

分 类 号 _____
学校代码 10542

密级 _____

基于 1:5000 国家基本比例尺地形图的
小流域自动提取算法研究
——以长沙市岳麓山为例

Automatic extraction algorithm of small
watershed based on 1 : 5000 scaled national
basic rigorous map: A case study from Yuelu
Mount, Changsha

论 文 编 号 200813020063

学 科 专 业 自然地理学

研 究 方 向 区域自然地理

湖南师范大学学位评定委员会办公室
二零一一年四月



摘 要

随着国家空间信息基础设施建设的发展, 1:50000 和 1:10000 比例尺地形图在全国范围内已经普遍建设完成, 在大城市或部分地区还编绘了 1:5000 的地形图。1:5000 比例尺地形图是我国基本比例尺地形图中比例尺最大、精度最高的基础数据, 可以用来研究小范围内地表详细信息(地貌、资源、灾害状况等)。而目前针对高精度 DEM 的研究通常是基于 1:10000 或 1:50000 比例尺地形图, 而基于 1:5000 比例尺地形图的研究比较少。

岳麓山在江南丘陵区具有很强的代表性, 它具有亚热带常绿阔叶林下风化层厚和沟谷发育的典型地貌特征。由于植被茂盛、坡面陡峭、小沟谷比较多, 定量观测较难正常进行。通过 1:5000 比例尺地形图生成更高精度 DEM, 进而进行小流域的自动提取研究, 可以直观地观察与分析岳麓山不同部位的地貌特征和坡面、沟谷演化规律, 定量地对地貌特征进行量测, 从而节省大量人力、物力和财力。同时, 该项研究对进一步分析小流域沟谷等地貌特征和坡面、沟谷演化规律, 也能为广大丘陵地区的资源开发、保护和灾害防治提供基础数据。

本文在参阅有关 DEM 流域提取算法的基础上, 针对目前存在的主要问题等, 探讨将基于地形表面几何形态分析和地形表面流水模拟

分析的原理与方法相结合的方法来提取小流域。其技术路线是：(1) 先采用几何分析的方法根据特征点的特性从 DEM 数据中提取可能是山脊点和山谷点的候选点；(2) 然后再用流水模拟的方法提取流域；(3) 再采用缓冲区分析和叠置分析生成新的栅格图层，然后对沟谷进行二次提取以处理伪沟谷；(4) 最后以候选山谷点作辅助对小流域沟谷进行修正，实现小流域特征信息的自动提取。

本文以岳麓山作为例对小流域特征信息进行了提取，主要技术环节如下：

(1) 洼地填充。本文假设所提取到的洼地都不是真实的地表形态，对所有洼地都进行填平处理。采用“高程增量迭加法”进行洼地填冲，通过增高洼地栅格单元的高程值，使得 DEM 中每一个栅格的高程值不比周围八个栅格中的高程值中至少一个低，然后就可以计算水流方向。

(2) 流域阈值。先计算所有栅格流量的平均值，然后分别在该平均值两侧各选两点为阈值，生成沟谷。通过对阈值与沟谷总长度进行线性函数、对数函数、多项式函数、乘幂函数进行分析，从而计算出合理的阈值。

(3) 伪沟谷的处理。首先把 Strahler 三级分级方案中的第二级别和第三级别的沟谷标记为主沟谷；然后对标记的主沟谷进行缓冲区分

析,缓冲区范围内的栅格单元降低一定的高程值,缓冲区范围外的栅格值赋为零,生成新的栅格图层;再将其与原始 DEM 进行叠加分析,对叠加后生成的 DEM 重新进行小流域提取,小流域提取后的沟谷数据就是进行伪沟谷删除的结果。

(4) 洼地匹配。读取提取的洼地图层中的洼地单元,然后对山谷点图层进行筛选,删除山谷点图层中与洼地点坐标相同的山谷点;然后提取沟谷线图层上的每一条沟谷的转折点,删除被该洼地单元包含的转折点(不能是沟谷线的起点或终点)。

本文通过流域自动提取算法实验所提取出的流域特征信息与手工提取的进行比对,其精度达到预期的要求。本文的主要研究结论有:

(1) 选取五组数据对阈值与沟谷总长度进行相关性分析,发现两者的相关性与乘幂函数 $Y=404115 \cdot X^{-0.557}$ 的拟合度最好, R^2 的值达到 0.9995。当阈值设置为 145.54 时,生成的沟谷与实际沟谷的长度比较接近,沟谷密度较近似,阈值较合理。通过进行相关性分析,可以快速地选择恰当阈值,使生成的沟谷与实际沟谷更加相符,避免了反复实验阈值的繁琐过程,也保证了模拟结果的准确性。

(2) 通过对岳麓山提取到的小流域沟谷网络分析,发现填充处理后的洼地仍然对小流域的沟谷产生影响,使沟谷明显弯曲,结合填充前提取的洼地对转折点进行删除,使沟谷更加合理,更符合实际。

(3) 比 1:10000 比例尺地形图相比, 岳麓山 1:5000 比例尺地形图的精度更高, 这样通过 1:5000 比例尺地形图更高精度 DEM 数据来进行小流域的提取, 所提取到的小流域沟谷网络更加精确。而且通过流域二次提取, 实现对伪沟谷进行了自动处理, 在小流域范围内实现完全删除伪沟谷, 减少了手工工作量, 使小流域沟谷网络更加合理。

本文以岳麓山作为例对小流域特征信息进行了提取, 选取范围比较小, 还要更进一步多选取一些典型区域进行分析, 使算法更加成熟, 能够广泛应用于丘陵区域小流域的提取。

关键词: 岳麓山, 小流域, DEM, 1:5000 地形图, 沟谷, 阈值

ABSTRACT

Together with the development of national spatial information infrastructure, 1:50000 and 1:10000 scale basic rigorous map construction has been generally completed in the country. As for large cities and coastal regions the 1:5000 scale basic rigorous maps are compiled. 1:5000 scale basic rigorous maps are the largest scale, most accurate basic data. It can be used to research the details of terrain (topography, resources, disaster situation, etc.). At present, high-precision DEM for the study is usually based on 1:10000 or 1:50000 scale basic rigorous maps. But relatively a few studies based on 1:5000 scale basic rigorous maps.

The Yuelu Mount is typical in the southern rolling ground. It represents the typical landform features which happened in regions with sub-tropical evergreen broad-leaved forest, thick weather, valley development. Yuelu Mount with high vegetation coverage, steep slope and many small gullies bring the disadvantage condition to the on-the-spot quantitative investigation. So the small watershed research extracted by 1:5000 scale DEM can be directly observed the topographic features、slope and valley evolution in Yuelu Mount. It can measure the landscape features of Yuelu Mount and save a lot of material and financial resources. It is benefit to analysis topographic features such as watershed and valley slope, valley evolution. Furthermore, It provide basic data for development, protection and disaster prevention for the majority of the southern rolling ground.

In order to deal with the problems in watershed extraction, this research refers to the algorithms for the extraction of watershed characteristics from DEM, and researched the combination approach about geometry analysis and water simulation analysis to extract the watershed. The way includes: 1, the feature points which may be the candidate of the ridge and valley points are extracted from the DEM data through geometric analysis method; 2, the watershed is extracted through simulation analysis method; 3, the watershed is extracted secondly from

the new raster data produced by further analysis and overlay analysis method. In this way pseudo-valley was corrected; 4, the watershed network is corrected with the candidate of the valley point. So the watershed characteristics extraction is complete.

As the case the watershed extraction of Yuelu Mount, the main technical aspects are as follows:

1、filling local depression. The local depression assumed it is not really of the surface are filled up. The method is incremental elevation superposition method. It increases the value of grid cell elevation and makes at least one elevation of the surrounding grid is lower than the elevation of center grid. Then water flow can be calculated and the watershed can be extracted.

2、the watershed threshold. First calculate the average of all grids, and then were selected in each side of the average of two points as the threshold value in order to generate the valley. Comparing to the length of the valley, the reasonable threshold can be calculated by a linear function, logarithmic functions, polynomial functions, power function analysis.

3、pseudo-valley treatment. The second level and third level of the valley classified by the Strahler classification are saved as the main valley. After the buffer analysis of the main valley, the elevation of the grid in the buffer size is reduced by a certain value and the elevation of the grid outside the buffer size is assigned zero. a new raster layer is generated by this way. Overlaying the new raster layer and the DEM data the small watershed can be extracted. Now the valley is the result of pseudo-valley treated.

4、the treatment of Matching with the local depression . After extracting the local depression in the depression layer, the channel layer is filtered with the local depression .Then the valley points whose point coordinate is the same with the local depression are deleted. Last the turning points in the valley in the valley layer are checked. The turning points contained by the local depression are deleted. But the turning point must be not at the start or the end of the valley.

Comparing to the watershed extracted by hand, the accuracy of the watershed extracted by algorithm automatically achieves the desired requirements. The main conclusion of this research is following:

1、 After the correlation analysis for the threshold and the length of the valley, the power function $Y=404115 \cdot X^{-0.557}$ is best and the value of R^2 is 0.9995. When the threshold is 145.54 the length and the gully density of the valley is close to the actual value. So the threshold selected is more reasonably. Through the correlation analysis, you can quickly choose the appropriate threshold, so that the valley generates more consistent with the actual valley, avoiding the tedious trial and error process of the threshold and ensure the accuracy of the simulation results.

2、 Though the analysis of the watershed valley, you can discover the local depression that still affect the watershed valley after filling. So the turning point of the valley which is in the local depression need to be deleted so that the valley is more reasonable and more realistic.

3、 Comparing to the scale of 1:10000, the accuracy of the 1:5000 scale basic rigorous map of the Yuelu Mount is higher and the watershed is more reasonably. And through secondary extraction, the pseudo-valley were processed automatically and deleted completely. The manual work is reduced and the small watershed valley is more reasonable.

In this research the watershed was extracted from the 1:5000 DEM data of Yuelu Mount. But the area of the Yuelu Mount is relatively small. So the research need more typical regions to analysis and test in order to making the algorithm is more mature and can be widely in the small watershed extraction in rolling ground.

Key words: the Yuelu Mount, small watershed, DEM, valley, threshold, 1:5000 scale basic rigorous map

目 录

摘 要	I
ABSTRACT	V
第一章 绪 论	1
1.1 研究的目的和意义	1
1.2 研究现状与发展趋势	3
1.3 研究内容	13
第二章 研究材料与方法	16
2.1 研究区概况	16
2.2 DEM 数据	18
2.3 研究基础方法	21
第三章 算法设计与实现	26
3.1 算法设计	26
3.2 算法实现	38
3.3 算法评估	40
第四章 总结与展望	49
4.1 总结	49
4.2 展望	51
参考文献	53
附录	57
1 程序代码	57
2 数据表	63
致谢	64

第一章 绪 论

1.1 研究的目的是和意义

1.1.1 研究的目的

随着国家空间信息基础设施建设的发展, 1:50000 和 1:10000 比例尺地形图在全国范围内已经普遍建设完成, 在经济比较发达部分地区还建立了 1:5000 的地形图, 如长沙市。1:5000 比例尺地形图是我国基本比例尺地形图中比例尺最大、精度最高的基础数据, 可以用来研究小范围内详细信息和评价地形。而目前针对高精度 DEM 的研究通常是基于 1:10000 或 1:50000 比例尺地形图, 而基于 1:5000 比例尺地形图的研究比较少, 而通过 1:5000 比例尺地形图生成更高精度 DEM 进行研究, 有利于进一步分析小流域沟谷等地貌特征和坡面、沟谷演化规律。

岳麓山的地貌特征在整个江南丘陵地区具有典型代表性, 代表了亚热带常绿阔叶林植被下风化层地壳厚、降水量大和地表径流比较活跃的典型地貌过程; 由于植被茂盛、坡面陡峭、小沟谷比较多, 定量观测较难正常进行; 因此利用大比例尺 DEM 进行小流域的自动提取研究, 可以直观地观察岳麓山的整体特征, 定量地对地貌特征进行量测, 从而节省大量人力、物力和财力。

目前应用于流域分析的地理信息系统或水文分析的软件大多是基于栅格结构的,如 ArcGIS 系统。因此基于 DEM 数据提取流域特征信息,从提取效率和结果的精度来说都是切实可行的。但从微观上讲,已有的算法或软件还无法确保结果与一些特殊地形区域的实际沟谷情况相一致,如谷地或平坦区域。

本文利用岳麓山区域 1:5000 的国家基本比例尺地形图构建数字高程模型 (DEM),然后基于高精度的 DEM 数据,对该小流域进行自动提取的研究,分析以高精度数据为基础影响流域提取的因素以及相应的处理方法,设计与实现小流域的自动提取算法,揭示不同阈值对流域特征信息的影响,提取岳麓山的小流域特征信息,为江南丘陵地区的小流域研究提供参考。

1.1.2 研究的意义

地貌是自然环境的重要组成部分之一,地貌通过其表面形态(如坡度、坡向、地形起伏等)、空间展布(如高程、高差、距离等)以及各种形态的类型组合(如高原、平原、山地、丘陵、盆地等)组成千姿百态的地表景观。地貌的形态、特征不仅影响着物质和能量的再分配,从而影响水文过程及其结构,而且影响土壤与植被的形成和发育过程,决定着自然资源及其组成的生态环境的类型以及利用方式和可以利用程度。因此,对区域范围的地貌特征进行分析,是研究自然资源及其组成的生态环境的形成、发展、演变的重要基础,对认识自然资源、保护自然资源以及开发的自然资源都具有重大实践意义。

沟谷是整个流域地貌系统中最活跃的部分之一。在沟谷中,流水的侵蚀作用和重力作用都很活跃,导致沟谷的地貌形态变化比较明显,尤其是在纵横剖面形态上的变化。同时由于沟谷的地貌形态以及形成过程都比较复杂,使得沟谷地貌研究成为流域地貌系统研究中的相对薄弱的环节^[29]。

流域的自动划分在实际应用中具有重要意义。例如,区域水土保持的重要工作之一是小流域治理(刘黎明,1998),其产流汇流过程的模拟(刘昌明,1998)均离不开小流域这一精确划分治理模型及其评价的基本单元^[23]。

因此,基于 1:5000 国家基本比例尺地形图生成更高精度 DEM 进行小流域自动提取研究,有利于进一步分析小流域沟谷等地貌特征和坡面、沟谷演化规律,也为“数字流域”提供一定的参考价值。

1.2 研究现状与发展趋势

1.2.1 数字高程模型 (DEM)

