CHAPTER 3

A Brief Overview on Geographic Information System

3.1 Introduction

The Information System is used to input, store, retrieve, manipulate, analyze and produce output geographically referenced data or geospatial data, in order to help in making decision for planning, management of land use, natural resources, environment, transportation, urban facilities, and other administrative records [23].

Geographical Information Systems (GIS) are computer-based systems that enable users to collect, store, and process, analyze and present spatial data [23, 25]. It provides an electronic representation of information, called spatial data, about the natural and man-made objects of the Earth. A GIS generally refers the real-world spatial data into a coordinate system. The features can be separated into different layers. The GIS system generally stores the information for each category in a separate "layer" for easy maintenance, analysis, and visualization about the earth object. For instance, different layers can be used to represent different information such as census data, demographics information, environmental and ecological data, roads, land use, river direction, drainage system of a city and flood. A GIS can also be used to store attribute data of the map features which is generally descriptive information. It also helps to examine both spatial and attribute data at the same time. It allows the user to search the information of attribute data and relate it to the spatial data. Thus, a GIS is used to produce maps and reports by combining geographic and other types of data, which enable the users to manage, collect and interpret information about a specific location in a planned and organized way. A GIS can be defined as a computer system capable of assembling, storing, manipulating, and displaying geographically referenced information about the earth observation.

The GIS systems are very dynamic which allow rapid updating, analysis, and display the earth observation by the satellite. It uses the data from different sources such as satellite imagery, aerial photos, maps, ground surveys, and global positioning systems (GPS) [25].

3.2 Geographic Information System

GIS is generally a computerized information system like any other database, but with a significant difference: all information in GIS must be linked to a geographic (spatial) reference (latitude/longitude, or other spatial coordinates).

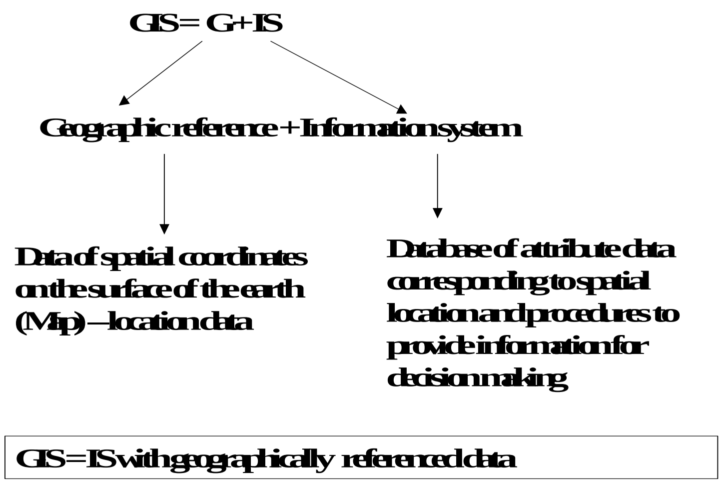

Fig 3.1: Geographic Information system

There are many different definitions of GIS, as different users stress different aspects of its use.

For example:

ESRI defined GIS as an organized collection of computer hardware, software, geographic data and personnel designed to efficiently capture, store, update, manipulate, analyze and display geographically referenced information.

(ESRI also provided a simpler definition of GIS as a computer system capable of holding and using data describing places on the earth's surface).

Duecker defined GIS as a special case of information systems where the database consists of observations on spatially distributed features, activities or events, which are definable in space as points, lines or areas. A GIS manipulates data about these points, lines or areas to retrieve data for ad hoc queries and analyses.

The United States Geological Survey (USGS) defined provided A GIS as a computer hardware and software system designed to collect, manage, analyze and display geographically (spatially) referenced data. This definition is a fairly comprehensive and is suitable for agricultural applications of GIS.

3.3 Basic Functions of GIS

3.3.1 Data Collection & Processing

The collection of data about earth observation is closely related to the disciplines of engineering, photogrammetry, remote sensing and the processes that convert analogue data into digital representations. Remote sensing is the field which converts the spatial data set from the raw base data. Surveys of the study area often need to be conducted for the data that cannot be obtained with remote sensing techniques, or to validate data thus obtained.

Traditional methods for obtaining the spatial data are from paper sources, which included manual digitizing and scanning of data set. Table 3.1 shows the main techniques and devices used for data collection. In recent days, there has been a significant availability and sharing of digital (geospatial) data from different sources. Various media and computer networks play an important role in the dissemination of this data, particularly the internet.

