

Development and Application of Automatic Extraction Software for Section Information in Water Conservancy Engineering

Xinyong Ma

Xinjiang Changji Fanghui Hydropower Design Co., Ltd., Changji, Xinjiang, 831100, China

Abstract

In river dredging work, cross-sectional data of hydraulic engineering plays a crucial role. The collection methods of cross-sectional data have significantly improved from traditional manual measurement to modern computer technology, but there are still many links in the process that require manual participation. Starting from this, this study aims to explore the automation of manual operation through automatic logical decomposition, while analyzing the principle of adaptive recognition and classification of graphic objects in the secondary development process of AutoCAD using C++ language. Practical application has proven that this software not only ensures the accuracy of data, but also significantly improves the automation and efficiency of data extraction, demonstrating broad application potential.

Keywords

water conservancy engineering; Section information; Automatic extraction software; Software Development

基于水利工程中断面信息自动提取软件开发与应用

马新勇

新疆昌吉方汇水电设计有限公司, 中国·新疆 昌吉 831100

摘要

在河道疏浚工作中, 水利工程的横断面数据扮演着至关重要的角色。从传统的人工测量到现代的电算技术, 断面数据的收集手段已显著进步, 但流程中依旧存在许多需要人工参与的环节。以此为出发点, 本研究旨在探讨将人工操作环节通过自动逻辑分解实现自动化, 同时解析在利用C++语言对AutoCAD进行二次开发过程中, 图元对象的自适应识别与分类原理。实际应用证明, 该软件不仅确保数据的准确性, 还大幅提升数据提取的自动化程度和效率, 展现出广阔的应用潜力。

关键词

水利工程; 断面信息; 自动提取软件; 软件开发

1 引言

在水利建设领域, 河道清淤等项目在施工前, 往往需要参照断面纵/横图纸来进行工作。高效且精确地获取这些图纸中的测量数据, 成为测绘工作中的一个关键环节, 其直接关系到工程进度与效率。本研究采用C++语言进行二次开发, 以实现断面信息的自动化和批量提取。最后, 将所开发的软件(方法)应用于《新疆某河流地形测量项目Ⅲ标段》, 经与已有的软件进行比较验证, 验证了其有效性, 为提高外业工作效率提供了有力保障。

2 全要素批量提取方法与流程

以AutoDM软件版本V3.0的运作机制为例, 我们可以

将其工作过程归纳为数据预处理和信息获取两个主要环节。

以数据预处理环节为例, 工作人员需要手工测量节点至起点的距离, 以此来确定每一条线段的精确里程。这一过程不仅耗时费力, 而且极易因测量工具的精度、操作人员的技能水平以及视觉判断的差异而产生误差。同样, 在信息获取环节, 手工输入断面线的起始点和终点坐标, 也是一项重复且枯燥的任务。这些操作几乎全靠人工手动完成, 不仅效率低下, 而且稍有不慎就可能造成数据录入错误, 为后续的数据分析埋下隐患。

在数据预处理工作圆满结束的基础之上, 我们可以借助AutoCAD软件, 通过其二次开发技术, 实现图形元素的自动识别、精确分类与高效组合。这一技术的运用, 能够极大地减轻人工操作的负担, 将原本需要手工完成的复杂步骤转化为自动化流程, 从而显著提高工作效率。

具体来说, 通过编写定制化的脚本或插件, AutoCAD

【作者简介】马新勇(1985-), 男, 中国甘肃民勤人, 本科, 副高级工程师, 从事水利设计研究。

能够自动识别图纸中的各类图形元素，如线条、节点、标注等，并根据预设的规则对它们进行分类和整理。随后，这些经过处理的图形数据可以被自动导出为标准格式的 TXT 文档，为后续的数据分析、建模或报告生成提供准确、可靠的原始资料^[1]。

在数据预处理工作之后，首要任务便是关闭那些与当前分析无关的图层，以减少干扰并优化处理资源。这一步骤能帮助我们聚焦于核心数据，避免冗余信息对结果的潜在影响。直接选取所有对分析至关重要的断面信息要素，这些要素是构建模型、分析趋势或验证假设的基础^[2]。随后，软件便开始在后台默默施展其强大的数据处理能力。它依据不同图形元素的独特特征，如形状、颜色、属性标签等，进行精细的识别工作，确保每一个元素都能被准确无误地归类。这一过程不仅考验着软件的算法精度，也体现了对图形学原理的深刻理解。