从 1955 年到 1960 年期间,美国麻省理工学院的摄影测量实验室主任 Chaires.L.Miller 教授在美国麻省土木工程部门和美国交通部门研究工作期间,首次将计算机与摄影测量技术结合在一起,比较成功地解决了道路工程中的计算机辅助设计问题^[28]。Miller 在解决道路计算机辅助设计这一特殊工程课堂的同时,和 LaFlamme 提出了一个一般性的概念:数字地面模型 (DTM:Digital Terrain Model)。自从 DTM 的概念提出以后,国际上相继又出现了许多其他近似的术语,如

DHM(Digital Height Model)、DGM(Digital Ground Model)、DTEM(Digital Terrain Elevation Model)和DEM(Digital Elevation Model)等^[22]。这些近似的术语实质上表达的意思差别很小。通常认为DTM趋向于表达比DEM和DHM更广意义上的内容,如DTM包括河流、山脊线、山谷线、断裂线等,而DEM不包括在内。现在普遍认为:DTM是地面属性特征信息空间分布的有序数值阵列,是对地表属性的数字化表达;数字高程模型(Digital Elevation Model,简称DEM)是DTM中最基本的部分,是当地形属性为高程时的DTM;DEM是对地球表面形态的一种离散的数字表达^[28]。

自从DEM提出以来,随着计算机技术的发展与数据库的广泛应用,对DEM的认识不断深入,DEM成为标准的基础地理信息产品。DEM的应用范围也随之而扩大,并且通过与地理信息系统的结合将DEM概念扩充到表示与地表相关的各种属性,与GIS联合进行空间分析,路线选址,水文分析,地貌分析,结合其它特征(如降水量等)进行综合分析,模拟虚拟现实等。

1.2.1.1 DEM 的含义

DEM是对某一种或多种地面特性空间分布的数字描述,是叠加在二维地理空间上的一维或多维地面特性向量空间,是地理信息系统(GIS)空间数据库的某类实体或所有这些实体的总和。DEM的本质共性是二维地理空间定位和数字描述^[22]。

DEM是对地球表面地形地貌的一种离散的数字表达,是表示区

域 DTM 上的三维向量有限序列, 用函数的形式描述为下面的公式:

$$V_i = (X_i, Y_i, Z_i) \quad (i=1, 2, 3, \dots, n) \quad \text{公式 1-1}$$

公式 1-1 中, X_i, Y_i 是平面坐标, V_i 是 (X_i, Y_i) 对应的高程。当该序列中各平面向量的平面位置呈规则格网排列时其平面坐标就可省略。此时 DEM 就简化为一维向量序列 $\{Z_i, i=1, 2, 3, \dots, n\}$ ^[22]。

1.2.1.2 DEM 的形式

DEM 表达形式有多种, 其中主要的有三种形式: 等高线、规则矩形格网与不规则三角网。其中规则矩形格网是目前应用最广泛的一种表达形式。

等高线 DEM 是指经过数字化处理以后的一系列的等高线集合与它们的高程值结合在一起就构成的 DEM。等高线 DEM 通常采用附加地形特征数据, 如特征点、山脊线、山谷线等, 从而构成完整的 DEM, 从而实现完整地表示地貌。规则矩形格网 DEM 是指在高斯投影平面上一系列在 X, Y 方向上等间隔排列的地形点的平面坐标 (X, Y) 及其高程 (Z) 的数据集。规则矩形格网 DEM 存储量最小, 可以有效地压缩存储, 便于使用, 容易管理, 但是地形简单的区域存在大量冗余数据; 对山丘、丘陵地或地貌比较破碎的地区的内插高程与实际高程有一定的误差。不规则三角网 DEM, 通常称三角网 DEM 或 TIN (Triangulated Irregular Network) 是通过从不规则分布的数据点生成的连续三角面来逼近地形表面, 是模拟地形最基本的一种形式。TIN 能较好地顾及地貌特征点、线, 表示复杂地形比矩形格网精确。但是,

TIN 的数据量较大, 数据结构较复杂, 使用与管理也较复杂, 因此不能得到广泛的应用^[22]

1.2.1.3 影响 DEM 数据质量与应用精度的因素

DEM 数据采集的模式主要有以下三种: (1)野外实地测量; (2)地图扫描矢量化。 (3)数字摄影测量。地图扫描矢量化的采集方法不仅能够充分利用原有的纸质地形图成果, 而且生产成本经济低廉、采集速度比较快, 易于批量作业, 因此我国基础地理数据库中各级的比例尺 DEM 均是在对相应比例尺的纸质地形图进行等高线数字化的基础上建成的^[22]。

因为 DEM 是基于原始的等高线地形图基础上建立的, 而原始地形图是有一定比例尺的, 所以等高线比例尺的差异必然导致所生成 DEM 也保留了地图制图综合、取舍的烙印。因此 DEM 的精度与其原始地形图精度之间存在着必然的联系。这个比例尺也决定了 DEM 精度和分析适宜性和所包含的地形信息容量是不同的^[22]。

DEM 对地形的表达采用了栅格这样一种离散点模拟的形式, 栅格分辨率的大小, 在很大程度上影响了地形描述的精度, 也是不确定性产生的重要根源之一。

DEM 的栅格分辨率按照其在空间的分布, 可以划分为两种基本类型—垂直分辨率和水平分辨率。

水平分辨率是指 DEM 栅格在水平方向边长的大小。一般来讲, 分辨率比较高的水平分辨率能够更好地拟合真实地球表面, 误差也相

对较小;而分辨率比粗略的水平分辨率对地球表面的概括程度更高,误差也相对较大,在地形上相对比较破碎的地区对地形描述精度的影响更加明显。如果格网间距太大,会产生大量数据,造成分析不便利,速度比较慢;但是如果格网间距太小,又不能够很好的表达地形描述精度,所以必须根据研究目的和具体情况来选择合适的 DEM 分辨率。

垂直分辨率是一个相对比较容易被人们忽略的概念。垂直分辨率是指 DEM 所记录高程数据表示的有效位数。垂直分辨率对地形描述精度的影响是随地貌类型的不同而有相当大的差异。在地形有很大起伏的山区,在小数点后比较细微的高程变化一般对总体地形描述不会产生较大或者明显的影响;但是在地面平坦的平原地区,微小的高程差异也会给地面汇流特征的描述带来较大的影响,从而最终影响诸多相关地形因子的提取与分析。

1.2.2 DEM 提取算法的应用性研究

1.2.2.1 宏观尺度上的应用性研究

目前从宏观尺度上,利用 DEM 对地貌的特征及其空间分布的差异进行的研究比较多。刘爱利、汤国安利用 1:100 万的 DEM 提取地形起伏度、地表切割度、地表粗糙度、高程变异系数、平均坡度、平均高程 6 个地形因子,对中国地貌的基本形态进行了多维信息综合分类^[34];刘新华利用 1:100 万的 DEM 提取了中国水土流失地形起伏度,完成了中国水土流失地形起伏度制图,最后对中国水土流失地形起伏度进行了适用性分析^[2];王雷等利用 1:25 万 DEM 提取了

有关坡度、坡向、地形起伏度等的地貌特征要素,并进行定位表达与特征统计分析,获得了对本区地貌特征的定位与定量化的总体认识,为农业规划、水土流失、土壤侵蚀、地质灾害等研究提供了基础^[1];莫申国利用格网精度为 100 m×100 m 的 DEM 提取秦岭的主分水岭,确定秦岭地貌的分类分级指标,对秦岭的起伏度和高度数据进行分类、编码,研究秦岭的数字地貌格局^[3]。

1.2.2.2 小区域大比例尺的应用性研究

利用 DEM 在小区域大比例尺方面的研究也取得了许多进展。赵健等利用 1:10000 大比例尺 DEM 提取万花乡向阳沟小流域的沟谷特征和其它常用的流域特征信息^[4];朱红春等利用 1:10000 大比例尺 DEM 提取陕北黄土高原不同地貌区的地形起伏度^[5];朱红春等利用 1:10000 大比例尺 5m 分辨率的高精度 DEM 提取黄土高原地区的 6 个典型地貌区域内沟壑密度、平均坡度、平均坡长和地形起伏度等基本地形信息因子,研究了沟壑密度与平均坡度、平均坡长和地形起伏度之间的线性相关关系^[6];陈浩等利用 1:13000 的 DEM、DOM 和 1:5000 的大比例尺地形图,研究分析王家沟流域黄土丘陵沟壑区域坡向上沟谷侵蚀演化的差异性,探讨了流域坡面与沟谷侵蚀演化的关系^[7]。

由此可见,大量的多尺度的利用 DEM 开展地貌学特征及其过程的分析研究,深化了已有的理论与研究,探索出许多新的研究方法,为进一步的研究奠定了必要的基础。

1.2.3 DEM 提取算法研究

1.2.3.1 基于图像处理技术的算法

最初, 因为规则格网 DEM 数据事实上是一种栅格形式的数据, 格网大小对应图像像素大小, 而高程值对应灰度值, 可以利用数字图像处理中的技术来设计算法。利用数字图像处理技术设计的算法大都采用各种滤波算子进行边缘提取。这种方法提取山脊山谷线的主要过程有两步: 首先提取特征点 (如山脊点、山谷点等); 然后将特征点连成特征线(如山脊线、山谷线)^[30]。

T.K.Pecuker 和 D.H.Douglas 提出了一种简单移动窗口的算法, 通过 2×2 窗口对 DEM 格网矩阵的扫描, 把 DEM 中可能的特征点全部提取出来。L.E.Band 改进了这一算法, 标示河流段上游的源点和下游的出水口, 然后进行细化处理, 从而将特征点连接成线。这些算法简单快捷, 但是在提取特征点时必须排除 DEM 中噪声的影响, 在将特征点连接成线的算法设计比较困难。余生晨等通过对 DEM 进行富里叶变换处理, 然后在频率域 (或波数域) 上每隔 5 度求一个方向导数 (沿着垂直与山谷线、山脊线的方向) 来检测山脊点和山谷点, 这种方法比较简单, 易于实现, 但是计算量比较大^[8]。

1.2.3.2 基于地表几何形态分析的算法

基于地表几何形态分析的典型算法是断面极值法。其基本思路是: 首先找出 DEM 的横向和纵向的两个断面的极大值点和极小值点,

然后根据一定的条件或准则将这些极值点划归各自所属的地形特征线，而地形断面上高程的极大值点就是分水点，而高程的极小值点就是汇水点。这种方法忽略了地形特征性的曲率变化，导致后续跟踪地形特征线比较困难，确定出的地形特征线也与实际的地形特征线有一定的差异，有时候会出现遗漏现象^[30]。

局部窗口法是通过在一个小矩形窗口内对高程值进行局部计算，然后在 DEM 上有序的移动该窗口来确定上该小矩形窗口的平面是否凹面或 V 型面，然后提取在这些凹面或 V 型面的底部的格网单元，从而生成水道或沟谷。Puecker、Douglas 和 Jenson 应用这种方法来生成沟谷。Band 也应用这种方法来分析流域的地形地貌，从而提取出流域。这种方法最严重的缺陷就在于生成的水系不连续性，而且对于地形起伏较小或者地形比较复杂的区域，这种不连续性更加突出，因此，导致了这一技术的应用范围有限。

1.2.3.3 基于地表流水物理模拟分析的算法

基于地表流水物理模拟分析的基本思路是：根据流水从高到低的自然规律，顺序计算每一个栅格单元的汇水量，然后按汇水量单调增加的顺序，由高到低找出区域内的每一条汇水线；再根据得到的汇水线找出各自汇水区域的边界线，就得到了分水线。这种方法采用了 DEM 整体追踪分析的方法，分析结果的系统性比较完整，便于进一步进行相应的径流成因分析。但是由于汇水量的计算与高程相关，在汇水线两端的汇水量计算容易产生误差，并且得到的分水线是闭合区

域, 与实际的地形线不完全符合; DEM 存在的局部凹陷的和平坦的栅格单元, 也会导致无法确定水流方向^[30]。

O'Callaghan 和 Mark(1984)提出的坡面流模拟方法就是基于地表流水物理模拟分析的, 根据流水总是沿着斜坡最陡方向流动的原理, 确定 DEM 栅格单元的水流方向; 设置合适的水道给养面积阈值来确定沟谷。Jenson 和 Dominique(1988)、Martz 和 de Jong(1988)等许多研究者也采用了这种方法。

已有许多学者研究过如何处理 DEM 洼地处的水流流向问题。Band (1986)认为可以采用垫高填平的方法, 通过增加洼地处的高程值, 使单元上的水流不返回至原洼地单元, 而流到相邻的单元, 这种方法仅仅能有效的处理最简单的地形。O' Callaghan 和 Mark (1984)采取先对 DEM 数据进行平滑, 减少洼地的数目; 然后确定其它的洼地的面积范围和这些洼地周边坡面流的最低点, 则标示洼地单元的水流向那个最低点。这种方法可以比较好地处理简单的凹陷洼地, 却会使生成的水道发生扭曲, 而且不能有效的处理复杂洼地。同时对 DEM 数据的平滑处理减少了 DEM 所包含的信息内容。

作为地表流域特征信息的重要体现, 水流流向反映了地表径流在流域空间内的流向趋势, 并且可以根据它来计算出该趋势并将其以数字的形式表达出来。因此, 如何确定 DEM 栅格单元的水流流向, 也有许多学者通过在 DEM 数据建立 3X3 的局部窗口的方法进行了研究。研究发现, 在 DEM 数据建立 3X3 的局部窗口内, 中心单元网格中的水流流向有八种可能性, 即流向该单元网格周围相邻的八个网

格。目前根据流向的不同, 流向判断的方法为单流向法和多流向法。单流向法和多流向法都是将中心单元格上产生的径流都流向周围的相邻单元栅格, 但是前者是流向周围坡度的最低的相邻单元栅格, 后者则是按一定的比例流向周围若干坡度相对较低的相邻单元栅格。

单流向法有 D8 法、Rho8 方法、DEMON 法和 $D\infty$ 法等, 应用最广泛的是 D8 法。D8 法是通过最陡坡度法来确定水流的流向, 关键在于计算中心网格与周围相邻的八个网格间的距离权落差(即网格相邻点与中心点高程差除以与网格中心点之间的距离), 中心单元栅格的流出网格就是距离权落差最大的网格, 中心网格的流向就是从中心网格到距离权落差最大的网格的方向^[36]。

多流向法有 FDS 与 FRho8 算法。FDS 与 FRho8 算法是根据一定的规则或准则, 使水流按照一定的比例流向周围比它低的所有相邻单元格, 可以用来进行水流分流建模。但是, 这样会产生的汇流面积有可能是不连续的, 并且几何对称性和边界约束都会影响到计算结果。多流向算法虽然避免了偏向某个栅格, 但是增加了存储量^[36]。

