**Introduction**

Traditional maps often plot a geographical environment on a scale, while digital cartography took the job forward by creating maps using computers that boasted of many additional interactive functions. Though the use of compass and other advanced magnetic storage devices allowed people to create more accurate maps, store and manipulate them digitally in the early days of map making, it was only in the 20th century that advanced electronic technology brought a revolution in digital cartography and computer mapping tools. The ready availability of computer peripherals like monitors, printers, scanners, analytic stereo plotters, accessibility of computer programs for database management, correct image visualization and processing, as well as proper spatial analysis had made the entire process of map making much easier. Maps to address the needs of new industries were possible with such technology by placing spatially located variables on the existing maps to create newer ones with enhanced features and potentials. With time, digital cartography has taken strides to become better, thereby offering modern businesses a lot of benefits as follows:

* Help visitors find your branch office locations: You can easily embed interactive online maps in your website that would help your customers to know about store locations all over the world, and help them reach any particular store with ease.

* Easy marketing in trade shows: If you are participating in a trade show, you can use digital mapping tools to help your potential clients locate your booth in the trade show or to guide them to the venue via an interactive street image.

* Using interactive image maps for offline use: You can easily make your offline multimedia presentations on the intranet or CDROM more engaging and interesting by using interactive image maps with clickable points and information box showing on mouse rollover. Some map software even allow you to export interactive maps to JPG that you may use for your document or PowerPoint presentations.

* Make the maps more informative and interactive: With digital cartography at hand, you can add labels, logos and icons to specific points, show routes with curves or lines, and add text descriptions in tool tip or info boxes. All these elements make the maps more easy to follow and interactive in nature.

* Other user-friendly elements of digital cartography: Elements like map legends, map chart and zoom function in digital maps make them a preferred choice over traditional paper maps.

* Floor plans for real estate dealers: For those in the real estate business, interactive maps for floor plans of a building or mouse over boxes showing information related to homes for sale and their prices can serve a lot.

* Benefits for tourism industry: People looking to book their vacations often search for information like the best deals, the easiest route to reach a destination, any specials and discounts on offer, cheap air fares etc. Using digital cartography, you can show all these information via various interactive features, thereby making it easier for travellers to get the required data at a glance.

* Advantages for colleges and universities: Showing campus details and sharing information about various courses and their venues becomes easy with digital maps. That’s how educational institutions can benefit a lot from the use of digital cartography.

No wonder that digital maps serve as useful guides for conveying multi-faceted information through various interactive features, which in turn make them a preferred choice over their traditional counterparts.

1. **Data Acquisition**

Data sources for GIS are broadly classified as *primary* or *secondary*. Primary data are those collected through first-hand surveys and can be termed raw data if they are unprocessed observations. Secondary data are those collected by others, perhaps even for a different purpose, or have been derived from
published/marketed sources. All data used in connection with GIS that have dimensionality can be categorized by measurement type and have characteristics of scale and resolution. Furthermore, the data may be an exhaustive compilation (e.g., census) or it may be a sample.

**Modes of Data Acquisition/Input**

- Keyboard entry for non spatial attributes
- Manual locating devices
  - User directly manipulates a device whose location is recognised by the computer e.g. digitizing
- Automated Devices
  - Automatically extract spatial data from maps and photography e.g. scanning
- Conversion from other digital sources

(a) **Digitizers**

Digitizers are the most common device for extracting spatial information from maps and photographs

**Hardware:**

The position of an indicator as it is moved over the surface of the digitizing tablet is detected by the computer and interpreted as pairs of \((x, y)\) coordinates. The indicator may be a pen-like stylus or a cursor (a small flat plate with a cross hair). Frequently, there are control buttons on the cursor which permit control of the system without having to turn attention from the digitizing tablet to a computer terminal.

Digitizing tablets can be purchased in sizes from 25 X 25cm to 200 X 150cm, at approximate costs from $500 to $5000.
Early digitizers were backlit glass tables. These operated on the basis of a magnetic field generated by the cursor and tracked mechanically by an arm located behind the table. The arms motion was encoded, coordinates computed and sent to the host processor. Some early low cost systems had mechanically linked cursors with the free cursor digitizer initially more expensive. The first solid-state systems used a spark generated by the cursor and detected by linear microphones. This had of course, problems generated by ambient noise.