The data which are collected once in digital format may not be ready for use in the computerized system, because the format of data obtained from the data collecting process is not quite similar to the format required for storage and further use that is why some data conversion need to be carried out. This problem may also arise when the captured data represents only raw base data, out of which the real data objects of interest to the system still need to be constructed. For example, semi-automatic digitizing may produce only line segments, while the application's requirements are non overlapping polygons are needed. A build-and-verification phase would then be needed to obtain these from the captured lines [25, 26].

Method	Devices
Manual digitizing	- Coordinate entry via keyboard - Digitizing tablet with cursor - Mouse cursor on the computer monitor (heads-up digitizing) - (digital) photogrammetry
Automatic digitizing	Scanner
Semi-automatic digitizing	Line-following software
Input of available digital data	- CD-ROM or DVD-ROM - Via computer network or internet (including geo-web services)

Table 3.1: Spatial data in-put methods and devices used

3.3.2 Data Storage and Retrieval

The storing and retrieving of plays an important role in the processing and the understanding of that data. In general the spatial data is organized in layers depending on theme and /or scale of the data. For example, the data may be stored and retrieved in thematic categories, such as land use, topography and administrative sub-divisions, or according to map scale of the data. In a GIS, the features of spatial data set are generally represented with their (geometric and non-geometric) attributes and relationships. The geometry features of spatial data set are usually represented with primitives of the respective dimension: a mill as a point, an agricultural field as a polygon, a road as a line etc. The primitives follow either the vector, or the raster approach.

The representation of Vector data types of an object is through its boundary, thus dividing the space into different parts that are occupied by the respective objects. The raster approach subdivides space into (regular) cells, mostly as a square of dimension two or three. The cell indicates the real world features of the object, if it represents a discrete field. If a continuous field is considered, then the cell holds a representative value for that field. Table 3.2 lists advantages and limitations of raster and vector representations of images [25, 26].

The raster image is stored in a file as a long list of values, one for each cell, preceded by a small list of extra data ('file header') that informs how to interpret the long list. The order of the cell values in the list can be interpreted from left-to-right or from top-to-bottom.

Raster representation	*Vector representation*
Advantages	
- Simple data structure - Simple implementation of overlays - Efficient for image processing	- Efficient representation of topology - Adapts well to scale changes - Allows representing networks - Allows easy association with attribute data
Limitations	
- Less compact data structure - Difficulties in representing topology - Cell boundaries independent of feature boundaries	- Complex data structure - Overlay more difficult to implement - Inefficient for image processing - More update-intensive

Table 3.2: Raster and vector representations

3.3.3 Measurement & Analysis

The most significant features of a GIS are measurement and analysis of spatial data, i.e. the spatial data operators use to derive a new geo-information. One of the important uses of GIS is to support spatial decisions and spatial queries process. Spatial decision support systems (SDSS) is an information system comprised of a database (DB), GIS software, models, and knowledge engine which allow users to deal specifically with locational problems.

GIS data are generally grouped into layer based themes of the spatial data. The measurement and analysis functions of a GIS use the spatial and non-spatial attributes of the data in a spatial database to provide result to user depending on the queries. The functions of GIS are used for maintenance of the data, and for analyzing the data in order to retrieve information from it. Analysis of spatial data can be defined as getting new information from the existing spatial data. The domain of road construction can be cited for instance. In a hill station this is a complex engineering task with many cost factors such as the number of tunnels and bridges needed to be constructed, the total length of the tarmac, and the volume of rock and soil to be moved to get a perfect road. GIS helps to analyze and compute such types of costs on the basis of an up-to-date digital elevation model and soil map [25, 26].

The exact nature of maintenance and analysis spatial data depends on the application requirements, but analytical and computational functions operate on both spatial and non-spatial data.

3.3.4 Output and Visualization

The presentation of spatial data, either in hard copy or on-screen, in maps or in tabular displays, or as ‘raw data’, is closely related to the disciplines of cartography, printing and publishing. The presentation of spatial data is an end-product, for example as a printed atlas, or an intermediate product, as in spatial data made available through the internet [25, 26].

3.4 Types of GIS

A GIS application can be classified under more than one type, the following are different types of GIS application. [25]

3.4.1 Multimedia GIS

GIS that allow the user to access a wide range of geo referenced multimedia data (sounds and videos etc.) by selecting images of geo referenced data from map base. A map serving as the primary index to multimedia data in a multimedia geo-representation is termed a hyper map. Multimedia and virtual geo-representations can be stored either in extended relational databases, object databases or in application-specific data stores.

3.4.2 Web GIS

GIS that allow widespread access to the Internet, the internet browsers and the explosion of commoditized geographic information have made it possible to develop new forms of multimedia geo-representations on the Web. Many current geometrics' solutions are Web-based overtaking the traditional Desktop environment and most future ones are expected to follow the same direction.