1.2.3.4 基于地表几何形态分析和地表流水物理模拟分析相结合的算法

由于基于地表几何形态分析原理和基于地表流水物理模拟方法的算法都有一定的缺陷, 如流域沟谷线的不连续或与实际之前存在偏差, 因此, 有些学者提出了将这两种方法结合起来进行流域的提取并取得了一些成果。基于地表几何形态分析和基于地表流水物理模拟相

结合的这种方法的基本思路是：首先按照流水物理模拟的方法对 DEM 提取汇流网络；然后基于形态学方法对 DEM 进行扫描，提取所有形态上的沟谷；再根据得到的主汇水线改进上一步得到的沟谷段集，去掉其中孤立的沟谷段，达到在空间上连续，并使用标准的二进制细化方法细化沟谷段，得到单网格宽的汇流网络；最后基于形态分析提取沟沿线，并从流域出口处递归搜索整个汇流网络直至遇到沟沿线栅格单元，如此则得到完整的沟谷网络^[38]。

闫国年（1998）等基于地表几何形体分析和基于地表流水物理模拟，设计了完整的基于两者相结合的沟谷网络自动提取算法，以典型黄土沟壑丘陵区山西离石王家沟小流域的 DEM 数据为例，通过结果证实提取得到的沟谷网络与实际情况相符^[32]。

随着对 DEM 认识的深入以及 DEM 在许多领域的广泛应用，针对 DEM 提取算法中遇到的问题，探索出了许多有益的研究方法，取得了许多成就。基于地表几何形体分析和基于地表流水物理模拟相结合的方法成为目前流域沟谷网络提取中应用最广泛的。

1.3 研究内容

1.3.1 研究思路

本文在参阅国内外有关 DEM 流域提取算法的基础上，研究基于地形表面几何形态分析和基于地形表面流水模拟分析，以及存在的问题等，通过对各模型和算法的优缺点的比较，理解到这两种分析方法都存在一定的缺陷，因此采取了两者的混合方法。其算法思路

是：先用几何分析的方法根据特征点的特性从 DEM 数据中提取可能是山脊线和山谷线上点的候选点；然后再用流水模拟的方法提取流域特征信息；再采用缓冲区分析和叠置分析生成新的栅格图层，然后对沟谷进行二次提取以处理伪沟谷；最后用得到的流域沟谷提取其与等高线的交点，辅助已得到的山谷线候选点进行识别存在同一栅格单元内的山谷点、使该交点与识别出的山谷点归类合并，实现对流域特征信息的修正，使之更加符合实际地貌。

1.3.2 主要内容与章节安排

本文研究了 DEM 流域提取方面的算法，研究基于地形表面几何形态分析原理和基于地形表面流水模拟分析原理，对部分算法及实现进行了改进，以提高信息的提取效率和精度。

本文以 Windows XP 作为系统平台，以 Visual Studio 2008 作为开发环境，基于 ArcGIS9.3 平台完成各算法的实现和软件开发，实现基于数字高程模型的流域特征信息的提取。论文共分为以下七章：

第一章绪论，包括（1）研究目的与意义；（2）DEM 的研究现状与发展趋势；（3）本文的研究内容。

第二章研究材料与方法，包括（1）研究区概况，介绍岳麓区的地理位置、地质地貌特征、气候、土壤以及植被情况；（2）DEM 数据，介绍 DEM 的构建和 DEM 误差分析；（3）研究的基本方法，介绍了 DEM 的修正原理；流域的提取原理。

第三章算法设计与实现，包括（1）流域的自动提取算法的设计

思路, 介绍 DEM 数据的预处理; 特征点的提取; 流域的提取; 流域的修正; (3) 算法的实现, 包括实现功能和实现流程; (3) 算法验证; (4) 算法评估。

第四章总结和展望, 对研究取得的成果和存在的不足之处进行总结, 并对下一步的研究工作提出展望。

第二章 研究材料与方法

2.1 研究区概况

2.1.1 地理位置

岳麓山地处湖南省长沙市市内，东临湘江，包括岳麓山的主体部分和周围相关联的丘陵部分。本文中的研究对象是岳麓山的主体部分，面积约 6km^2 ，最低海拔 55m，最高海拔 299.9m，位于东经 $112^\circ54'55''\sim112^\circ56'37''$ ，北纬 $28^\circ10'5''\sim28^\circ12'22''$ ，南北长约 4km，东西宽约 3km。

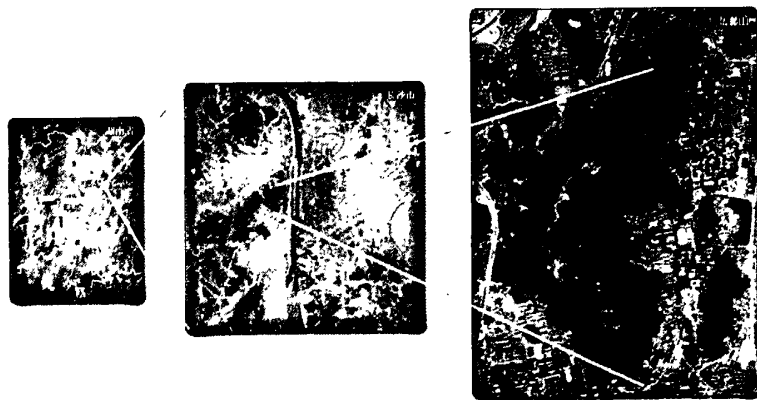


图 2-1 岳麓山的地理位置影像图（源于 GoogleEarth）

Fig.2-1 The location of yuelu Mountain

2.1.2 地质地貌特征

岳麓山脉主要是由扭动构造带中的新华夏系构造形成，并且地

貌格局基本上定型于强烈的燕山运动期的新华夏系构造体系,以拗陷和断裂为主,形成了东北—东南向的山脉延伸状况,这种上升运动,导致了侵蚀基准面的下降,为重力侵蚀、沟谷溯源侵蚀和流水侵蚀提供了有利的条件^[61]。

岳麓山山脉中部海拔高,向南北方向海拔逐渐降低;东坡坡面平缓,而东南坡坡面陡峻。岳麓山的山脊线以云麓峰为转折处,云麓峰以南沿东南—西北方向延伸,截止至左家垅的靳江口;云麓峰以北沿东北方向延伸,截止至溁湾镇。岳麓山的山体呈半月弧型,凹面朝向长沙市。

岳麓山的地层岩石主要是晚新元古代板溪群紫红色板岩和一些古生代(主要为泥盆系)的砂岩,以及一些第四系沉积物^[41]。第四系沉积物是长沙地区广泛发育的地质单元,主要岩性为砂砾岩、砂岩、砾岩和砂页岩组成的混杂堆积,为河流相沉积物^[59]。

2.1.3 气候

岳麓山属亚热带季风性湿润气候,春温多变,夏秋多晴,全年严寒期短,暑热期长,年平均温度 17℃,7 月份平均气温 28.6℃,极端高温 40.6℃,一月份平均气温 4.6℃,极端低温-11.3℃,与长沙市中心地区相比,山中平均气温低 4-6℃,季节往往推迟 10d 左右,年平均降雨 1200~1400mm^[59]。

2.1.4 土壤和植被概况

岳麓山的主要土壤类型为丘陵山地红壤。岳麓山属于亚热带季风区，它的红壤的发育因为受到亚热带高温多雨气候的影响，土壤中盐基大多淋失，具有深度的富铝化特性。

岳麓山植物种类丰富，主要以典型的亚热带常绿阔叶林和亚热带暖性针叶林最具代表性，区系属于华中地区，地带性科则以常绿阔叶树（如壳斗科、冬青科、樟科、山矾科和松科等）占优势，森林总面积达到 533.33hm^2 ，森林覆盖率为 96%^[41]。

2.2 DEM 数据

随着计算机技术水平的提高，DEM 的数据获取方法、储存和处理速度等都已经取得了很大的进展。DEM 已经成为标准的数据基础产品，取代了传统等高线地形图。DEM 是一种典型的数字化产品，具有广泛的实用价值^[28]。

2.2.1 数据采集

构建数字高程的重要基础是采集地表数据。目前多种数据采集方法都有其各自的优缺点。按照标准格网通过模拟航测仪器采集高程值的方法，由于模拟航测仪器比较昂贵，成本过高而不适合在生产中应用；通过全数字摄影测量工作站按照像片上的规则格网、利用数字影响匹配，既可以利用高程值直接求解的影响匹配方法，也可以按照模型上的规则格网进行数据采集，但是其所得的结果需要进行编辑才能得到较好的精度，生成 DEM 时只有任意采集地貌特征点线，才能保

证 DEM 的高精度。特别是在平坦地区、森林覆盖地区或者房屋密集的城区, 仍然需要人工干预和手工编辑。因此, 选择 DEM 合适的采集方法不仅要考虑目的需求, 而且考虑到精度要求, 还要考虑到设备条件、经费条件等多方面因素。因此本文使用地图扫描矢量化方法将地形图中的等高线矢量化, 用不规则三角网法插值生成标准格网的 DEM 数据的方法与前两者相比具有精度、效率和成本上的优势, 所以本研究采取该方法构建数字高程模型^[22]。

本研究中主要探讨 DEM 在地形破碎程度较高岳麓山地形分析应用, 因此, 主要考虑水平分辨率对 DEM 精度的影响。

2.2.2 DEM 数据

本文所采用的原始数据是四幅 1:5000 纸质地形图。这四幅 1:5000 纸质地形图是 2000 年航摄 2003 年成图的, 采用高斯-克里格投影, 西安 1980 坐标系, 1958 国家高程标注, 图幅号分别为: H49H183158、H49H183159、H49H184158、H49H184159, 其中 H49H183158 和 H49H184158 两幅图幅等高距为 5 米, H49H183159 和 H49H184159 两幅图幅等高距为 2.5 米。

本文所采用的 DEM 是高丽芳 2007 年构建的。其构建的具体方法是: (1) 首先, 利用高精度扫描仪逐一将纸质地形图扫描, 分辨率为 300dpi, 扫描后以 tiff 格式存储为 CNKI 图像文件, 然后输入到计算机, 保存扫描的数字化栅格地图。(2) 利用 ArcGIS 软件对扫描到的数字栅格图像上的等高线和离散高程点进行手工矢量化。矢量化完

毕之后，在等高线矢量图层上设置控制点，输入其坐标值，把坐标转换为大地坐标，拼接分幅图层。拼接完成之后，检验等高线是否正确连接，高程值是否正确匹配（3）利用 3D 分析模块，根据等高线数据和离散高程点数据生成不规则三角网（TIN）模型，然后设置采样间隔为 2.5 米，将不规则三角网（TIN）模型转化为比例尺为 1:5000 的规则格网（GRID）数字高程模型（见图 3-1）。

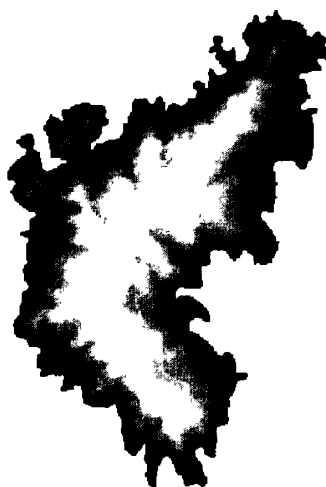


图 2-2 岳麓山的数字高程模型
Fig 2-2 the DEM data of yuelu Mountain

2.2.3 数据精度分析

DEM 的精度主要是指 DEM 精度的总体状况以及 DEM 是否与实际地形相吻合。本文使用任意点法进行精度评定，其数学公式 2-1：

$$R = \frac{1}{n} \sum_{k=1}^n (R_k - Z_k)^2 \quad \text{公式 2-1}$$

公式 2-1 中，R 是指 DEM 的中误差；n 是指抽样检查点数； Z_k ($k=1, 2, \dots, n$) 是指检查点的高程真值； R_k ($k=1, 2, \dots, n$) 是指通过 DEM 内插法得到的检查点的高程值。

根据附表 1:5000 比例尺 DEM 误差统计表可以看出,根据样区所建立的 DEM 中,抽样检查点的高程中误差比相应比例尺地形图的 2/3 等高距小,DEM 精度达到了国家规定的标准^[60]。

2.2.3 数据等高线分析

对 DEM 所提取等高线的分析,是诊断 DEM 数据误差和 DEM 对地形结构表达的有效工具,具体方法就是用目视解译与原始地形图对比,发现 DEM 生成的等高线图上的不合理部分,从而找出数据中的误差。经过分析,DEM 的质量较好,符合进行研究的精度要求。

2.3 研究基础方法

基于数 DEM 的地形特征进行提取的,通常采取两种基本的方法:

(1) 局部窗口扫描法; (2) 地形表面流水模拟法。

局部窗口扫描法是设置一个合适的矩形窗口来扫描 DEM 矩阵,通过比较矩形窗口内中心栅格单元与周围相邻的栅格单元的高程值来判别是否为洼地,并将位于洼地内的那些栅格单元标记,由这些标记的栅格单元组成水系。这种方法产生的水流线不连续,需要来把它们连接起来,通过修剪和细化产生合理的流域沟谷体系。

地形表面流水模拟法是通过模拟地表径流按照从高到低的流动来生成流域。这种方法根据 DEM 矩阵的中心栅格单元和周围相邻的八个单元格之间的最大坡降来确定水流方向,然后根据水流方向计算每个栅格单元的流量,设置合适的阈值对 DEM 矩阵进行筛选,由通

过筛选的栅格单元组成水系。这种方法产生连续的河流线段。所以目前采用这种方法来提取水系得到广泛的应用。

2.3.1 DEM 的修正原理

由于在 DEM 生成过程中必然存在误差以及 DEM 的不同分辨率的影响,因此 DEM 中洼地的存在是不可避免的,从而造成在计算水流方向以及生成流域特征信息时会出现水流逆流或断流的情况,给后面进行的水流追踪带来麻烦,并致使产生的河流线段不连续以及流域其他特征信息的不准确。因此,在提取流域特征信息之前要先对生成的 DEM 进行修正,其中包括对存在的洼地和平坦区域进行判别和消除。