Contemporary tablets use a grid of wires embedded in the tablet to generate a magnetic field which is detected by the cursor. Accuracies derived from such are better than 0.1mm which is better than the accuracy with which the average operator can position the cursor. The functions for transporting coordinates are sometimes built into the tablet and used to process data before it is sent to the host.

The Digitizing Operation

The map is affixed to a digitizing table and three or more control points ("reference points", "tics", etc.) are digitized for each map sheet. These are easily identifiable points such as intersections of major streets, major peaks or points on a coastline. The coordinates of these points will be known in the coordinate system to be used in the final database, e.g. lat/long, State Plane Coordinates or military grid. These control points are used by the system to calculate the necessary mathematical transformation parameters to convert all coordinates to the final system. The more control points, the better the accuracy in computing the parameters and the better the quality of the final coordinates.

Digitizing the map contents can be done in two different modes:

- **Point mode:** The operator identifies the points to be captured explicitly by pressing a button on the cursor
- **Stream mode:** Points are captured at set time intervals (typically 10 per second) or on movement of the cursor by a fixed amount

Advantages and disadvantages:

- In point mode the operator selects points subjectively. As a result, two point mode operators will not code a line in the same way.
- Stream mode generates large numbers of points, many of which may be redundant.
- Stream mode is more demanding on the user while point mode requires some judgement about how to represent the line.

Most digitizing is currently done in point mode.

Problems with digitizing maps

These arise since most maps were not drafted for the purpose of digitizing. Paper maps are unstable: each time the map is removed from the digitizing table, the reference points must be re-entered when the map is affixed to the table again. If the map has stretched or shrunk in the interim, the newly digitized points will be slightly off in their location when compared to previously digitized points. As a result errors occur on these maps, and are entered into the GIS database as well. The level of error in the GIS database is directly related to the error level of the source maps. In other words, the GIS database cannot be more accurate than the map it was created from.

Maps are meant to display information but they do not always accurately record locational information, for example, when a railroad, stream and road all go through a narrow mountain pass, the pass may actually be depicted wider than its actual size to allow for the three symbols to be drafted in the pass. Discrepancies across map sheet boundaries can cause discrepancies in the total GIS database e.g. roads or streams that do not meet exactly when two map sheets are placed next to each other. User error causes overshoots, undershoots (gaps) and spikes at intersection of lines. User fatigue and boredom also contribute to poor quality results from the digitizing process.
Editing errors from digitizing
Some errors can be corrected automatically such as small gaps at line junctions and overshoots and sudden spikes in lines. Error rates depend on the complexity of the map and are high for small scale, complex maps.

(b) Scanners
An image scanner—often abbreviated to just scanner—is a device that optically scans images, printed text, handwriting, or an object, and converts it to a digital image. Common examples found in offices are variations of the desktop (or flatbed) scanner where the document is placed on a glass window for scanning. Mechanically driven scanners that move the document are typically used for large-format documents, where a flatbed design would be impractical.

Modern scanners typically use a charge-coupled device (CCD) or a Contact Image Sensor (CIS) as the image sensor, whereas older drum scanners use a photomultiplier tube as the image sensor. A rotary scanner, used for high-speed document scanning, is another type of drum scanner, using a CCD array instead of a photomultiplier. Other types of scanners are planetary scanners, which take photographs of books and documents, and 3D scanners, for producing three-dimensional models of objects.

Another category of scanner is digital camera scanners, which are based on the concept of reprographic cameras. Due to increasing resolution and new features such as anti-shake, digital cameras have become an attractive alternative to regular scanners. While still having disadvantages compared to traditional scanners (such as distortion, reflections, shadows, low contrast), digital cameras offer advantages such as speed, portability and gentle digitizing of thick documents without damaging the book spine. New scanning technologies are combining 3D scanners with digital cameras to create full-color, photo-realistic 3D models of objects.