识别完成后，软件会进一步根据元素间的逻辑联系，如空间位置关系、属性相似性或是用户自定义的规则，对识别出的元素进行分类与组合。这一步骤是数据整合的关键，它能够零散的数据点编织成一张紧密相连的信息网，为后续的分析提供有力支撑。最终，经过一系列复杂的运算与处理后，软件会将精心整理的结果保存为 TXT 文件。这种格式的选择既便于数据的存储与传输，也兼容了多种后续分析工具的需求，为数据的广泛应用打下了坚实的基础。

值得注意的是，上述提到的识别、分类、组合等具体运算过程，并非简单的线性操作，而是基于元素间复杂而微妙的逻辑联系，通过一系列高度智能化的算法实现的。这些算法的细节与实现机制，将在后续的内容中得到详尽而深入的阐述。

3 智能化阈值分析

在智能优化过程中，关键性的改进点集中在“里程追加”和“信息抓取”两大功能模块。

3.1 里程追加部分

在 CAD 系统中，里程标注（表现为文本元素）一般需按照面向河流下游的方向进行排列，并且要沿着断面线的中轴线位置进行设置。

在实施该步骤时，对起始点至各节点的距离 L 以及其旋转的角度 α 进行精确的测量至关重要。节点间距的累积方法 AGO 可用于确定距离 L ；至于旋转角度，则需依照断面线的方向来设定。若假定终点的坐标 (x_1, y_1) 和 (x_2, y_2) 已知，借助公式 (1) 和公式 (2) 即可计算出该点相对于水平线的角度 β 以及必要的旋转度数 α 。

$$\beta = \arctan\left(\frac{|y_2 - y_1|}{|x_2 - x_1|}\right) \quad (1)$$

$$\alpha = \begin{cases} \beta & y_2 > y_1, x_2 > x_1 \\ 360 - \beta & y_2 < y_1, x_2 > x_1 \\ 180 - \beta & y_2 > y_1, x_2 < x_1 \\ 180 + \beta & y_2 < y_1, x_2 < x_1 \end{cases} \quad (2)$$

3.2 信息提取

在 N 条特定的断面轨迹中，每一轨迹都对应一个唯一的起始原点和里程数值，因此我们拥有 N 组原点与里程数据。考虑到高程点与轨迹之间为非单一对应，实际上有 M 个高程数据点（包括非有效点）。在数据的前处理环节，关键要素一般会被调整到轨迹两侧。本系统支持以下规则设定：

- ①位于断面轨迹 5 米内的高程点被认为是有效数据。
- ②断面轨迹 10 米内所对应的里程数值与相关高程点相对应^[3]。
- ③断面轨迹 1 米以内的圆形被视为相关圆；以圆心作为基准点，若高程点与圆心的距离在 0.1 米以内，则该点被认定为原点（目的是收集高程信息）。

超出上述范围的图形要素均视为无效。基于这些规则，本文进一步展示了识别、分类以及归纳的流程逻辑，具体如图 3 所示。

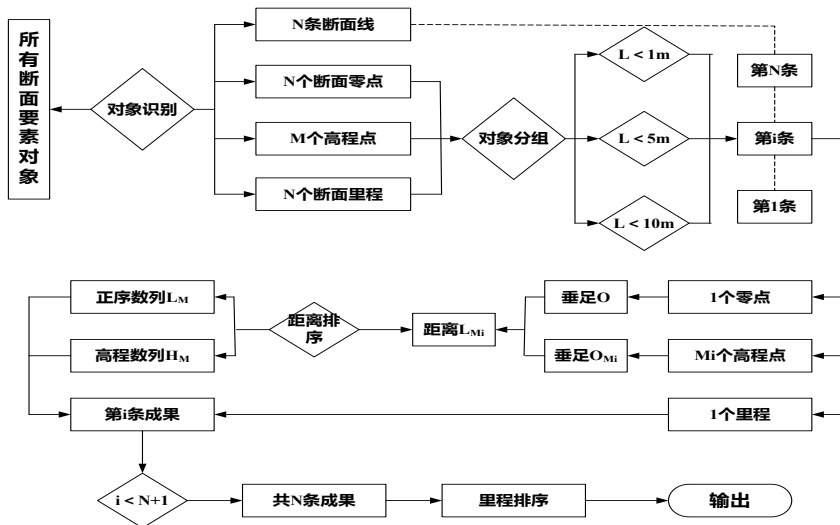


图 3: 运行逻辑图示

如图3所示，L 指标表示物体嵌入位置与断面轨迹的间隔距离； M_i 则指的是那些达到阈值标准的高程数据点；O 点代表起始点在断面轨迹上的垂足，而 O_{M_i} 则对应 M_i 个高程数据点在断面轨迹上的垂足位置； L_{M_i} 指的是 M_i 个高程数据点的垂足 O_{M_i} 与起始点垂足 O 的间距；LM 是将所有 L_{M_i} 值依照递减顺序排列形成的数组；HM 则表示与 LM 数组中各值对应的高程数据序列。