2.3.1.1 闭合洼地和平坦区域的处理方法

DEM 中存在闭合洼地或平坦区域,一般有两种看法:第一种看法认为闭合洼地或平坦区域都是真正的地表形态;第二种看法认为闭合洼地或平坦区域是假的地形。目前的许多处理方法都认同第二种看法,通常利用一种光滑方法对可能存在问题的地形予以消除,而剩下的没有问题的地形予以除去或调节。其具体过程分为两步:第一步,采用垫高填平来矫正 DEM,即增加闭合洼地内网格单元的高程值,使该网格单元的高程达到最低出口网格单元的高程,这样以来就可以用平坦区域代替填满所有的闭合洼地。第二步确定平坦区域的所有网格单元(包括由洼地填平的平坦区域和原 DEM 中存在的平坦区域)

的流向。这种方法的既对到有问题的区域予矫正消除,也对没问题的 DEM 的网格单元以矫正消除,导致了数字高程信息量减少,使得流域沟谷提取时与实际沟谷不完全拟合。

2.3.1.2 平坦区域流向及河道的处理方法

目前,在确定平坦区域流向的方法中最广泛使用的是由 Jenson 和 Domingue 提出的。这种方法首先以有流向的网格单元为中心,寻找出与之相邻的那些网格单元,然后确定这些相邻的网格单元的流向。重复执行这一步骤,最后确定所有的平地网格单元的流向为止。这种方法将流路约束在平坦区域内,致使有可能出现多出口的情况^[35]。这种方法通常会生成不真实的平行河道。

2.2.3 流域的提取原理

2.2.3.1 利用“高于”的概念确认沟谷

最初确定沟谷的方法是根据地形表面的几何形体进行分析的,其具体方法是在 DEM 中依次选择某一单元为中心单元,比较与其周围的相邻单元的高程,如果周围的某些相邻单元的高程大于该中心单元的高程,则确定该单元为谷地单元,继而通过确定谷底单元来进行的。这种方法利用了中心单元与周围的某些相邻单元之间高程关系。根据选择周围的某些相邻单元不同,又分为两种:第一种是通过比较某单元的周围的四个直线方向(南、北、东、西)的单元的高程与其本身的高程来确定该单元是不是谷地单元;第二种是在此基础上增加了与

对角线方向（东北、西南、西北、东南）的四个单元的高程比较来确定该单元是不是谷地单元^[35]。

这种方法生成的谷地单元通常是不连续的，而且生成的谷线也比实际谷线要向上坡延伸得太远。为了解决这个问题，从水文学中引进了坡度阈值的概念，同时在考虑中心栅格单元周围的八个相邻单元的基础上还考虑到每一个侧面更远的单元^[35]。

2.2.3.2 利用水流方向信息识别沟谷

随着基于地形表面水流模拟方法的产生，基于地表水流沿着坡度最大方向流动原理的连续沟谷生成法由于其水文意义上进行分析，得到了非常广泛的应用，并且被认为是最普遍实用的方法^[35]。

这种连续沟谷生成的方法分为三步；首先确定每一个栅格单元的水流方向，即水流流出单元最大坡度的方向；其次利用水流方向统计汇集到每一个栅格单元的所有栅格单元的总数；最后选择一个合适的汇水面积阈值，将统计到的总数不小于阈值的所有栅格单元都作为沟谷的一部分。这种方法必须假定所生成的沟谷是以地表完全不渗透水、无植被区域上空间降雨均匀分布为前提。这个假定非常重要，因为自然界中地表基本上没有完全不渗透水和无植被覆盖的区域，所以根据这种假定生成的沟谷很大可能是于野外计算不相同的，但是如果广泛应用到水文模拟的过程中这种自动提取的沟谷就具有重要意义 (Tribe, 1992)^[35]。

这种连续沟谷生成的方法的缺点也比较明显。首先不能应用于平

坦的区域，如平原等。其次，在相对比较平坦的谷地也往往会出现平行河道现象，即形成“伪河道”，现实中这种平行河道是不存在的。再者，这种方法采用了汇水面积阈值的概念，根据不同的阈值将得到不同的沟谷，选择合理汇水面积阈值形成稳定沟谷的还有待于进一步研究^[35]。

第三章 算法设计与实现

3.1 算法设计

3.1.1 设计思路

本文通过研究基于地形表面几何形态分析原理和基于地形表面流水模拟分析原理，理解到这两种分析方法都存在一定的缺陷，因此采取了两者相结合的混合方法，其算法思路是：先用几何分析的方法根据特征点的特性从 DEM 数据中提取可能是山脊线和山谷线上点的候选点；然后再用流水模拟的方法提取流域；再采用缓冲区分析和叠置分析生成新的栅格图层，然后对沟谷进行二次提取以处理伪沟谷；最后用得到的流域的汇水线辅助对已得到的山谷线候选点进行识别、归类，顺序提取各条流域的汇水线。

3.1.2 DEM 数据预处理

在对 DEM 数据进行流域提取之前，先检查 DEM 数据中是否存在空值数据。如果存在空值数据就无法对 DEM 数据进行洼地填充，要先对栅格空值数据赋值。

洼地是 DEM 中那些被周围具有较高高程的栅格单元所包围的局部地形单元。在 DEM 数据中的洼地，大多是由于数据获取和内插处

理时的误差所造成的。洼地的存在会在地形分析中导致不正确的结果，如导致流域网络和地形结构线的间断或不正确连通，成为了对 DEM 分析的障碍。所以在进行分析之前要对 DEM 进行洼地填充。

洼地是水流方向不合理的地方，可以通过水流方向来判断哪些地方时洼地，并进行填充。但是并不是所有的洼地都是由于数据存在的误差所造成的，有一些洼地是地貌形态的真实反映，所有在进行洼地填充之前，要先对洼地进行填充。洼地填充有两种方法：第一是先提取洼地，计算洼地深度，然后通过纸质地形图进行对比分析，判断哪些地区是由于数据的误差所造成的，而哪些地区是真实的地貌形态的反映，从而根据洼地深度确定合适的填充阈值对洼地填充；第二是对所有洼地完全填充。填洼时采用“高程增量迭加法”进行处理，通过增高洼地栅格单元的高程值，使得 DEM 中每一个栅格的高程值不比周围八个栅格中的高程值中至少一个低，然后就可以计算水流方向，进行小流域提取。

流程见图 3-1。

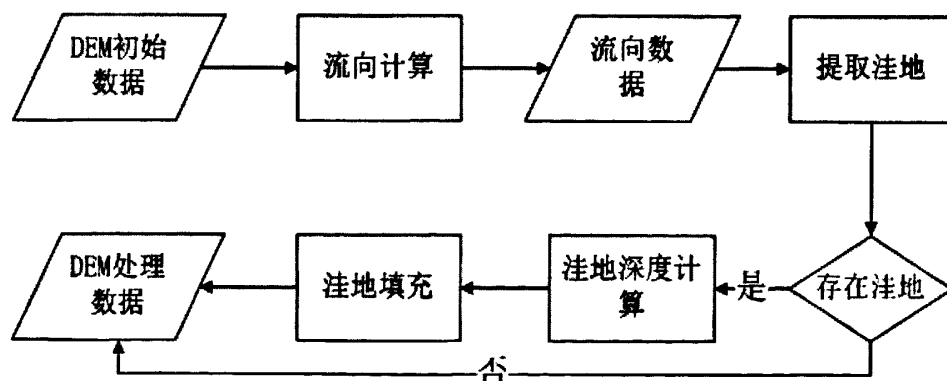


图 3-1 DEM 数据预处理流程图

Fig.3-1 the flow chart of preprocessing

3.1.3 特征点提取

地形特征点主要包括山顶点、凹陷点、脊点、谷点、鞍点，平地
点等。本文采用了局部窗口法，其具体方法为：设置当前中心点行列
号为 (i, j) ，以该中心单元为中心，扑捉上下左右四个单元；如果该
中心单元扑捉到的上下左右四个周围单元没有高程为 0 的，就可以建
立一个 $(2m+1) \times (2m+1)$ 的局部窗口，并且在这个局部窗口内，
根据中心单元的高程与其周围单
元高程之间的关系，判断该中心
单元是否为地形特征点；如果该
中心单元扑捉到的 8 个周围单元
存在着高程为 0 的，则认为该中
心单元为研究区域的边界点，不

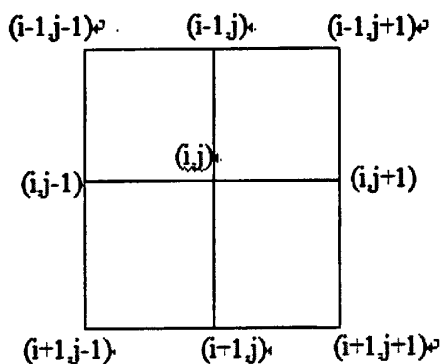


图 3-3 窗口示意图

Fig 3-3 the schematic for the windows

是所要获取的地形特征单元点。

假设有一个如图 3-3 窗口示意图所示的 3×3 窗口。则：

如果 $(Z_{i,j-1} - Z_{i,j})(Z_{i,j+1} - Z_{i,j}) > 0$

(1) 当 $Z_{i,j+1} > Z_{i,j}$ 则 $VR(i, j) = -1$

(2) 当 $Z_{i,j+1} < Z_{i,j}$ 则 $VR(i, j) = 1$

如果 $(Z_{i-1,j} - Z_{i,j})(Z_{i+1,j} - Z_{i,j}) > 0$

(3) 当 $Z_{i+1,j} > Z_{i,j}$ 则 $VR(i, j) = -1$

(4) 当 $Z_{i+1,j} < Z_{i,j}$ 则 $VR(i, j) = -1$

如果 (1) 和 (4) 或 (2) 和 (3) 同时成立，则 $VR(i, j) = 2$

如果以上条件都不成立，则 $VR(i, j) = 0$

其中，

$$VR(i, j) = \begin{cases} -1, & \text{表示谷点} \\ 1, & \text{表示脊点} \\ 2, & \text{表示鞍点} \\ 0, & \text{表示其他点} \end{cases}$$

流程见图 3-2。

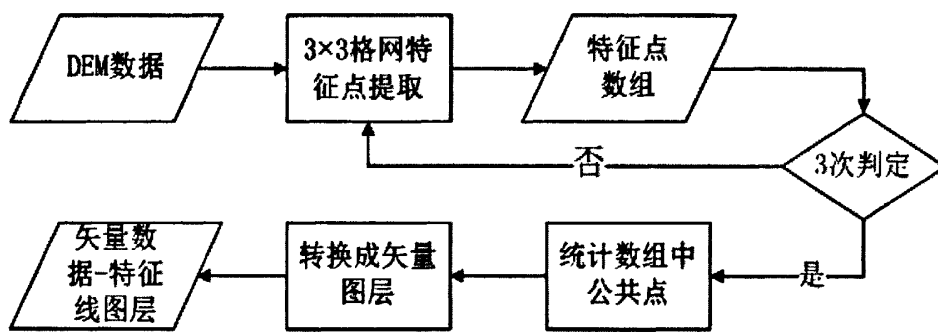


图 3-2 特征点提取流程

Fig.3-2 The flow chart of the feature points

3.1.4 坡度和坡向的提取

坡度和坡向是指局部地表坡面在空间上的倾斜程度和朝向。

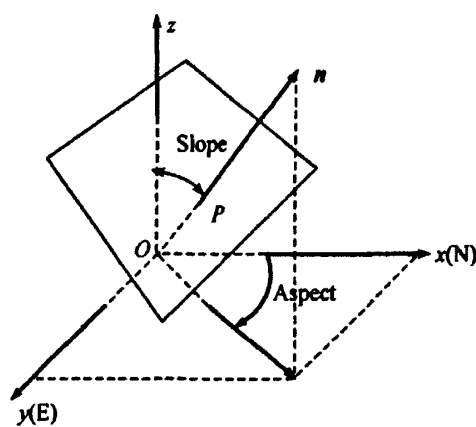


图 3-3 地表单元坡度和坡向

Fig.3-3 the slope and the aspect

严格来讲，地表面任意一点的坡度是指过该点的切平面与水平地

面的夹角（见图 3-3）。坡度以向量的方式表示了地表面在该点的倾斜程度。在进行坡度计算时，通常采用简化的差分公式：

$$\text{Slope} = \arctan \sqrt{f_x^2 + f_y^2} \times 180/\pi$$

公式 3-1

公式 3-1 中， f_x 为 x 方向高程的变化率， f_y 为 y 方向高程变化率。

在实际应用中，坡度有两种表示方式一坡度和坡度百分比。坡度是指水平面与地形面之间的夹角。坡度百分比是指高程增量与水平增量之间的百分数。

坡向的定义是指地表面上的任意一点的切平面的法线矢量在水平面的投影与过该点的正北方向的夹角（见图 3-4）。坡向表征了该点高程值该变量的最大变化方向。其数字表达公式为

$$\text{Aspect} = \arctan\left(\frac{f_y}{f_x}\right)$$

公式 3-2

公式 3-2 中， f_x 为 x 方向高程的变化率， f_y 为 y 方向高程变化率。

3.1.4 小流域提取

水流模拟方法的基本思想是按照流水从高到低的自然规律，顺序计算每一栅格点的汇水量，然后按汇水量单调增加的顺序，由高到低找出区域中的每一条汇水线；根据得到的汇水线，通过计算找出各自汇水区域的边界线，就得到了分水线。

流程见图 3-5。

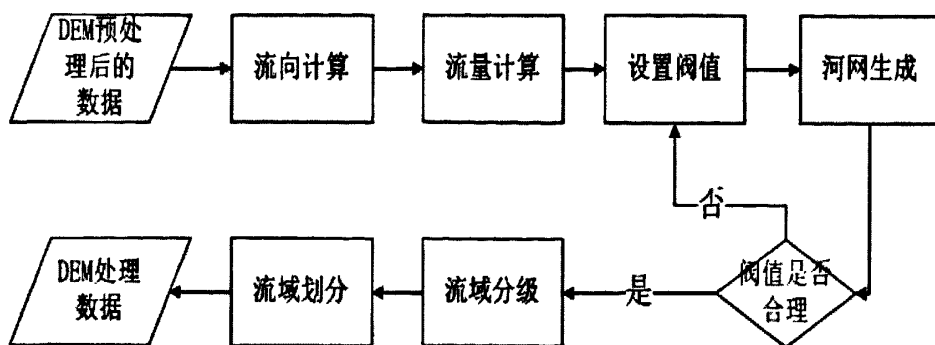


图 3-5 流域提取流程

Fig.3-5 The flow chart of the watershed

3.1.4.1 流向统计

流域是气候和下垫面各种自然地理因素综合作用的产物，下垫面主要包括地貌、地质、植被、湖泊和沼泽等。本文认为虽然多流向水流模拟方法比单流向水流模拟方法更能模拟地表水流的运动，但是单流向模拟方法更能突出表现流域的河谷网络。

