"A decision is as good as the information that goes into it."
John F. Bookout, Jr.

Chapter 4 Data Capture

How to feed maps into your computer

Data: the fuel

The geographic data are information about the earth's surface and the objects found on it. Data are fuel to a GIS. How can we feed data such as a map into a GIS? Data capture is the process of putting information into the system. A wide variety of sources can be used for creating geographic data.

Types and sources of geographic data

Geographic data are generally available in two forms: analogue data and digital data. Analogue data are a physical product displaying information visually on paper, e.g. maps. Digital data are information in a computer-readable form, e.g. satellite data (Figure 4.1).

There are various sources for obtaining these types of data. For example, as shown in Figure 4.2, the sources are maps, aerial photographs, satellite images, existing tabular data (in analogue and digital format) and field data (GPS). A GIS is able to capture these different types of data from various sources. Creating a database, i.e. capturing the data, is the initial and time-consuming stage of a GIS project.

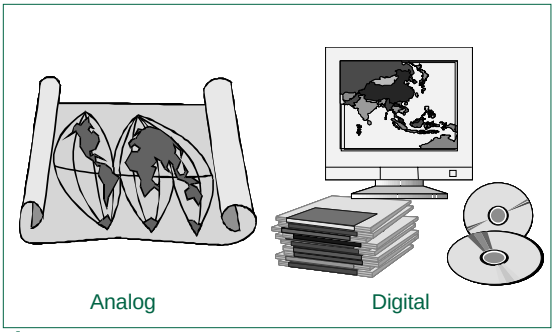

Figure 4.1
Analogue and digital data

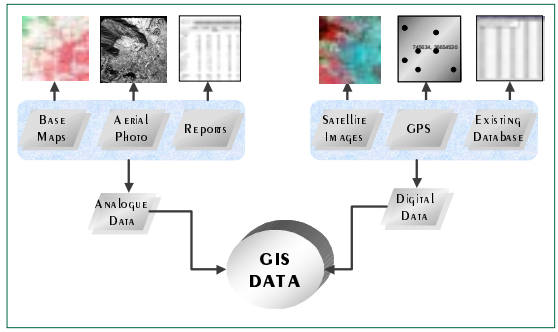

Figure 4.2
Data sources

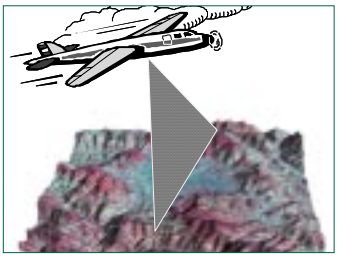

Figure 4.3a
Arial photography

Data capturing methods

Data capturing methods from various sources commonly used in a GIS are briefly discussed below (Figure 4.3).

Photogrammetric compilation

The primary source used in the process of photogrammetric compilation is aerial photography. Generally, the process involves using specialised equipment (a stereoplotter) to project overlapping aerial photos so that a viewer can see a three-dimensional picture of

the terrain. This is known as a photogrammetric model. The current technological trend in photogrammetry is toward a greater use of digital procedures for map compilation.

Figure 4.3b
Digitiser

Digitising

A digitising workstation with a digitising tablet and cursor is typically used to trace digitise. Both the tablet and cursor are connected to a computer that controls their functions. Most digitising tablets come in standard sizes that relate to engineering drawing sizes (A through E and larger). Digitising involves tracing with a precise cross hair in the digitising cursor features on a source map that is taped to the digitising tablet and instructing the computer to accept the location and type of the feature. The person performing the digitising may input separate features into map layers or attach an attribute to identify the feature.

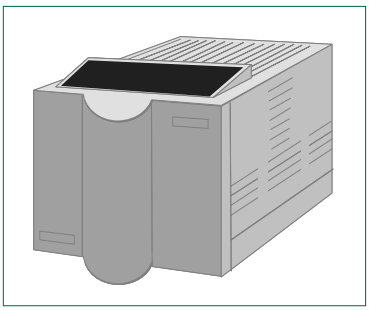

Figure 4.3c
Map scanner

Map scanning

Optical scanning systems automatically capture map features, text and symbols as individual cells or pixels and produce an automated product in raster format. Scanning outputs files in raster form, usually in one of several compressed formats to save storage space (e.g. TIFF 4, JPEG). Most scanning systems provide software to convert raster data to vector format that differentiates point, line and area features. Scanning systems and software are becoming more sophisticated with some ability to interpret symbols and text, and store this information in databases. Creating an intelligent GIS database from a scanned map will require vectorising the raster data and manual entry of attribute data from a scanned annotation.

Figure 4.3d
Satellite data

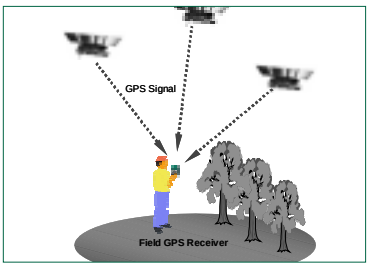

Figure 4.3e
GPS

Satellite data

Earth resources satellites have become a source of huge amounts of data for GIS applications. The data obtained from satellites are in digital form and can be imported directly into a GIS. There are numerous satellite data sources such as LANDSAT or SPOT. A new generation of high-resolution satellite data, that will increase opportunities and options for GIS database development, are becoming available from private sources and national governments. These satellite systems will provide panchromatic (black and white) or multi-spectral data in the 1-m to 3-m ranges as compared to the 10-m to 30-m range available from traditional remote sensing satellites.

Field data collection

Advances in hardware and software have greatly increased opportunities for capture of GIS data in the field (e.g. utility sign inventory, property surveys, land-use inventories). In particular, electronic survey systems and the global positioning systems

(GPS) have revolutionised surveying and field data collection. Electronic distance measurement services allow for survey data to be gathered quickly in an automated form for uploading to a GIS. Sophisticated GPS collection units provide a quick means of capturing the coordinates and attributes of features in the field.

Tabular data entry

Some of the tabular attribute data that are normally in a GIS database exist on maps as annotation or can be found in paper files. Information from these sources that is required for GIS applications has to be converted to a digital form through keyboard entry. This kind of data entry is commonplace and relatively easy to accomplish.

Document scanning

Smaller format scanners can also be used to create raster files of documents such as permit forms, service cards, site photographs, etc. These documents can be indexed in a relational database by number, type, date, engineering drawings, etc., and queried and displayed by users. GIS applications can be built that allow users to point to and retrieve for display a scanned document (e.g., tax parcel) interactively.

Translation of existing digital data

Existing automated data may be available from existing tabular files maintained by outside sources. Many programmes are available that perform this translation. In fact, there are many GIS packages with programmes that translate data to and from several 'standard' formats that are accepted widely by the mapping industry. They have been used as intermediate 'exchange' formats for moving data between platforms (e.g. Intergraph SIF, TIGER, Shapefile and AutoCAD DXF).