4 实际工程应用

该部分内容深度剖析软件的最佳效能，以及对用户交互界面和程序编码核心要素的详尽介绍。

4.1 交互界面设计

本软件采用 C++ 语言开发，构建一个图形窗口应用程序。该程序成功实现了第二部分所述的两大功能模块，并且针对断面信息提取过程中的技术细节进行了专项开发，以应对起始里程不为零或起始编号非标准的情况。另外，在执行《新疆某区域地形测绘项目 III 标段》时，对点属性信息的详细度要求极高。由于在 TXT 格式文件中手动输入这些详细信息极为费时，因此我们尝试将展点功能的相关代码独立出来，这要求我们必须精确掌握高程点与代码嵌入点之间的位置对应关系。

在 1:1000 这一特定比例尺下，高程点坐标与以文本形式展现的展点代码插入点坐标之间存在的细微偏差。这一发现深刻揭示不同比例尺之间存在的内在联系与转换规律，其中，S 作为 CAD 图像比例尺的代名词，扮演着至关重要的角色。在实际操作中，深知文本插入点的完全重合是一种理想状态，难以实现。但正是这种不完美，促使寻找到另一种解决方案——利用比例尺的独有属性。通过精准地把握比例尺的特性，能够确保在数据转换或信息提取过程中，属性信息的完整性得到最大程度的保留。这一方法不仅巧妙地规避因坐标偏差导致的信息丢失问题，更为在不同比例尺间自由切换、灵活应用数据提供坚实的保障。

4.2 实际应用展现

在某新疆区域的一条河流 III 标段中，该标段被细分为 A、B、C 三个区段，总长达到 160km，共布置了 130 个测

量断面。以 A 洪道作为参考，每两公里设立一个主要横断面，共计布置了 45 个这样的横断面。在这些横断面上，平均每个都分布着 120 个监测点，并且这 45 个横断面线的长度均值均超出两公里。

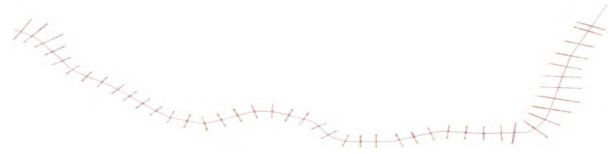


图 4: A 洪道断面整理数据图示

利用 AutoDM 软件版本 3.0 对断面数据进行常规效率的搜集以及添加地形标注(即断面点的属性信息)的过程中，大概耗费的时间为 3 到 8 个小时。而对于全部资料整理完毕，则通常需要 1 到 3 个工作日。通过直接选取所有断面元素(除中心线外)并按下 ENTER 键，使用计时器记录，运行时间约 100 秒，平均每个断面的处理时间仅为 2.3 秒，因此提取 160 个断面信息仅需 6 分钟。与原先“1 至 3 天”的时间相比，效率显著提高。

5 结语

综上所述，文章深入剖析水利领域断面信息自动化提取系统的构建原理及开发过程，并在篇末对该软件系统与市场上普遍应用的软件(如 AutoDM)在执行效率与稳定性两个关键指标上进行详尽的比较分析。该软件在确保稳定性的基础上，显著提高信息提取速度(对比现有软件处理单个断面的时间单位从分钟级降至秒级)。尽管本软件在智能化方面取得一定的成就，但依旧面临一些挑战，包括代码的优化简化以及复杂场景的应对能力等。未来计划针对这些问题进行优化升级，以促进软件的更广泛使用。

参考文献

- [1] 李磊,徐树元,罗时龙,等.水利工程中断面信息自动提取软件开发与应用[J].现代测绘,2023,46(6):58-61.
- [2] 王林森,王存堂,牛首印,等.基于VisualStudio下MES系统开发与应用[J].制造技术与机床,2024(8):97-104.
- [3] 吴国军.基于Ansys软件的大体积混凝土温度场仿真程序设计[J].湖南水利水电,2024(4):1-3.