中心网格与周围相邻的八个网格间的距离权落差(即网格相邻点与中心点高程差除以与网格中心点之间的距离)，中心单元栅格的流出网格就是距离权落差最大的网格，中心网格的流向就是从中心网格到距离权落差最大的网格的方向

水流方向的确定采用 D8 算法，计算 DEM 每一中心栅格单元与其周围相邻的八个单元之间的坡度，水流方向是指中心栅格单元流向周围相邻的坡度最陡的栅格单元的方向。其具体方法是：如果在 3×3 的局部窗口里，设中心格网为 c ，其流向（这里指水流的流出方向）在其相邻八个格网单元 $i(i=1, 2, \dots, 8)$ 中选择， i 满足条件：

$$\max \{ (k \times (Z_c - Z_i)) \}; i=1, 2, \dots, 8 \quad \text{公式 3-3}$$

公式 3-3 中,当 i 位于东西或南北直线方向时, $k=1$; 当 i 为四个对角线方向时, $k=1/\sqrt{2}$ 。并且 i 接受 c 的全部流量, 即流向是间隔 45 度的八个可能格网的方向之一, 中心网格单元的流量全部流入最陡(下降)方向上的下游网格单元中。

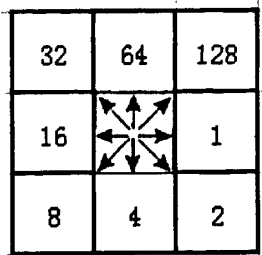


图 3-6 流向的数字表示

Fig.3-6 The mathematic of the flow direction

获得流向以后, 为了确定单元格的水流流向, 对每个方向赋以代码表示(见图 3-6), 标示的数值以右侧作为起点逆时针开始以 1 开始, 以 2 的倍数递增, 分别标示对应的流向。若单元内的水流流向矩阵的左边, 它的水流方向定为 16。

3.1.4.2 流量统计

流量是指每一个栅格单元的流水累积量。在水流模拟方法中, 流量是基于水流方向数据计算得到的。其基本思想是: 假设 DEM 中高程大于零的每一个栅格单元都有一个单位的水量, 首先按照高程从大到小的顺序对 DEM 中高程大于零的栅格单元排序; 然后排序后的 DEM 栅格单元依次根据水流沿着高到低的自然规律, 根据相对应栅格单元的水流方向(见表 3-1), 计算当前栅格单元所流过的水量数值, 即可以得到该栅格单元的流量(见表 3-2)。在此假设模拟前每点的水

量为 1 个计量单位，如果该点的水流没有流出，那么该点的流量为 0。

表 3-1 流向统计

2	2	2	4	4
2	2	2	4	4
1	1	2	4	8
128	128	1	2	4
128	128	1	1	4

表 3-2 流量统计

0	0	0	0	0
0	1	1	2	1
0	3	8	5	2
0	1	1	20	0
0	0	0	1	24

3.1.4.3 沟谷生成

根据水流从高处流向低处的自然规律，当水流的汇流量达到一定界限的时候，就会在地表产生地表径流，所有这些汇流量超过这个临界值的栅格单元就是潜在的水流路径，而这些水流径流所构成的网络结构就是沟谷。

在生成沟谷之前预先设定一个阈值，该阈值即临界值表示流域中点的最小积水面积，将流量数据高于此阈值的网格连接起来，便可形成流域沟谷网络。当阈值变少时，网络的密度便相应增加。如同计算每个网格的流向一样，流域沟谷网络同样从流域的汇水量最大的那些栅格单元开始向上游搜索。汇水量最大的那些栅格单元即是流域的终结点，该点可以标记为 1 号流域沟谷网络的起点，找流入该点的所有网格，如果其流量大于阈值则加入水系网中，如此直至某一个网格有不止一个流量大于阈值的流入网格，或者没有流量大于阈值的流入网

格为止。如果有两个或者两个以上流量大于阈值的流入网格，那么这个编号的河道结束，则可以让每个流量大于阈值的流入网格开始标记新的河道编号。如果没有流量大于阈值的流入网格，那么该编号的河道结束。当所有编号的河段的最后一个网格都没有流量大于阈值的流入网格时，此时流域沟谷提取完毕。

沟谷阈值的设置是满足三个方面的标准的：(1) 在三条以上的等高线弯曲之处，应该有流域沟谷；(2) 流域沟谷能够尽量接近源头(3) 流域沟谷的源头不能够进入到 0-15 度的坡度区域内。

3. 1.4.4 沟谷分级

沟谷分级是对一个线性表示的沟谷网络以数字标识的形式来划分级别。不同级别的沟谷所代表的汇流量不同，级别越高，汇流量越大，一般是主流，而级别比较低的沟谷一般是支流。通过对沟谷的分级，把沟谷的主流与支流区分开来。

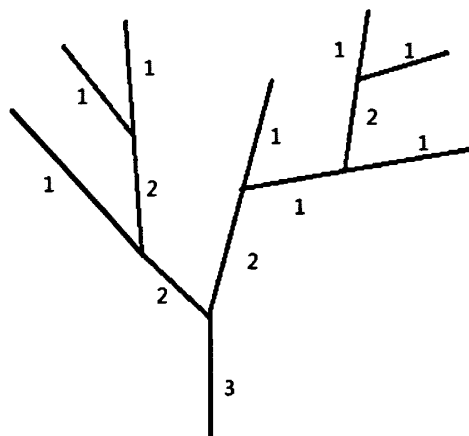


图 3-7 Strahler 河流分级方案

Fig.3-7 The Strahler stream classification scheme

本文采用 Strahler 于 1953 年提出的分级方案（见图 3-7），把沟谷段划分为四级，即所有的外部沟谷段（这里指没有其他沟谷段汇合的沟谷段）统计为第一级；两个第一级的沟谷段汇合而形成的新的沟谷段为第二级；依次，两个第三级的沟谷段汇合而形成的新的沟谷段为第四级；如果级别较低的沟谷段汇合到级别较高的沟谷段，那么其级别不变。

3.1.4.5 流域划分

流域是指流经其中的水流和泥沙等其他物质从一个公共出水口排出而形成的一个集中的排水区域，又被称为集水区域。流域能够显示每一个汇水区域的大小。出水口是流域内水流排除的出口，是流域的最低点。流域之间通过分水岭划分开。分水岭包围的区域称为一个水系的流域，流域分水岭所包围的区域面积就是流域的面积。

要想划分流域，先要确定出水口，即整个流域的最低点，然后结合水流方向，分析出这个出水口上游的所有流过该出水口的栅格，一直到分水岭的位置为止。这样计算的整个区域就是流域。

3.1.5 流域修正

流程见图 3-8。

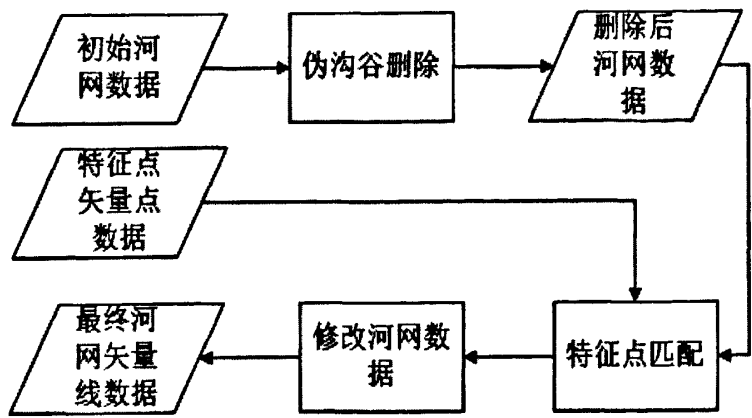


图 3-8 流域修正流程

Fig.3-8 The flow chart of the watershed correction

3. 1.5.1 伪河谷删除

由于沟谷提取是采用的最大坡降的方法，在谷底区域的水流方向存在随机的可能性，容易生出平行状的沟谷，这种平行状的沟谷称为伪河谷。伪河谷的存在使生成的沟谷与实际沟谷相比存在误差，应该被删除。通过研究发现：在选择合适的阈值以后，并没有发现伪沟谷，而在选择比该合适阈值较小的阈值时，产生的伪沟谷数目比较少，可以选择了人工进行伪沟谷的删除。

本文也对伪河谷自动删除进行研究，其具体方法是：首先把上一步骤中生成的采取 Strahler 三级分级方案的沟谷中的第二级别和第三级别的沟谷标记为主沟谷；然后对标记的主沟谷进行缓冲区分析；其次对缓冲区范围内的栅格单元降低一定高程值，缓冲区范围外的栅格值赋为零，然后生成新的栅格图层；再对生成的新的栅格图层与原始 DEM 进行叠加分析，对叠加后生成的 DEM 重新进行小流域提取，小流域提取后的沟谷数据就是进行伪沟谷删除的结果。

3.1.5.2 洼地匹配

获得区域内沟谷后,通过其与 DEM 预处理提取的洼地比较,发现:(1)山谷点中有一部分在洼地填充后的栅格单元的中心,即洼地填充后的栅格可能就是山谷点的候选点;(2)提取的沟谷线会经过填充的洼地栅格单元。

针对第一种情况,首先读取提取的洼地图层中的洼地栅格单元,根据该栅格单元的行列号来计算其所在的栅格位置(即从左到右、从上到下排列,该栅格是第几个),记录所有的洼地栅格位置到洼地栅格数据集,然后根据山谷点图层中的位置属性字段对山谷点图层进行筛选,删除山谷点图层中与洼地点坐标相同的山谷点。即如果改洼地栅格数据集中的数值中有与山谷点图层的位置属性字段值相等的数据,认为该点是洼地填充过的,而不是真正的山谷点,删除该点。

针对第一种情况,首先把洼地栅格图层转换成矢量点图层,然后获取沟谷线图层上的每一条沟谷的转折点,对该转折点做缓冲区处理,设置缓冲区半径为栅格单元长度的一半,在转换的矢量点图层中搜索是否存在被该缓冲区范围包含的点,如果点存在说明该转折点是洼地填充过的,删除该点,但是如果该搜索到的点是沟谷线的起点或终点,不删除该点,以免使沟谷线断裂。

3.1.5.2 特征点匹配

获得区域内沟谷后,首先对等高线进行从低到高排序;然后读取该沟谷中的第一条沟谷线,在排序后的等高线图层中搜索与该沟谷线

相交的各条等高线,然后分别求出该沟谷线与相交的各条等高线的交点;依次对交点做缓冲区处理,设置缓冲区半径为栅格单元长度的一半;在洼地匹配后的山谷点图层中搜索被该缓冲区范围包含的点集,在该点集中找出交点最近的一个山谷点,然后在沟谷线上添加该山谷点,如果没有,则对下一个交点进行处理;该沟谷线与相交的各条等高线的交点处理完成以后,依次处理该沟谷中的下一条沟谷线直至沟谷中的所有沟谷线处理完成。

3.2 算法实现

本文是以 Windows XP 作为系统平台,以 Visual Studio 2008 作为开发环境,以 CShape 语言为开发语言,利用 ArcEngine 组件对 ArcGIS 的功能进行扩展,基于 ArcGIS9.3 平台完成算法的实现和软件开发,简化了 ArcGIS 水文分析方法的操作,增加了流域水系与特征点的验证,实现基于数字高程模型的流域特征信息的提取,提高了流域数据的准确性。该模块既可以作为相关软件的功能扩充模块,也可以作为一个独立的 WINDOWS 应用程序。以岳麓山的高分辨率的 DEM 数据为数据源进行了验证,运行效率以及提取的流域特征信息的精度都有了很大的提高。

3.2.1 实现功能

算法设计的目的是在 DEM 数据支持下,生成小流域水系提取,主要应用功能包括以下几方面:

(1)特征点提取功能。通过局部窗开发对 DEM 栅格单元进行比较,判断其是否为山脊点、山谷点,并保存成矢量数据。

(2)小流域提取功能。通过对 DEM 栅格单元的流向统计、流量统计,生成沟谷,并进行沟谷分级,从而划分流域。

(3)流域修正功能。删除伪河谷、通过前面提取的特征点对流域沟谷进行修正,使之与实际地貌更加符合。

3.2.1 实现流程

程序的流程(见图 3-9)包括四个部分:第一部分是原始 DEM 经过数据的预处理,主要包括 DEM 数据洼地填充处理;第二部分是地形特征点的提取,主要包括局部窗口法的计算、赋值和特征点矢量数据的生成算法;第三部分是沟谷生成,主要包括流向、流量统计;第四部分是特征线的修正。

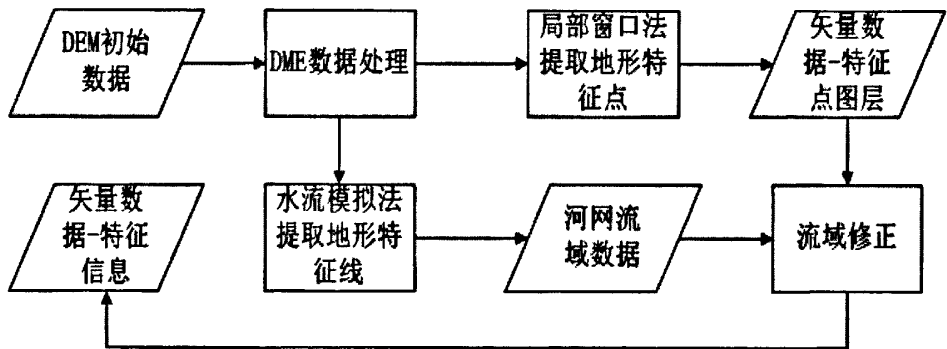


图 3-9 程序流程图

Fig.3-9 The flow chart

该程序的输入数据包括 mxd 文档(加载 GRID 型的 DEM 和等高线矢量数据),在进行流域提取时设置流域提取的控制参数阈值。最后程序可以直接将沟谷网络等特征信息输出到程序显示界面上,也可以以栅格或矢量的形式输出到文件中。

3.3 算法评估

3.3.1 原始数据

原始数据包括 1:5000 的 DEM 数据（属性见表 3-3）和等高线数据。其中等高线数据中共有 631 条线段。

表 3-3(1) 岳麓山流域的原始数据 DEM 的属性

DEM	行	列	横向行间距(米)	纵向行间距(米)
属性	242	323	11.14	11.14

表 3-3(2) 岳麓山流域的原始数据 DEM 的属性

DEM	X 最小值(米)	X 最大值(米)	Y 最小值(米)	Y 最大值(米)
范围	38393611	38396306.88	3117839.006	3121437.226

