测方法[J]. 计算机技术与发展, 2012, 22(4):5-8.
- [6] Cousot P, Cousot R. Abstract interpretation: An unified lattice model for static analysis of programs by construction or approximation of fix points[C]//Proc of the 4th POPL. Los Angeles: ACM Press, 1977:17-19.
- [7] 姬孟洛, 王怀民, 李梦君, 等. 一种基于抽象解释和通用单

参考文献:

- [1] 孙学金, 王晓蕾, 李 浩, 等. 大气探测学[M]. 北京: 气象出版社, 2009.
- [2] 温泉彻, 彭 宏, 黎 琼. 基于内容的图像检索关键技术研究[J]. 微计算机信息, 2007, 23(1-3):278-280.
- [3] 陈渭民. 卫星气象学[M]. 北京: 气象出版社, 2005.
- [4] Loncaric S. A survey of shape analysis techniques[J]. Pattern recognition, 1998, 31(8):983-1001.
- [5] 孙君顶, 赵 珊. 图像低层特征提取与检索技术[M]. 北京: 电子工业出版社, 2009.
- [6] 李俊山, 李旭辉. 数字图像处理[M]. 北京: 清华大学出版社, 2010.
- [7] Zhang D S. Image retrieval based on shape[D]. Australia: Monash University, 2002.
- [8] 王水璋. 基于纹理的图像检索技术研究[D]. 太原: 太原理工大学, 2008:140-150.
- [9] 杨政武, 方 涛. 基于 Zernike 矩的图像归一化技术的研究[J]. 计算机工程, 2004, 30(12):34-36.
- [10] 李金泉, 王建伟, 陈善本, 等. 一种改进的 Zernike 正交矩亚像素边缘检测算法[J]. 光学技术, 2003, 29(4):500-503.
- [11] Li S Z. Matching: Invariant to translations, rotations and scale changes[J]. Pattern recognition, 1992, 25(6):583-594.
- [12] Lu G J, Sajjanhar A. Region-based shape representation and similarity measure suitable for content-based image retrieval[J]. Multimedia system, 1999, 7(2):165-174.
- [13] Zhang D S, Lu G J. Evaluation of similarity measurement for image retrieval[C]//Proc of IEEE international conference on neural networks & signal processing. Nan Jing, China: [s. n.], 2003:928-931.

调数据流框架的值范围分析方法[J]. 计算机研究与发展, 2006, 43(11):2020-2026.

- [8] Cousot P. Abstract interpretation based formal methods and future challenges[C]//Informatics 10 years back, 10 years ahead. London: Springer Verlag, 2001:138-156.
- [9] 王 伟, 刘久富, 娄坚波, 等. 基于多 Agent 的软件测试系统设计[J]. 计算机技术与发展, 2011, 21(4):37-39.
- [10] 赵修伟. 基于抽象解释的实时软件 WCET 研究[D]. 大连: 大连理工大学, 2009.
- [11] Shannon C E. A mathematical theory of communication[J]. Bell system technical journal, 1948, 27:379-423.
- [12] 陈立前, 王 戟, 刘万伟. 基于约束的多面体抽象域的弱接合[J]. 软件学报, 2010, 21(11):2711-2724.
- [13] Pincus S M. Approximate entropy (ApEn) as a complexity measure[J]. Chaos, 1995, 5(1):110-117.
- [14] Miné A. The octagon abstract domain[J]. Higher-order and symbolic computation, 2006, 19(1):31-100.

形状特征的卫星云图检索方法研究

作者：[夏士明](#)，[李骞](#)，[谢凯翔](#)，[XIA Shi-ming](#)，[LI Qian](#)，[XIE Kai-xiang](#)
作者单位：[解放军理工大学 气象海洋学院](#)，[江苏 南京](#)，[211101](#)
刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2014(4)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201404015.aspx