3.4.3 Virtual Reality GIS

Virtual Reality GISs have been developed to allow the creation, manipulation and exploration of geo-referenced virtual environments, e.g., using VRML modelling (Virtual Reality Modelling Language). Virtual Reality GIS can be also Web-based. Applications include 3D simulation for planning (to experiment with different scenarios).

3.5 Data Types:

Data types used in the GIS domain are mainly known as spatial data and non-spatial data. Spatial data either may be in raster form or in vector form. The representation of Vector data is in the form of points, lines and polygons. And the representation of Raster data is the presence of continuous cells or pixels of specific size in the form of digital numbers. They are usually found in the form of scanned images, satellite images or aerial photographs. Non-spatial data are all statistical data interlinked to a vector data in the attribute table. For example - the demography of cities, length of a road, direction of drainage, length of a river, size of school building polygon etc. There are various advantages and limitations of both raster and vector data. The Vector and raster forms are the major representation models for geometric information [23-25].

3.5.1 Vector Form and its Data Structure

The representations of objects in a map are combinations of a point, edge and area. The vector form of spatial data is provided by the above geometric factors and the attributes are assigned to points, edges and areas. The data structure is specified for the vector form as follows [24, 25].

- ✓ A point is represented by geographic coordinates.
- ✓ An edge is represented by a series of line segments with a starting point and an end point.
- ✓ A polygon is represented by the sequential edges of a boundary.

The inter-relationship between points, edges and areas is called a topological relationship. Any change in a point, edge or area will affect the other factors through the topological relationship. Thus the data structure should be specified properly to fulfill the relationship,

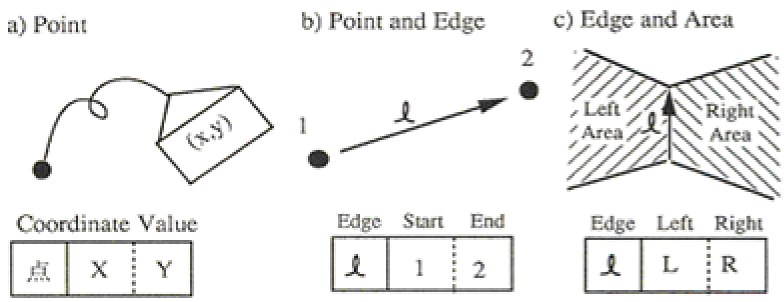

Fig 3.2: Data Structure of Vector Data (Topological Relations)

3.5.2 Raster Form and its Data Structure

An object is represented in the raster form by dividing it into a group of regularly pixels to which the attributes are assigned. The remote sensing data format is basically identical to the raster data format. As the pixels are generated regularly the coordinates correspond to the pixel number and line number, which is usually represented in a matrix form as shown below [25, 26].

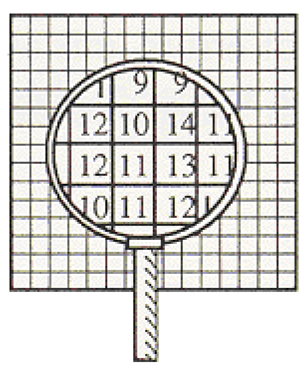

Fig 3.3: Data Structure of Raster Data

3.6 Application of GIS in Real World

There are number of domain areas where GIS plays an important role, some of the major areas of GIS application can be categorised as follows [24 - 26]:-

1) ***Facilities Management*****:** Large scale precise maps and network analysis are used mainly for utility management.
2) ***Environment and Natural Resources Management*****:** The overlay techniques in combination with aerial photographs and satellite images are used for impact analysis on environmental and natural resource management.
3) ***Street Network*****:** For the vehicle movement, locating house, school, hospital, street etc in a city the large or medium scale maps and spatial analysis are required.
4) ***Planning and Engineering*****:** Large or medium scale maps are used by civil engineering department for planning a model town, city.
5) ***Disaster Management*****:** GIS is being used to manage environmental problems and specifically in disaster relief.
6) ***Health / Medical Resource Management*****:** GIS is important for the proper planning and analysis of the health services, e.g. the NHS (National Health Service). The package is used to plan and examine a number of issues including GP surgeries.

3.7 Chapter Summary

GIS has the ability to analyze a large amount of spatial data within a short time. Without the use of GIS these voluminous data would have become useless. To analyze a one time remote sensing data may take years - thus before completion of processing the earlier data a number of multi date data may be demanded for analysis [26].