表 3-3(3) 岳麓山流域的原始数据 DEM 的属性

DEM	格式	高程最小值(米)	高程最大值(米)	平均高程(米)
高程	GRID	54.98077393	299.0515137	145.5422024

3.3.2 小流域提取

3.3.2.1 流域阈值

先计算出所有栅格流量的平均值为 145.54，然后分别在该平均值的两侧分别选取两点 60、100 和 200、250 为阈值，通过这五个选取的阈值来生成沟谷，然后统计分析沟谷参数信息，分析其阈值与流域总长度之间的关系，利用统计分析的方法来选择合适的阈值。

表 3-4 不同阈值时沟谷参数

分级	60	100	145.54	200	250
1 级(条)	177	107	72	50	40
2 级(条)	76	48	26	18	12
3 级(条)	18	4	0	0	0
总长(km)	27.291	19.840	15.279	12.550	10.677
流域面积(km ²)	4.151	3.913	3.780	3.524	3.412
河网密度(km/km ²)	0.000	0.000	0.000	0.000	0.000

从表 3-4 可以看出, 设置阈值的大小与沟谷总长度存在一定的相关性。通过对阈值与沟谷总长度进行线性函数、对数函数、多项式函数、乘幂函数进行分析, 发现两者的相关性与乘幂函数 $Y=404115 \cdot X^{-0.557}$ 的拟合度最好, R^2 的值达到 0.9995 (见图 3-10)。由此可以算出, 当阈值设置为 145.54 时, 生成的沟谷与实际沟谷的长度比较接近, 沟谷密度较近似, 阈值较合理。

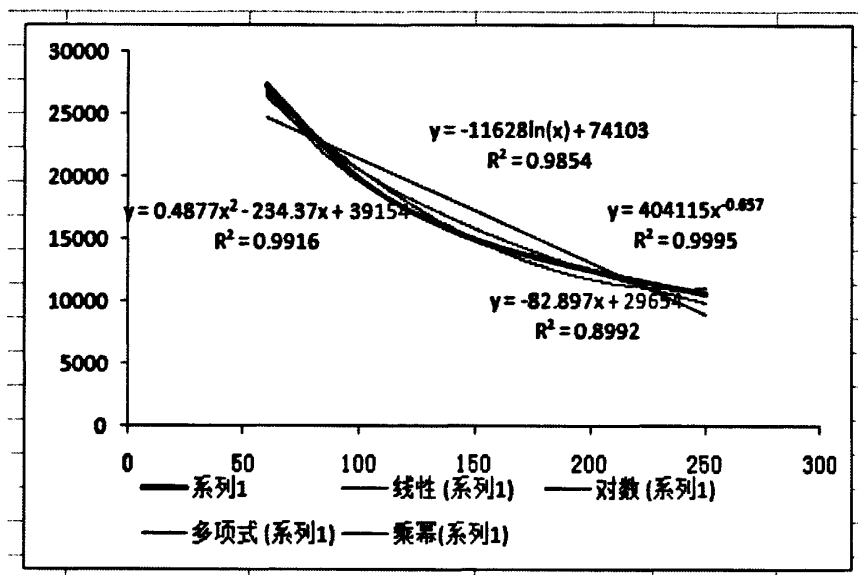


图 3-10 相关性函数

Fig 3-10 the ranking function

3.3.2.2 沟谷网络

由于算法实现是基于 ArcEngine 组件进行的, 因此, 设置阈值为 145.54 以后, 与手工提取流域沟谷相比, 生成的流域数据的精度达到了同样的程度。

流域数据包括沟谷分级图层和流域区域图层。沟谷分级图层 (见图 3-11) 用矢量线数据表示, 共有 98 条沟谷线。流域区域图层 (见

图 3-12) 用矢量面数据表示, 共有 128 个集水区。

表 3-4 结果数据统计量

特征信息	数据	统计量
沟谷流域	沟谷沟谷图层	98
	流域区域图层	128

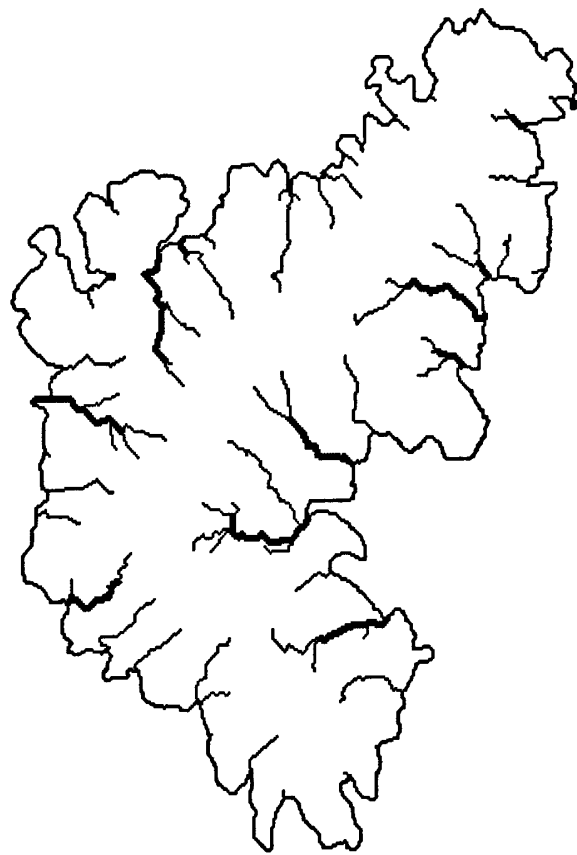


图 3-11 沟谷分级图层

Fig 3-11the valley classified layer

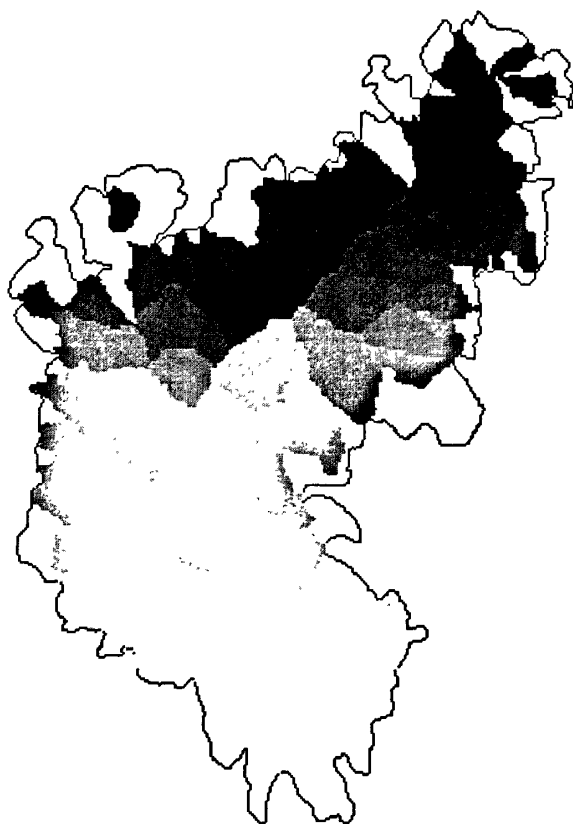


图 3-12 流域区域图层

Fig 3-12 the watershed region layer

3.3.3 流域修正

通过与手工提取的流域数据相比,发现算法实现流域提取的流域数据的精度达到了同样的程度,同时也还存在一些问题,尤其是在高精度 DEM 数据基础上,如洼地对沟谷仍然存在较显著的影响,沟谷线有明显的扭曲。因此,在小流域提取以后,对流域沟谷网络进行了修正处理。

3.3.3.1 伪沟谷删除

由于利用高精度 DEM 对小流域进行提取,在设置阈值为 145.54

后,生成的流域内并没有出现伪沟谷。在采用阈值为 100 时,小流域伪沟谷也仅仅有一条。通过对阈值为 100 时提取的沟谷网络进行伪沟谷删除处理,设置缓冲区半径为五个栅格单元长度,对缓冲区范围内的栅格单元降低高程 50 米,缓冲区范围外的栅格值赋为零,然后生成新的栅格图层,再与原始 DEM 进行叠加分析,对叠加后生成的 DEM 重新进行小流域提取。该方法可以有效地消除那一条伪沟谷,但是数据量过少,无法对该方法做出有效的验证。

3.3.3.2 洼地匹配与特征点匹配

通过对 DEM 的洼地提取(见图 3-13),计算洼地深度,得到洼地的最小值 0.028 米和最大值 9.121 米。在此,本文假设所提取到的洼地都不是真实的地表形态,对所有洼地填充。

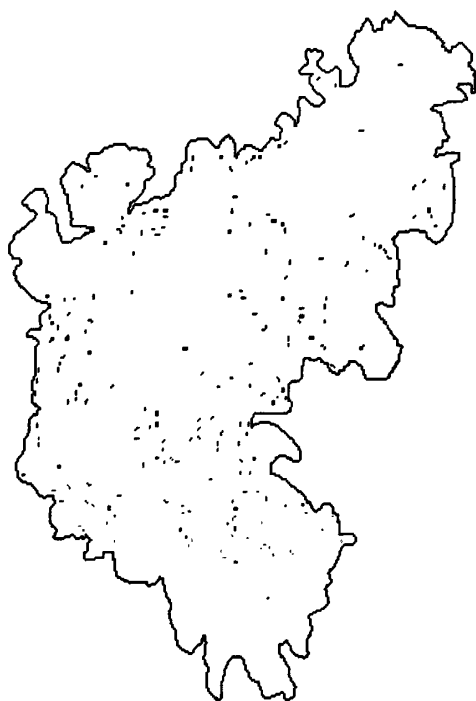


图 3-13 洼地

Fig 3-13 the local depression

经过生成的沟谷与山谷点比较（见图 3-14），研究发现：通过对洼地进行填充处理后，使其高程达到周围的最低网格单元的高程，从而以平坦区域进行流向计算。这样的填充处理，使洼地成为了山谷点的后选点，致使山谷点的数量与实际的山谷点的数量大，对下一步使用山谷点数据造成明显的错误。因此，要对山谷点图层进行筛选，通过洼地匹配来清除那些是洼地的山谷点。



图 3-14 沟谷与山谷点

Fig 3-14 the watershed vally and the channel points

经过沟谷与洼地比较（见图 3-15），研究发现：利用高精度 DEM 生成沟谷，对洼地进行填充处理，然后以平坦区域的形式确定 DEM 栅格单元流向，从而生成沟谷。这种方法可以生成连续的沟谷，但是由于洼地存在，致使水流方法存在明显的误差，沟谷发生扭曲。这种扭曲在高精度 DEM 比较明显，尤其是在由 1:5000 国家基本比例尺地形图构建的 DEM。填充后的洼地也对流域沟谷提取造成比较大的影响。因此，要对生成的沟谷进行洼地匹配来。经过处理后，效果如图

3-16, 通过洼地匹配方法可以拉直洼地造成的扭曲沟谷段。

同时由于洼地对山谷点的影响, 通过特征点匹配进行处理, 可以使沟谷与等高线的弯曲形状的拟合度更高, 与实际地形基本上相符合。

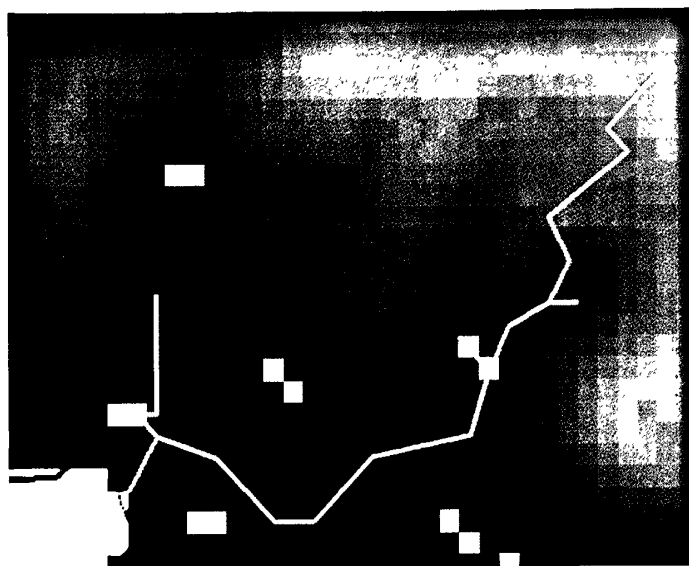


图 3-15 沟谷与洼地

Fig 3-15 the watershed vally and the local depression



图 3-16 沟谷与洼地

Fig 3-16 the watershed vally and the local depression

3.3.4 性能分析

ArcGIS 水文分析过程是：首先加载 DEM 数据，然后对 DEM 进行完全填充处理；对填充后的 DEM 进行流向分析，生成流向栅格图层；根据流向计算流量，生成流量栅格图层；设置流域阈值，提取流域的沟谷网络，再把沟谷网络转成矢量线图层；根据流向生成流域出水口栅格图层，再根据流向和流域出水口栅格生成流域沟谷的集水区域，再把集水区域转成矢量面图层。

本文所研究的小流域自动提取算法也是采用同样的流程，所不同的有三个方面：（1）在 DEM 填充之前，先提取洼地，进行洼地计算，人工判断采取完全填充或者自定义填充模式；（2）手工填入不同的流域阈值，生成不同的流域沟谷矢量线图层和流域集水区域矢量面图层，提取属性表，然后人工分析、判断、选择合适的阈值；（3）在选择合适阈值以后，生成合适的流域沟谷和流域集水区域，然后对沟谷进行洼地匹配和特征点匹配，使沟谷与实际的拟合度更高。

以 1:5000 岳麓山 DEM 为例，进行小流域提取以后获得的沟谷网络，与手工提取流域相比，通过算法实现流域提取操作更加简单方便，小流域提取以后获得的沟谷网络数据的精度达到了同样的程度，而且也对存在一些问题（如洼地填充后的影响，沟谷线的扭曲）进行了处理，使流域沟谷与实际的拟合程度更高（见图 3-17，其中青线为未修正的沟谷线，蓝线为修正后的沟谷线）。

