

水文年鉴数字化研究

肖卫, 时昶

(长江水利委员会水文局, 湖北 武汉 430010)

摘要:水文年鉴是精装硬壳封面、线装书刊,在数字化处理时不宜将其拆散,再重新装订,为了加强对水文年鉴的保护和利用,针对不拆卷档案的特殊性,建立了一套特殊档案数字化加工系统。该系统具有纸质媒体整理、扫描、影像纠错、书目索引以及自动上载的功能,能兼容市面上主流书刊扫描仪及无边距扫描仪,涵盖了任务分配、图像扫描、图像处理、原文挂接和数据验证等整个加工业务流程,保证了纸质档案完整性和扫描数据与目录数据挂接一致性。应用该系统完成了长江上游、中游和下游干流等区域水文年鉴的数字化工作,为洪水预报、水文计算、水生态保护和防灾减灾等工作提供准确、高效的服务。

关键词:水文年鉴;数字化;扫描;影像

中图分类号:TP391 **文献标志码:**A **文章编号:**1674-3075(2011)04-0149-04

水文年鉴是对海量水文观测原始记录进行整理、分析,编制成系统数据和图表,按照统一的要求和规格,并按流域和水系统一编排卷册,逐年刊印的水文资料。水文年鉴为水利工程的建设和管理、水资源管理和开发利用、水环境保护和水生态资源建设等提供不可缺少的基础资料,发挥着重要的作用。蓄水工程、调水工程、流域或区域的防洪规划、水资源综合规划等都是基于水文年鉴中的原始数据,进行泥沙、降水等水文数据分析为基础,从而确定工程规模或规划方案(彭立波和鲁青,2009)。水文模型建立需要的资料一部分可以通过遥感等手段获取,更多的还是要依赖水文年鉴整编资料(彭立波和鲁青,2009)。张强等(2009)在分析水文年鉴等资料的基础上,对天津市暴雨在海河干流沿岸市郊区的径流系数、径流量进行了估算,并对海河干流排沥原则、排沥计算及海河水位、蓄量的预测方法做了说明,为海河干流行洪排沥、防洪减灾提供了参考依据。陈静生等(1998)根据水文年鉴中的水质监测资料和全球水质监测计划中长江站的水质监测资料,对长江干流1960-1980年代水质变化进行了研究。孙延锋等(2011)根据中苏历史水文年鉴,普查了冰坝河流的分布和凌汛洪水过程,分析了寒冷山区性河流冰坝形成的条件和类型。

20世纪80年代及以前的水文年鉴只有刊印本,20世纪90年代开始,水文年鉴的资料整理和储

存开始应用计算机技术,并进行了水文数据库建设的研究和实践。王海军和张潮(2007)研究了基于XML(extensible markup language)技术制作报表的方法,并应用于水文资料整汇编软件研制中。张然(2011)开发出基于VBA开发环境的程序,将水文整编成果表转换成符合数据库结构的标准格式,并录入到标准数据库中。

长江水利委员会水文年鉴利用率非常高,由于长时间的频繁借阅利用,对纸质水文年鉴的损害非常严重。一些20世纪50年代印刷的年鉴已经破损严重,有许多封面开裂、内页破损或脱落。水文年鉴是精装硬壳封面、线装书刊,在数字化处理时不能将其拆散,再重新装订。2010年,长江水利委员会水文局引进了特殊档案数字化加工系统,对水文年鉴进行数字化加工扫描并载入数据库,以推进水文年鉴电子信息化管理,实现水文年鉴资源依法共享,为长江流域综合治理和水生态环境保护提供重要基础资料支撑。

1 系统结构和功能

1.1 系统结构

特殊档案数字化加工系统硬件包括服务器端和客户端,并通过网络互连。服务器端完成任务分配等作业,客户端连接扫描仪进行影像扫描操作和处理,并将电子文档上载到档案数据库服务器。系统结构见图1。

1.2 系统功能

特殊档案数字化加工系统是根据水文年鉴的特殊要求而设计的,具有界面美观、功能键统一,操作

收稿日期:2011-07-12

作者简介:肖卫,女,1962年生,工程师,研究方向为水文信息化。E-mail:xiaowei1024@hotmail.com

简便等特点。系统能兼容市面上主流书刊扫描仪及无边距扫描仪,涵盖了任务分配、图像扫描、图像处理、原文挂接和数据验证等整个加工业务流程,保证了纸质档案完整性和扫描数据与目录数据挂接一致性。系统支持 TIFF、JPG、PDF 等多种文件格式。

