National Mapping Heritage Project

A Digital Library for Mapping Products

By Krishna Kumar R Nair, GIS Analyst, CGIS.

The Centre for GIS (CGIS) is the national base mapping and positioning service agency for the State of Qatar. It develops and maintains the digital topographic vector database, the digital ortho image database, and the digital terrain model database for the entire State of Qatar. It is also responsible for the Geodetic Datum and network for the State of Qatar.

The Centre has inherited a large repository of mapping products from its various national mapping activities over the past few decades, which are stored in the Centre's Mapping office. The mapping repository contains different types of maps, aerial photo products and related records. They include color, black & white maps, transparency maps, aerial photo negatives, aerial photo prints, mosaic negatives, map indexes, flight diagrams, and technical/administrative records. The National Mapping Heritage Project was initiated for digitally preserving its voluminous collection of mapping products.

The project objectives include digitally scanning all the mapping products, creating a spatially indexed digital library for the products with metadata in a relational database management system, and developing an electronic data discovery and data browsing tool to provide rapid access to this large volume of archived data including metadata.

Overview

Scanning technology allows storing, accessing and restoring hard copy information in digital environment. It will also dramatically reduce the on-site storage and handling as well as conserve maps, photos and documents from deterioration and information loss. Most importantly digital capture and storage of this information would provide on-line access to multi users.

Processing, compression & storage of scanned mapping materials

A comprehensive database of mapping products and all related information is being created in Oracle. The data discovery and data-browsing tool is being developed using Visual Basic and ESRI MapObjects/ArcView GIS. The tool will be capable of tracking what items are in the map library through attribute or spatial selection, browse any selected scanned image and produce hard copy outputs/information reports. This tool will work in list and graphical mode independently.

Data capture and post processing

A detailed study of the existing maps photos, images and other mapping products were carried out to understand the functional requirements of data automation and extent of related metadata. Before scanning, it is being ensured that the products are free of dusts and kinks.

VIDAR TruScan Select large format scanner with geometric accuracy +/- 0.1% is used for scanning maps and photo mosaic prints. The Scanner software VIDAR TruInfo does despeckling thresholding and deskewing functionalities.

For scanning the photogrammetric products viz aerial photo negatives, oblique negatives, markup prints etc. a high resolution, radiometrically and

Maps being scanned by large format scanner

geometrically precise Photogrammetric Scanner will be used.

The scanned image file is closely examined and the data is cleaned using image processing software Photoshop. The Image is then edited to add any attribute information and cropped to remove unwanted edges.

Since size of the scanned maps and photos will run into hundreds of Gigabytes, an effective compression technique need to be evolved. All images are compressed and retrieved using MrSID (Multi-resolution Seamless Image Database) compression software - a new technology for efficient storage and retrieval of digital images of very large dimension. MrSID Publisher utilizes a wavelet transform-based algorithm to achieve this.

Data storage relates to digital storage space requirements, data format, storage management software and hardware, and interface to indexing software. The data consists of two parts, the frequently accessed metadata and randomly accessed original high resolution scanned data. The high volume image data is stored using Iomega's portable 2GB Jaz drive and disks. The metadata database stored in the local Server will be fully on-line on Local Area Network (LAN), while the image data on Jaz drive will be identified by the indexing tool and will be loaded. All data backups are maintained in digital tapes also.

Product database and Spatial Indexing

There are two major classes of queries for spatially indexed information. First the users will be interested to discover the phenomena that occur at specific locations. i.e. "what is here?" phenomena. Secondly the users want to discover the locations at which the specified phenomena occur. i.e. "where is this?". Geographic locations can be specified both directly (coordinates) and symbolically (address). When dealing with a mix of such representations, a useful intermediate abstraction is the location's footprint i.e. the geographical extent on the map. Thus both the above direct and symbolic representation can be represented by point or polygon data models on the map. In this way, every product will inherit a locational footprint. After knowing which data exist and where, users may want to examine the actual map or air photo by viewing this selected data on the screen, or may browse the area for the set of data.

Towards building the metadata schema, detailed study was carried out for identifying the full set of related information of the mapping products. It is important to construct a coherent metadata schema from the current array of international standards. Related tables were created in Oracle for all information entities.

The metadata for the information system were selected from a broad classification of metadata topics i.e. identification information, data quality information, spatial data organization information, spatial reference information, entity and attribute information, distribution information, metadata reference information, citation information, time period information, contact information, product location (hardcopy and digital) and product identifiers.

Along with scanning, this full set of data is collected and populated in the Oracles tables using forms with built-in data integrity.

The main focus of the indexing software is data discovery. The indexing software is being created in Visual Basic with the spatial data handling components provided by MapObjects/ArcView GIS. The location foot prints of the mapping products depicting the geographical extent of the maps and photos will be created in Arc/Info. Individual footprints are linked using unique digital product identification key to individual rows in the Oracle table containing the metadata. The tool integrates elements of user interface such as indexing dimensions search capabilities and report formats.

Data maintenance and onward growth

As an archival digital library this system will mainly store unchanging data that will require little update. But there are cases when few modifications may happen to the metadata, which need to be modified. Also as time passes, more mapping products will gets added to this system, and the present design of the system is such that more types of data products can be accommodated in future. The system should be able to accept future technology upgradation also.

Conclusion

The National Mapping Heritage Project accommodates an incremental and evolutionary design. Completion of the project creates a very stable product for archival of all national mapping products, and an exhaustive information system where by all related information regarding any mapping product, can be queried spatially indexed. It can also locate the physical location of the hard copy products. The design of the database product hierarchy is such that it can be expanded to accommodate products of different types in future.