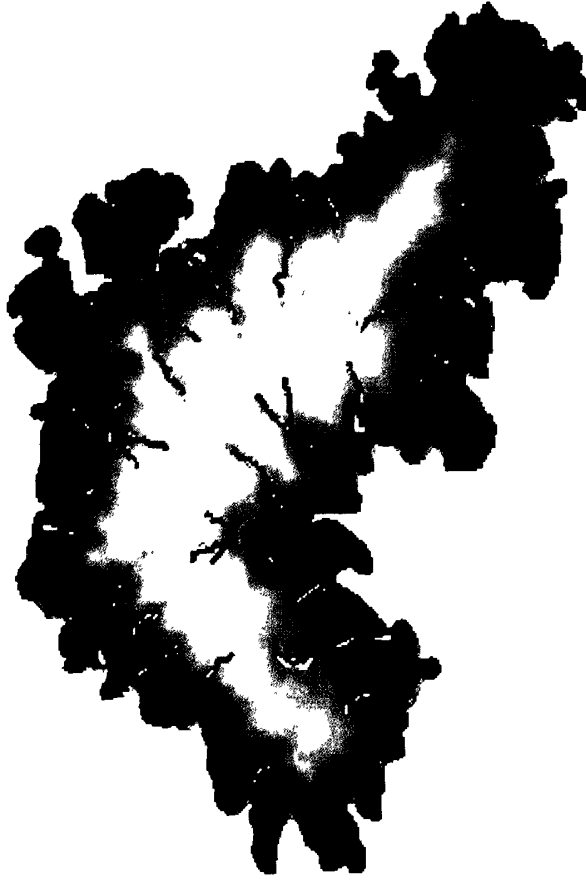


图 3-17 沟谷

Fig 3-17 the local depression

第四章 总结与展望

4.1 总结

本文在参阅国内外有关 DEM 提取流域特征信息的各种模型和算法的基础上,研究基于地形表面几何形态分析原理和基于地形表面流水模拟分析原理以及存在的问题等,采取了两者相结合的混合方法,基于 ArcEngine 实现小流域的自动提取。

本文以岳麓山作为例对小流域特征信息进行了提取,主要技术环节如下:

(1) 洼地填充。本文假设所提取到的洼地都不是真实的地表形态,对所有洼地都进行填平处理。填充洼地时采用“高程增量迭加法”,通过增高洼地栅格单元的高程值,使得 DEM 中每一个栅格的高程值不比周围八个栅格中的高程值中至少一个低,然后就可以计算水流方向,进行小流域提取。

(2) 流域阈值的选取。先计算所有栅格流量的平均值,然后分别在该平均值两侧各选两点为阈值,生成沟谷。通过对阈值与沟谷总长度进行线性函数、对数函数、多项式函数、乘幂函数进行分析,从而计算出合理的阈值。

(3) 伪沟谷的处理。首先把上一步骤中生成的采取 Strahler 三级分级方案的沟谷中的第二级别和第三级别的沟谷标记为主沟谷;然后对标记的主沟谷进行缓冲区分析;其次对缓冲区范围内的栅格单元降

低一定高程值,缓冲区范围外的栅格值赋为零,然后生成新的栅格图层;再通过生成的新的栅格图层与原始 DEM 进行叠加分析,对叠加后生成的 DEM 重新进行小流域提取,小流域提取后的沟谷数据就是进行伪沟谷删除的结果。通过缓冲区分析与叠置分析,对沟谷进行二次提取,不仅比较容易对伪沟谷进行处理,使沟谷结构比较合理,而且可以完全消除伪沟谷。

(4) 洼地匹配与特征点匹配处理。读取提取的洼地图层中的洼地单元,然后对山谷点图层进行筛选,删除山谷点图层中与洼地点坐标相同的山谷点;然后提取沟谷线图层上的每一条沟谷的转折点,删除被该洼地单元包含的转折点(不能是沟谷线的起点或终点)。

本文通过流域自动提取算法实验所提取出的流域沟谷网络与手工提取的进行比对,其精度达到同样的水平,并且进一步进行流域的修正处理,使流域沟谷与实际更相符合。本文的主要研究结论有:

(1) 选取五组数据对阈值与沟谷总长度进行相关性分析,发现两者的相关性与乘幂函数 $Y=404115 \cdot X^{-0.557}$ 的拟合度最好, R^2 的值达到 0.9995。当阈值设置为 145.54 时,生成的沟谷与实际沟谷的长度比较接近,沟谷密度较近似,阈值较合理。通过进行相关性分析,可以快速的选择恰当阈值,使生成的沟谷与实际沟谷更加相符,避免了反复实验阈值的繁琐过程,也保证了模拟结果的准确性。

(2) 通过对岳麓山提取到的小流域沟谷网络分析,发现填充处理后的洼地仍然对山谷点和小流域的沟谷线产生影响,通过匹配,删除洼地单元上的山谷点和沟谷的转折点,然后结合山谷点对沟谷线进行处理,使沟谷线与等高线的弯曲处更相符合,使沟谷更加合理,更符合实际。

(3) 比黄土高原 1:10000 比例尺地形图相比, 岳麓山 1:5000 比例尺地形图的精度更高, 这样通过 1:5000 比例尺地形图更高精度 DEM 数据来进行小流域的提取, 所提取到的小流域沟谷网络更加精确。但是, 在高精度 DEM 上生成的沟谷的弯曲程度就显得比较重要, 要进一步处理。通过洼地匹配和特征点匹配处理, 使小流域沟谷网络更加合理。

本文以岳麓山作为例对小流域特征信息进行了提取, 选取范围比较小, 还要更进一步多选取一些典型区域进行分析, 使算法更加成熟, 能够广泛应用于丘陵区域小流域的提取。

4.2 展望

目前针对高精度 DEM 的研究通常是基于 1:10000 或 1:50000 比例尺地形图, 而基于 1:5000 比例尺地形图的研究比较少。而且岳麓山在江南丘陵区具有很强的代表性。通过 1:5000 比例尺地形图生成更高精度 DEM, 进而进行岳麓山小流域的自动提取研究, 对进一步分析小流域沟谷等地貌特征和坡面、沟谷演化规律, 也能为广大丘陵地区的资源开发、保护和灾害防治提供基础数据。

本文应用岳麓山 DEM 的数据研究基于 DEM 的流域自动提取技术, 但由于学术水平、时间以及自身等因素的限制, 本研究还存在许多不足之处:

(1) 以岳麓山作为例对研究选取范围比较小, 还要更进一步多选取一些典型区域进行分析, 使算法更加成熟, 能够广泛应用于丘陵区域小流域的提取

(2) DEM 分析仅仅考虑了部分因素的影响而忽视其他相关因素的影响。如只考虑了水平分辨率的因素，而忽略了垂直分辨率因素的影响；只考虑了地形的因素，而忽略了人文因素的影响；在水流方向的确定中，采用了单流向方法，忽略了其他方向上的流向的影响。

(3) DEM 小流域提取特算法在水流方向的确定中，采用了单流向方法，认为水流方向是唯一的，这是对实际水流方向理想化，忽略了其他方向上的流向的影响。

所以，今后，要增加对国内外文献阅读，进一步完善和深化流域水系提取分析理论；多选取一些典型区域进行分析，加强模块对 DEM 数据通用性和适用性的研究，深入研究其流域特征点和特征线之间存在的关系，使算法更简洁高效。

参考文献

- [1] 王雷, 朱杰勇, 周燕. 基于 1:25 万 DEM 昆明地区地貌形态特征分析. 昆明理工大学学报(理工版), 2007, 32(1):6-9
- [2] 刘新华, 杨勤科, 汤国安. 中国地形起伏度的提取及在水土流失定量评价中的应用[J]. 水土保持通报, 2001, 21(1):57-62
- [3] 莫申国, 基于 DEM 的秦岭数字地貌格局研究, 华东师范大学学报(自然科学版), 2008(2):8-14
- [4] 赵健, 贾忠华, 罗统. ArcGIS 环境下基于 DEM 的流域特征提取[J]. 水资源与水工程学报, 2006(1): 75-76
- [5] 朱红春, 陈楠, 刘海英, 汤国安, 自 1:10000 比例尺 DEM 提取地形起伏度——以陕北黄土高原的实验为例, 测绘科学, 2005, 8(4):86-88
- [6] 朱红春, 陈楠, 刘海英, 汤国安, 基于 DEM 的流域地形因子提取与量化关系研究——以陕北黄土高原的实验为例, 测绘科学, 2007, 32(2):138-140
- [7] 陈浩, 方海燕, 蔡强国等. 黄土丘陵沟壑区沟谷侵蚀演化的坡向差异-晋西王家沟小流域为例[J]. 资源环境, 2006, 28(5): 177-184
- [8] 余生晨、刘大有、刘洪, 山脊线和山谷线的计算机自动检测, 中国图像图形学报, 1999.4(8):693-695
- [9] 周毅, 汤国安, 王春, 肖晨超, 董有福, 孙京禄, 基于高分辨率 DEM 的黄土地貌正负地形自动分割技术研究, 地理科学, 2010, 3(20):261-265
- [10] O' CALLAGHAN J F, MARK D M. The Extraction of Drainage Networks from Digital Elevation Data[J]. Computer Vision, Graphics, and Image Processing, 1984, 28(4): 323-344
- [11] JENSON S K, DOMINGUE J O. Extraction Topographic Structure from Digital Elevation Data for Geographic Information System Analysis[J]. Photogrammetric Engineering and Remote Sensing, 1988, 54 (11): 1593-1600
- [12] TARBOTON D G, BRAS R L, RODRIQUEZ I I. On the Extraction of Channel Networks

- from Digital Elevation Data[J]. Hydrologic Processes, 1991, 5(1):81-100
- [13] Tribe. Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method [J]. Journal of Hydrology, 1992, 139: 263-293.
- [14] Martz W, Garbrecht J. Short Communication, Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method—Comment [J]. Journal of Hydrology, 1995, 167: 393-396.
- [15] Stanley S W, Pierre C. U.S. soil erosion rates—myth and reality[J]. Science, 2000, 289: 248-250
- [16] 赵亚萍, 黄岩, 邱道持, 数字流域河网提取中的阈值问题研究, 信阳师范学院学报(自然科学版), 2008, 21(2):232-235
- [17] 李辉;陈晓玲;张利华;李长安, 基于三方向搜索的 DEM 中洼地处理方法, 水科学进展, 2009, 20(4):473-478
- [18] 周贵云, 刘瑜, 郭伦, 基于数字高程模型的水系提取算法, 地理学与国土研究, 2000, 16(4):77-81
- [19] 孔凡哲, 芮孝芳, 处理 DEM 中闭合洼地和平坦区域的一种新方法, 水科学进展, 2003, 14(3):290-294
- [20] 孔凡哲, 芮孝芳, 数字高程模型在流域水文模型应用中的若干问题, 水文, 2002, 22(5):1-4
- [21] 刘光, 李树德, 张亮, 基于 DEM 的沟谷系统提取算法综述, 地理与地理信息科学, 2003, 19(5):11-15
- [22] 虞继进, 数字高程模型(DEM)的建立与应用, 江苏测绘, 1999, 22(3):33-36
- [23] 陈恒, 叶斌, 大比例尺数字高程模型制作的研究, 首届长三角科技论坛数字区域建设与地理空间技术论坛优秀论文集, 2004, 230-238
- [24] 张会平;杨农, 张岳桥等. 基于 DEM 的岷山构造带构造地貌初步研究[J]. 国土资源遥感, 2004, 12(4):54-57
- [25] 黄陪之, 提取山脊线和山谷线的一种新方法, 武汉大学学报. 信息科学版, 2001,

26(3):247-252

[26]陈永良,刘大有,虞强源,从 DEM 中自动提取自然水系,中国图象图形学报,2002,7(1):91-96

[27]朱庆,赵杰,基于规则格网 DEM 的地形特征提取算法,测绘学报,2004,33(1):77-82

[28]靳海亮,高井祥,康建荣,基于矢量等高线数据提取山脊线山谷线的研究,测绘通报,2005,10:54-56

[29]陈永良,刘大有,一种新的山脊线和山谷线的自动提取方法,中国图形图象学报,2001,6(12):1230-1234

[30]任立良,数字高程模型在流域水系拓扑结构计算中的应用,水科学进展,1999,10(2):129-134

[31]任立良,刘新仁,数字高程模型信息提取与数字水文模型研究进展,水科学进展,2000,11(4):463-469

[32]闫国年,钱亚东,陈钟明.黄土丘陵沟壑区沟谷网络自动制图技术研究[J].测绘学报,1998,27(2):131-137

[33]LI Zhi-ling, ZHU Qing. Digital Elevation Model [M]. Wuhan: Wuhan Technical University of Surveying & Mapping Press, 2000. (in Chinese)

[34]李丽,郝振纯,基于 DEM 的流域特征提取综述,地球科学进展,2003,18(2):251-256

[35]周启明、刘学军,数字地形分析,北京:科学出版社,2006

[36]李志林、朱庆,数字高程模型,武汉大学出版社,2001

[37]陆中臣,贾绍凤,袁宝印等.流域地貌系统[M].大连:大连出版社,1991.42-43

[38]张宏、温永宁、刘爱利等,地理信息系统算法基础,北京:科学出版社,2006

[39]汤国安、杨昕, ArcGIS 地理信息系统空间分析实验教程,北京:科学出版社,2006

[40]汤国安、陈正江、赵牡丹、刘万青、刘咏梅, ArcView 地理信息系统空间分析方法,北京:科学出版社,2006

[41]高丽芳,基于 DEM 的岳麓山地貌特征的分析,长沙,湖南师范大学,2007

[42]刘爱利.基于 1:100 万 DEM 的我国地形地貌特征研究[D].陕西,西北大学,2004,6

[43]孙友波,基于 DEM 的水文特征信息的提取和可视化的研究与实现,北京,首都师范大学,2005,4

- [44] 魏南. 中国 1:100 万数字地貌(黑河幅)地貌形态特征提取方法研究[D]. 东北师范大学, 2006, 5
- [45] 刘华, 数字流域水文特征信息提取软件研究及应用, 江苏, 河海大学, 2007
- [46] 徐涛, 基于 DEM 的地形特征信息提取技术的研究, 武汉, 中国地质大学, 2005
- [47] 聂振钢, 基于 DEM 的流域和水系提取系统研究, 北京, 北京林业大学, 2008
- [48] 陈于林, 基于 DEM 的水系提取及水系网多级分解, 四川, 西南交通大学, 2006
- [49] 祁晨, RS_GIS 技术支持下的近 35a 艾丁湖演变研究, 新疆, 新疆师范大学, 2008
- [50] 杨万春, 基于小波多尺度分解的数字高程模型研究, 贵州, 贵州大学, 2007
- [51] 明安发, 基于小波地形简化技术的三维地形可视化的研究与实现, 贵州, 贵州大学, 2006
- [52] 徐涛, 基于 DEM 的地形特征信息提取技术的研究, 武汉, 中国地质大学硕士论文, 2005
- [53] 苏丹丹, DEM 中线性信息提取方法综合研究, 吉林, 吉林大学, 2006
- [54] 李建柱, 滦河流域分布式降雨径流模拟研究, 天津, 天津大学, 2005
- [55] 刘永强, 新疆融雪洪水预警 DSS 关键技术及实现方式研究, 新疆, 新疆大学, 2007
- [56] 戴红毅, 地理信息系统三维建模及可视化技术研究, 湖南, 中南大学, 2004
- [57] 董雱, 数字高程模型在陷落柱林三维空间形态表达和剥蚀回溯中的应用, 太原理工大学 2003
- [58] 郭超颖, 基于 DEM 的马莲河流域数字地形分析, 北京, 北京林业大学, 2008,
- [59] 兰烟, 亦农. 湖南省风景名胜区岳麓山揽胜[M]. 湖南: 湖南地图出版社, 1992
- [60] 国家标准. 数字测绘产品质量要求(第 1 部分)[M]. GB/T17941.1-2000.
- [61] 湖南省国土委员会. 湖南国土资源【M】. 湖南: 湖南省国土委员会, 1985