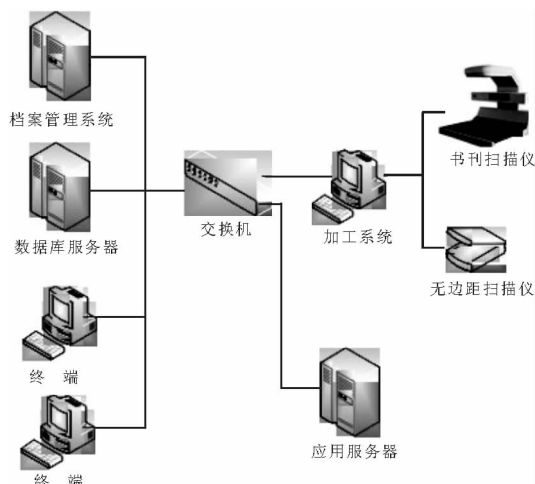


图1 特殊档案数字化加工系统结构

Fig.1 Overall structure of special file digitization and processing system

1.2.1 任务分配 可进行工程管理(新建工程、打开工程、保存与接受工程、标记与撤消工程等)和批次管理(新建或删除1个批次、批次划分、重命名批次、标记与撤消批次等)。

1.2.2 影像扫描 可以按照要求进行扫描影像、重扫影像(新的影像替换当前影像)、插扫单页或多页影像(新的单页或多页影像将插在当前影像之后)、导入已有影像、批量删除导入的影像等。

1.2.3 影像处理影像优化 自动对影像进行“纠偏”和“去黑边”操作、清除选定区域内或外部影像、按照中缝设定的宽度自动对影像进行“去中缝”操作。

影像校正:线型弯曲校正,文字变形校正。

影像的排列方式:按顺序或反序排列切分影像。

影像旋转:影像左旋90°、右旋90°、旋转180°或旋转一个任意角度。

放大与缩小影像:放大或缩小影像、局部放大影像、影像交换、影像移动。

编辑影像:擦除选定区域内或外影像,切割、剪切、反色、手动纠偏等操作,撤消或恢复操作等。

保存与删除影像:保存或删除预览窗口中所选中的影像。

影像纠错:影像修正(如对黑边、歪斜、污点的

处理)、缺页重扫、窗体的缩放、翻页码、错页调整(提升上页、降低下页)、移动一个或多个影像、交换一个或多个影像等。

自动影像遮幅:自动去除影像多余的黑边使其变为白色。

手指影像删除:自动删除手指影像。

影像对中:自动对影像进行对中校正。

影像文件合并:根据扫描目录信息,合并成多页TIF图像文件。

影像预览:对扫描影像进行浏览、查看,可选择单幅或双幅影像浏览方式。

批处理:对一批影像进行重扫、旋转、镜像翻转与自动纠偏等操作;对一批影像进行去黑边、歪斜、污点等操作;对一批影像进行由彩色转化为灰度或黑白影像的操作等。

1.2.4 原文挂接 手动或定时上载方式将原文挂接到档案管理系统对应的条目下。

1.2.5 书目著录 根据书刊目录进行识别挂接。

1.2.6 手动、定时和批量上载 可以手动或定时上载电子文档,还可以手动上载多个电子文档。

1.2.7 数据验证 通过档案查询系统逐页检查。

2 系统操作流程

特殊档案数字化加工系统可以按照不同的要求灵活定制操作流程。在操作过程中,可以根据纸质幅面大小设置扫描幅面,还可以根据纸质不同的色彩设置彩色或黑白扫描。这样,既能真实地记录纸质档案的原貌,还能有效地节省存贮空间,便于快速打开和阅览电子文档。系统操作流程见图2。

3 系统特点

3.1 适用于不可拆分、倒置困难的档案或书籍

系统采用书刊扫描仪,对于精装、硬壳封面的档案或书籍则避免了拆分,扫描时不需倒置档案或书籍,只需将页面朝上,平放于托书台。

3.2 可以扫描不规则的纸张

系统既可以加工A3、A4、B5等标准规格的纸质文件,还可以加工不规则的纸质文件和贴有标签的纸张,甚至是装订纸张。

3.3 高精度彩色、灰度和黑白扫描

通过高分辨率CCD(电荷耦合器件图像传感器),系统实现了高精度完美的扫描。可处理不同的扫描模式,包括彩色、灰度及黑白模式,都可保证最佳的扫描效果。

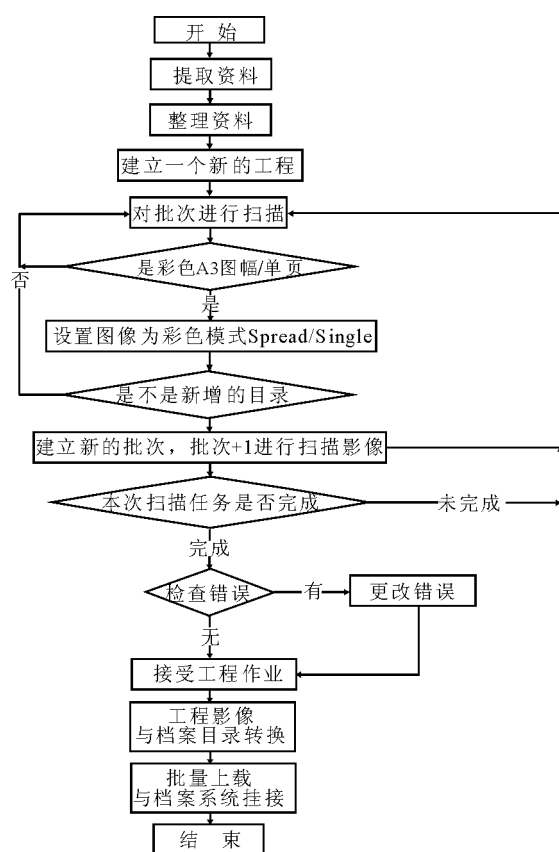


图2 特殊档案数字化加工系统操作流程

Fig.2 Flow chart of a special file digitization and processing system

3.4 保护珍贵的文件

系统具有自动识别和修正书本文字弯曲变形的功能,扫描时不需压平书本以防止文件影像因书本弯曲而扭曲变形。

3.5 自动修正

系统具有书脊弯曲校正、文字变形校正、自动中缝删除、自动遮幅、自动影像对中、自动手指印删除等,使扫描电子文档能真实再现原稿本色。

3.6 3D 扫描

系统具有3D影像扫描功能。因扫描时不直接接触实物,物体可以垂直放置,可轻松、灵活地扫描3D图像。

3.7 扫描电子文档与目录自动挂接

水文年鉴一般按照说明资料、考证资料、水位资料、流量资料、输沙率资料、泥沙颗粒级配资料、水温资料、冰凌资料、水化学资料、地下水资料、降水量资

料、蒸发量资料、水文调查资料和补刊资料等进行目录编排和刊印。根据水文年鉴的特殊性,为了方便电子文档的浏览和查阅,在进行水文年鉴的扫描过程中,按目录顺序建立批次,利用批次逐一将其进行合理划分,按批次划分顺序形成工程任务的电子文档(即影像文件)。将该工程任务的电子文档上载到档案管理系统后,与对应的档案条目自动关联,并将水文年鉴案卷目录与工程任务的电子文档的批次相关联,导入到档案管理系统,实现对应条目与电子影像文档一一对应。

3.8 中文界面,简单易用,操作轻松

独特的扫描软件是基于用户需求而开发的,使系统不仅能够设置扫描功能,也可以保存多种格式的影像,如BMP、JPG、PDF、TIFF、批量或多页文件,还可以进行影像后处理、纠偏、分页浏览和缩略图浏览等。

4 小结

特殊档案数字化加工系统适应了不拆卷档案的特殊性,既加强了对水文年鉴的保护和利用,又实现了水文年鉴由纸质资料向电子文档的转化。长江水利委员会水文局现已应用该系统完成了馆藏长江上游、中游和下游干流等区域水文年鉴的数字化工作,并为洪水预报、水文计算、防灾减灾、水生态保护和流域综合规划等工作提供了准确、高效的水文数字化服务,取得了很好的效果。

参考文献

- 陈静生,关文荣,夏星辉,等. 1998. 长江干流近三十年来水质变化探析[J]. 环境化学, 17(1):8-13.
- 彭立波,鲁青. 2009. 特殊的水利档案信息资源——水文年鉴[J]. 黑龙江水利科技, (6):135-136.
- 孙延锋,郑旭,李龙辉. 2011. 黑龙江松花江流域部分年份冰坝凌汛分析[J]. 黑龙江水利科技, (3):191-192.
- 王海军,张潮. 2007. XML技术在水文资料整汇编软件中的应用[J]. 长江科学院院报, 24(2):22-24.
- 张然. 2011. 基于VBA编程技术实现水文整编成果向基础水文数据库的转换与导入[J]. 江苏水利, (6):44-46, 48.
- 张强,王小琛,邢荣. 2009. 天津市暴雨对海河干流排涝及海河水位变化影响[J]. 海河水利, (2):48-49, 61.

(责任编辑 杨春艳)

Research on Digitization of Hydrology Yearbook

XIAO Wei, SHI Chang

(Bureau of Hydrology, Changjing Water Resources Commission, Wuhan 430010, China)

Abstract: The hydrology yearbook records basic hydrological data, which played an important role in production, scientific research and services. In order to improve the protection and utilization of hydrology yearbook, especially for those files that can't be taken apart, the bureau of hydrology developed a special file digitization and processing system, which can realize processing and scanning papery media file, image error correction, bibliography search and automatic upload. By transforming from papery yearbook into digital information for operation and management, it can not only effectively protect the hydrology yearbook and expand the range of utilization, but also improve the efficiency and realize share of information.

Key words: hydrology yearbook; digitization; scan; image