Video scanner
These are essentially television cameras, with appropriate interface electronics to create a computer-readable dataset. They are available in either black and white (monochrome) or color (16bit, 32 bit, 64bit etc) and are extremely fast (scan times of under 1 second). Video scanners are relatively inexpensive with prices ranging from $500 - $10,000. Video scanners produce a raster array of brightness (or color) values, which are then processed much like any other raster array. Typical data arrays from video scanners are of the order of 250 to 1000 pixels on a side. They typically have poor geometrical and radiometrical characteristics, including various kinds of spatial distortions and uneven sensitivity to brightness across the scanned field. Video scanners are difficult to use for map input because of problems with distortion and interpretation of features.

Electromechanical scanner
Unlike the video scanning systems, electromechanical systems are typically more expensive ($10,000 to 100,000) and slower, but can create better quality products. One common class of scanners involves attaching the graphic or map to a drum and is called a drum scanner. As the drum rotates about its axis, a scanner head containing a light source and photodetector reads the reflectivity of the target graphic, and digitizing this signal, creates a single column of pixels from the graphic. The scanner head moves along the axis of the drum to create the next column of pixels, and so on through the entire scan. This controls distortion by bringing the single light source and detector to position on a regular grid of locations on the graphic. Scanning systems may have a scan spot size of as little as 25 micrometers, and be able to scan graphics of the order of 1 meter on a side.
An alternative mechanism involves an array of photodetectors which extract data from several rows of the raster simultaneously. The detector moves across the document in a swath. When all the columns have been scanned, the detector moves to a new swath of rows. This is referred to as the flatbed scanner.

**Requirements for scanning**

- Documents must be clean (no smudges or extra markings)
- Lines should be at least 0.1 mm wide
- Complex line work provides greater chance of error in scanning
- Text may be accidentally scanned as line features
- Contour lines cannot be broken with text
- Automatic feature recognition is not easy (two contour lines vs. road symbols)

Special symbols must be recognized and dealt with. If good source documents are available, scanning can be an efficient time saving mode of data input.

**CONVERSION FROM OTHER DIGITAL SOURCES**

This involves transferring data from one system to another by means of a conversion program. This is useful since more and more data is becoming available in magnetic media such as:

- digital elevation models (DEMs)
- TIGER and other census related data
- data from CAD/CAM systems (AutoCAD, DXF)
- data from other GIS

These data generally are supplied on digital media that must be read into the computer.
Automated Surveying

This directly determines the actual horizontal and vertical positions of objects. There are basically two kinds of measurements are made namely distances and directions which are used in computing coordinates or locational information of features on the earth's surface. Traditionally, distance measuring involved pacing, chains and tapes of various materials with direction measurements were made with transits and theodolites. Modern surveyors have a number of automated tools to make distance and direction measurements easier. Electronic systems such as total stations measure distance using the time of travel of beams of light or radio waves. The data is downloaded to a host computer at the end of each session for direct input to GIS and other programs.

Global Positioning System (GPS)

This is a new tool for determining accurate positions on the surface of the earth. GPS is an all weather system that measures position in three dimensions (X, Y, Z) and with a fourth dimension of time. It operates on the concept of trilateration in which at a minimum of three satellites are required to determine position. GPS depends on precise information about the orbits of the satellites. A radio receiver with appropriate electronics is connected to a small antenna, and depending on the method used, in one hour to less than 1 second, the system is able to determine its location in 3-D space current GPS positional accuracies are order 2 to 10 m with standard equipment and as small as 1 cm with “survey grade” receivers. Accuracy will continue to improve as more satellites are placed in orbit and experts fine tune the software and hardware and the integration of GPS and GLONASS satellites by 3G GPS receivers.