附录

1 程序代码

//1、DEM数据预处理

```
//读取DEM数据,
IMap pMap = axMapControl1.Map;
IRasterLayer rasterLayer = (IRasterLayer)pMap.get_Layer(common.GetIndex(pMap,
"demcaijian"));
string filePath =
rasterLayer.FilePath.Remove(rasterLayer.FilePath.LastIndexOf('\\'));
IWorkspaceFactory workspaceFactory = new RasterWorkspaceFactoryClass();
IRasterWorkspace rasterWorkspace = workspaceFactory.OpenFromFile(filePath, 0)
as IRasterWorkspace;
IRasterDataset rasterDataset =
rasterWorkspace.OpenRasterDataset(rasterLayer.Name);
//生成流向图层
string flowdirpath = baserootpath + "flowdir0";
IRasterDataset pFlowDir;
if (!FlowDir(rasterDataset, flowdirpath, out pFlowDir))
{
    return;
}
//提取洼地
string sinkpath = baserootpath + "sink";
IRasterDataset pSink;
if (!Sink(pFlowDir, sinkpath, out pSink))
{
    return;
}
//判断洼地图层是否为空, 非空时对洼地填充
double[,] rasters = RasterValue.ReadRasterZvalue(pSink, ref pUPLeftPoint,
ref pBlockX, ref pBlockY, ref pColumns, ref pRows);
if (parseNotZero(rasters))
{
    IRasterDataset rasterSingleDataset = null;
    string single = baserootpath + "single";
```

```

        string singleexpression = rasterLayer.FilePath;
        if (!SingleOutputMapAlgebra(rasterLayer.FilePath, single, out
rasterSingleDataset))
        {
            return;
        }
        IRasterDataset rasterFillDataset = null;
        string filldem = baserootpath + "filldem";
        if (!Fill(rasterSingleDataset, filldem, out rasterFillDataset))
        {
            return;
        }
        IRasterLayer pFillRasterLayer = new RasterLayerClass();
        pFillRasterLayer.CreateFromDataset(rasterFillDataset);
        pMap.AddLayer(pFillRasterLayer);
    }
    //2、特征点提取
    //分别选择三种矩形窗口(3×3、4×4、5×5) 计算山脊点和山谷点,
    //山脊点赋属性值为1, 山谷点赋属性值为-1
    double[,] rasters = RasterValue.ReadRasterZvalue(rasterDataset, ref
pUPLeftPoint,
                ref pBlockX, ref pBlockY, ref pColumns, ref pRows);
    double[,] featurepoint1 = GetRasterFeaturePoint(rasters, 1);
    double[,] featurepoint2 = GetRasterFeaturePoint(rasters, 2);
    double[,] featurepoint3 = GetRasterFeaturePoint(rasters, 3);
    double[,] featurepoint = GetRasterFeaturePoint(featurepoint1, featurepoint2,
featurepoint3);
    //根据属性值分别保存山脊点图层和山谷点图层
    IFeatureClass ridgepointFeatureClass = CreateNewPointShapeFile("ridgepoint");
    if (!ArrayToPoint(featurepoint, ridgepointFeatureClass, 1))
    {
        return;
    }
    IFeatureClass channelpointFeatureClass =
CreateNewPointShapeFile("channelpoint");
    if (!ArrayToPoint(featurepoint, channelpointFeatureClass, -1))
    {
        return;
    }
    //2、流域提取
    IRasterLayer rasterLayer = (IRasterLayer)pMap.get_Layer(common.GetIndex(pMap,
"filldem"));
    string filePath =
rasterLayer.FilePath.Remove(rasterLayer.FilePath.LastIndexOf('\\'));
    IWorkspaceFactory workspaceFactory = new RasterWorkspaceFactoryClass();

```

```

IRasterWorkspace rasterWorkspace = workspaceFactory.OpenFromFile(filePath, 0)
as IRasterWorkspace;
IRasterDataset rasterDataset =
rasterWorkspace.OpenRasterDataset(rasterLayer.Name);
//流向计算
string flowdirpath = baserootpath + "flowdir";
IRasterDataset pFlowDir;
if (!FlowDir(rasterDataset, flowdirpath, out pFlowDir))
{
    return;
}
//流量计算
string flowacclpath = baserootpath + "flowaccl";
IRasterDataset pFlowAccl;
if (!FlowAccl(pFlowDir, flowacclpath, out pFlowAccl))
{
    return;
}
//河网分级, 并保存矢量图层
string strahler = baserootpath + "strahler";
IRasterDataset pStreamStrahler;
if (!StreamOrder(pStreamnet, pFlowDir, strahler, out pStreamStrahler))
{
    return;
}
string strahlerpolyline = baserootpath + "strahlerline";
IFeatureClass pStrahlerPolyline;
if (!StreamToFeature(pStreamnet, pFlowDir, strahlerpolyline, out
pStrahlerPolyline))
{
    return;
}
//生成流域, 并保存矢量图层
string streamlinkpath = baserootpath + "streamlink";
IRasterDataset pStreamlink;
if (!StreamLink(pStreamnet, pFlowDir, streamlinkpath, out pStreamlink))
{
    return;
}
string streamwatershedpath = baserootpath + "watershed";
IRasterDataset pStreamwatershed;
if (!Watershed(pFlowDir, pStreamlink, streamwatershedpath, out
pStreamwatershed))
{

```

```

        return;
    }

    string watershed = baserootpath + "watershed2";
    IFeatureClass watershedFeatureClass;
    if (!RasterToPolygon(pStreamwatershed, watershed, out watershedFeatureClass))
    {
        return;
    }

    //4、沟谷线与山谷点匹配
    IRasterLayer rasterLayer = (IRasterLayer)pMap.get_Layer(common.GetIndex(pMap,
    "filldem"));
    string filePath =
    rasterLayer.FilePath.Remove(rasterLayer.FilePath.LastIndexOf('\\'));
    IWorkspaceFactory workspaceFactory = new RasterWorkspaceFactoryClass();
    IRasterWorkspace rasterWorkspace = workspaceFactory.OpenFromFile(filePath, 0)
    as IRasterWorkspace;
    IRasterDataset rasterDataset =
    rasterWorkspace.OpenRasterDataset(rasterLayer.Name);

    IFeatureLayer pStrahlerFeatureLayer =
    (IFeatureLayer)pMap.get_Layer(common.GetIndex(pMap, "strahlerline"));
    IFeatureClass pStrahlerFeatureClass = pStrahlerFeatureLayer.FeatureClass;
    ModifyPolyline(pStrahlerFeatureClass);

    private void RepairPolyline(IFeatureClass pFeatureClass, IFeatureClass
    pTinFeatureClass, IFeatureClass pPointFeatureClass)
    {
        Sort(pTinFeatureClass, "contour");
        ISpatialFilter pSpatialFilter = new SpatialFilterClass();
        pSpatialFilter.GeometryField = pFeatureClass.ShapeFieldName;

        IFeatureCursor pCursor = pFeatureClass.Search(null, false);
        IFeature pFeature = pCursor.NextFeature();
        while (pFeature != null)
        {
            IPolyline pPolyline = (IPolyline)pFeature.ShapeCopy;
            IPolyline pEndPolyline = (IPolyline)pFeature.Shape;
            pSpatialFilter.SpatialRel = esriSpatialRelEnum.esriSpatialRelCrosses;
            pSpatialFilter.Geometry = pPolyline;
            IFeatureCursor sCursor = pTinFeatureClass.Search(pSpatialFilter,
            false);
            IPolyline[] contourArray = CursorsToPolylineArray(sCursor);
            System.Runtime.InteropServices.Marshal.ReleaseComObject(sCursor);

```



```

        if (contourArray == null) continue;
        for (int j = 0; j < contourArray.Length; j++)
        {
            IPolyline pcountourPolyline = contourArray[j] as IPolyline;
            IPointCollection pPointCollection =
GetTheCommonPoint(pcountourPolyline, pPolyline);
            for (int k = 0; k < pPointCollection.PointCount; k++)
            {
                IPoint pJoinPoint = pPointCollection.get_Point(k);
                ITopologicalOperator topo = (ITopologicalOperator)pJoinPoint;
                IPolygon pTempPolygon = (IPolygon)topo.Buffer(11.14 / 2);

                pSpatialFilter.SpatialRel =
esriSpatialRelEnum.esriSpatialRelContains;
                pSpatialFilter.Geometry = pTempPolygon;
                IFeatureCursor sCursor2 =
pPointFeatureClass.Search(pSpatialFilter, false);
                IPoint[] pointArray2 = CursorToPointArray(sCursor2);

                System.Runtime.InteropServices.Marshal.ReleaseComObject(sCursor2);

                if (pointArray2 != null && pointArray2.Length > 0)
                {
                    IPoint pToPoint = pointArray2[0] as IPoint;
                    if (!parseInPolyline(pToPoint, pEndPolyline))
                    {
                        MovePolylinePoint(pEndPolyline, pToPoint);
                    }
                }
            }
        }
        pFeature.Shape = pEndPolyline;
        pFeature.Store();
        pFeature = pCursor.NextFeature();
    }
    System.Runtime.InteropServices.Marshal.ReleaseComObject(pCursor);
}

private void RepairPolygon(IFeatureClass pFeatureClass, double number)
{
    ISpatialFilter pSpatialFilter = new SpatialFilterClass();
    pSpatialFilter.GeometryField = pFeatureClass.ShapeFieldName;
    pSpatialFilter.SpatialRel =
esriSpatialRelEnum.esriSpatialRelIntersects;

```

```

IFeatureCursor pCursor = pFeatureClass.Search(null, false);
IPolygon[] pArray = CursorsToPolygonArray(pCursor);
System.Runtime.InteropServices.Marshal.ReleaseComObject(pCursor);

for (int i = 0; i < pArray.Length; i++)
{
    IPolygon pPolygon = pArray[i] as IPolygon;
    if (pPolygon.Envelope.Width > number && pPolygon.Envelope.Height >
number)
    {
        continue;
    }
    else
    {
        pSpatialFilter.SpatialRel =
esriSpatialRelEnum.esriSpatialRelContains;
        pSpatialFilter.Geometry = pPolygon;
        IFeatureCursor sCursor = pFeatureClass.Search(pSpatialFilter,
false);
        IPolygon[] inArray = CursorsToPolygonArray(sCursor);
        for (int j = 0; j < inArray.Length; j++)
        {
            IPolygon pTempPolygon = inArray[j] as IPolygon;
            MovePolygonPoint(pPolygon, pTempPolygon);
        }
    }
}
}

```

2 数据表

1: 5000 比例尺 DEM 误差统计表

点号 N	原始高程值 A	原始高程值 B	差值的绝对值 C= A-B	差值的平方 D=C*C
1	73.2	72.5	0.7	0.49
2	85.8	85.31	0.49	0.2401
3	86.7	87.55	0.85	0.7225
4	62.6	60	2.6	6.76
5	165.8	165	0.8	0.64
6	55.2	56.05	0.85	0.7225
7	63	62.5	0.5	0.25
8	78.5	77.5	1	1
9	66.3	70	3.7	13.69
10	89.3	87.5	1.8	3.24
11	106.1	106.02	0.08	0.0064
12	101.2	101.35	0.15	0.0225
13	118.3	118.09	0.21	0.0441
14	197.8	197.99	0.19	0.0361
15	132.4	131.74	0.66	0.4356
16	288.5	286.87	1.63	2.6569
17	237.1	236.9	0.2	0.04
18	158	158.46	0.46	0.2116
19	178.1	178.03	0.07	0.0049
20	168.3	168.13	0.17	0.0289
21	59	60	1	1
22	113.7	112.81	0.89	0.7921
23	106.1	110	3.9	15.21
24	167.3	167.92	0.62	0.3844
25	160.7	161.46	0.76	0.5776
26	152	152.24	0.24	0.0576
27	241.8	242.81	1.01	1.0201
28	82.3	82.41	0.11	0.0121
合计			23.74	37.646

中误差=

$$\sqrt{\sum_{i=1}^N Di / N(N = 28) = 1.340256}$$

平均误差=

$$\sum_{i=1}^N Ci / N(N = 28) = 0.915714$$

致 谢

在这三年的学习、工作和生活中，韩广导师给予了我莫大的关怀和帮助，他严谨的治学态度和求实的科研作风以及勤勉的工作态度使我在学习期间受益匪浅，给我今后的工作和学习指引了方向。衷心感谢韩老师在我学习、工作和生活上给予的无微不至的关怀。

感谢王芳师姐和艾福利、庞西磊师兄，在学习和工作中给予了诸多有益的指导和教诲，在生活上得到了诸多关怀和帮助。

感谢孙广禄同学在写作过程中给予极好的建议与帮助。龙芳同学、贺致远同学、徐斌同学在论文编写过程中提供了热心的帮助和建议，特此致谢。

最后，感谢我的家人对我一直的理解和支持，这是我一直以来的精神动力和支柱。

刘冰

2011 年 4 月

湖南师范大学学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：刘冰 2011年6月2日

湖南师范大学学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，研究生在校攻读学位期间论文工作的知识产权单位属湖南师范大学。同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权湖南师范大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于

1、保密□，在-----年解密后适用本授权书。

2、不保密□。

(请在以上相应方框内打“√”)

作者签名：刘冰 日期：2011年6月2日

导师签名：解 日期：2011年6月2日