CRITERIA FOR CHOOSING MODES OF INPUT

- The type of data source
  - images favor scanning
  - maps can be scanned or digitized
- The database model of the GIS
  - scanning easier for raster, digitizing for vector
- The density of data
  - dense linework makes it difficult but not impossible for digitizing
- Expected applications of the digital data e.g. GIS applications

RASTERIZATION AND VECTORIZATION

Rasterization of digitized data

\[\text{for some data, entry in vector form is more efficient, followed by conversion to raster}\]
\[\text{we might digitize the county boundary in vector form by}\]
• mounting a map on a digitizing table
• capturing the locations of points along the boundary
• assuming that the points are connected by straight line segments
• this may produce an ASCII file of pairs of xy coordinates which must then be processed by the GIS, or the output of the digitizer may go directly into the GIS
• the vector representation of the boundary as points is then converted to a raster by an operation known as vector-raster conversion
  ➢ the computer calculates which county each cell is in using the vector representation of the boundary and outputs a raster
• digitizing the boundary is much less work than cell by cell entry
• most raster GIS have functions such as vector-raster conversion to support vector entry
  ➢ many support digitizing and editing of vector data

Vectorization of scanned images

• for many purposes it is necessary to extract features and objects from a scanned image
  o e.g. a road on the input document will have produced characteristic values in each of a band of pixels
  o if the scanner has pixels of 25 microns = 0.025 mm, a line of width 0.5 mm will create a band 20 pixels across
  o the vectorized version of the line will be a series of coordinate points joined by straight lines, representing the road as an object or feature instead of a collection of contiguous pixels
• successful vectorization requires a clean line scanned from media free of cluttering labels, coffee stains, dust etc.
  o to create a sufficiently clean line, it is often necessary to redraft input documents
    ➢ e.g. the Canada Geographic Information System redrafted each of its approximately 10,000 input documents
• since the scanner can be color sensitive, vectorizing may be aided by the use of special inks for certain features
• although scanning is much less labor intensive, problems with vectorization lead to costs which are often as high as manual digitizing
  o two stages of error correction may be necessary: 1. edit the raster image prior to vectorization 2. edit the vectorized features

G. INTEGRATING DIFFERENT DATA SOURCES

Formats

• many different format standards exist for geographical data
• some of these have been established by public agencies
  o e.g. the USGS in cooperation with other federal agencies is developing SDTS (Standard Data Transfer Standard) for geographical data, will propose it as a national standard in 1990
  o e.g. the Defense Mapping Agency (DMA) has developed the DIGEST data transfer standard
• some have been defined by vendors
  o e.g. SIF (Standard Interchange Format) is an Intergraph standard for data transfer
• see Unit 69 for more on GIS standards
• a good GIS can accept and generate datasets in a wide range of standard formats

Projections

• there are many ways of representing the curved surface of the earth on a flat map
  o some of these map projections are very common, e.g. Mercator, Universal Transverse Mercator (UTM), Lambert Conformal Conic
  o each state has a standard SPC (State Plane Coordinate system) based on one or more projections
see Unit 27 for more on map projections
- a good GIS can convert data from one projection to another, or to latitude/longitude
- input derived from maps by scanning or digitizing retains the map's projection
- with data from different sources, a GIS database often contains information in more than one projection, and must use conversion routines if data are to be integrated or compared

Scale

- data may be input at a variety of scales
- although a GIS likely will not store the scale of the input document as an attribute of a dataset, scale is an important indicator of accuracy
- maps of the same area at different scales will often show the same features
  - e.g. features are generalized at smaller scales, enhanced in detail at larger scales
- variation in scales can be a major problem in integrating data
  - e.g. the scale of most input maps for a GIS project is 1:250,000 (topography, soils, land cover) but the only geological mapping available is 1:7,000,000
  - if integrated with the other layers, the user may believe the geological layer is equally accurate
  - in fact, it is so generalized as to be virtually useless

Resampling rasters

- raster data from different sources may use different pixel sizes, orientations, positions, projections
- resampling is the process of interpolating information from one set of pixels to another
- resampling to larger pixels is comparatively safe, resampling to smaller pixels is very dangerous
Plotters
Vdu’s

Tablet PCS

- Stereo headset
- Mono headset
- Stereo speakers
- Internal antennas
- Jog dial (up/down, enter)
- Pen with barrel button
- Power switch
- Programmable launch buttons (Qmenu, esc, tab, PIM)
- 3 pen button (TPP, journal, rotate screen)
- VGA
- 2 USB
- RJ45/RJ11
- PC card/compact flash slot
- DC power
Data storage: