Chapter 13 Geographic Information System(GIS)

13.1 GIS and Remote Sensing

a. GIS in remote sensing

For the users of remote sensing, it is not sufficient to display only the results obtained from image processing. For example, to detect land cover change in an area is not enough, because the final goal would be to analyse the cause of change or to evaluate the impact of change. Therefore the result should be overlaid on maps of transportation facilities and land use zoning as shown in Figure 13.1.1. In addition, the classification of remote sensing imagery will become more accurate if the auxiliary data contained in maps are combined with the image data.

In order to promote the integration of remote sensing and geographic data, **geographic information system** (GIS) should be established in which both the image and graphic data are stored in a digital form, retrieved conditionally, overlaid on each other and evaluated with the use of a model.

Figure 13.1.2 shows a comparison between the computer assisted GIS and the conventional analog use of maps.

b. Function of GIS

The following three functions are very important in GIS.

(1) To store and manage geographic information comprehensively and effectively

(2) To display geographic information depending on the purpose of use

(3) To execute query, analysis and evaluation of geographic information effectively

At present, the following research and development have been undertaken. In this book the following technologies , a part from, visualization will be described.

(1) Model and data structure for GIS
(2) Data input and edition
(3) **Spatial query**
(4) **Spatial analysis**
(5) **Visualization**

Figure 13.1.1 An overlay of RS data and map data (elevation data)

Interpretation can be made easier by an overlay of geo-corrected RS data with a map data

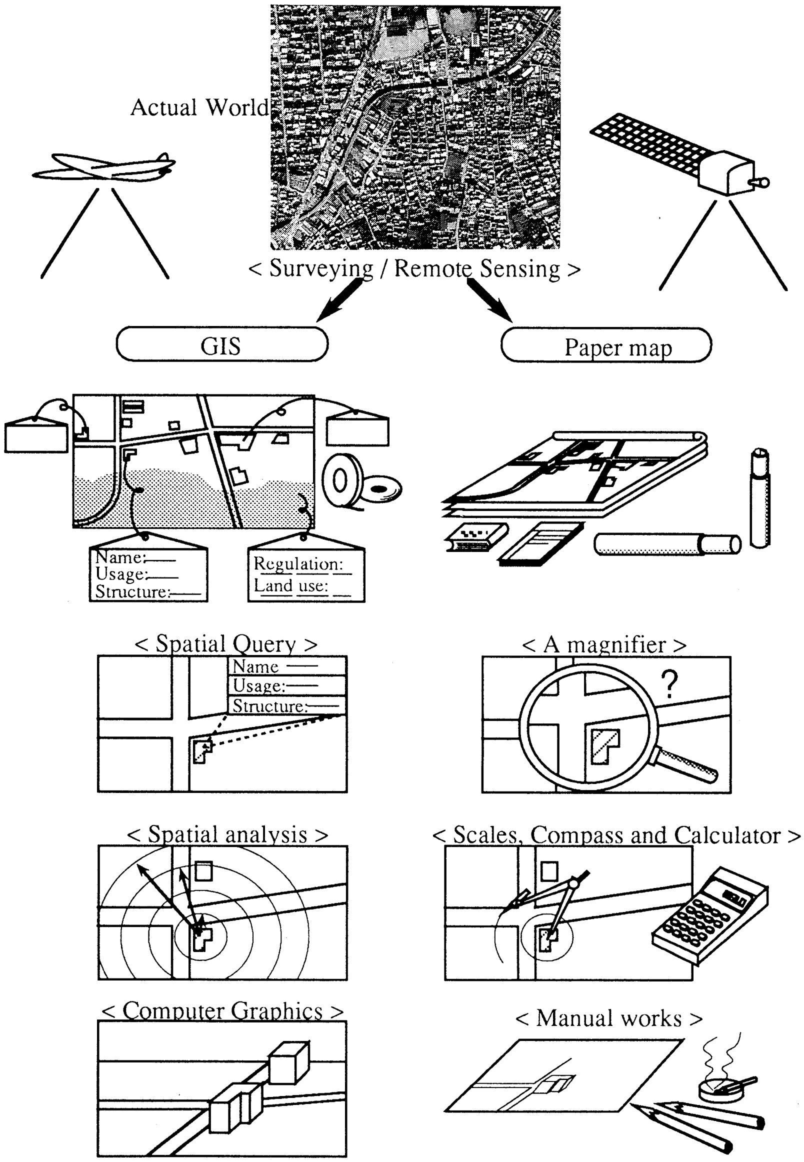


Figure 13.1.2 Geo-information management with GIS - A comparison with paper maps -

13.2 Model and Data Structure

a. Requirement of the Model and Data Structure

In order to process and manage geographic information by computers, it is necessary to describe the spatial location and distribution, as well as the attributes and characteristics, according to a specified form, termed a **spatial representation model** with a standardized **data structure**.

b. Modeling and Data Structure

Geographic information can be represented with **geometric information** such as location, shape and distribution, and **attribute information** such as characteristics and nature, as shown in Figure 13.2.1.

Vector and raster forms are the major representation models for geometric information.

(1) Vector form and its data structure

Most objects on a map can be represented as a combination of a **point** (or node), **edge** (or arc) and **area** (or polygon). The **vector form** is provided by the above geometric factors. The attributes are assigned to points, edges and areas.

The data structure is specified for the vector form as follows.

A point is represented by geographic coordinates. An edge is represented by a series of line segments with a start point and an end point. A polygon is defined as 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 influence other factors through the **topological relationship**. Therefore the data structure should be specified to fulfill the relationship, as for the example as shown in Figure 13.2.2.

(2) Raster form and its data structure

In the raster form, the object space is divided into a group of regularly spaced grids (sometimes called pixels) to which the attributes are assigned. The **raster form** is basically identical to the data format of remote sensing data.

As the grids are generated regularly, the coordinates correspond to the pixel number and line number, which is usually represented in a matrix form as shown in Figure 13.2.3.

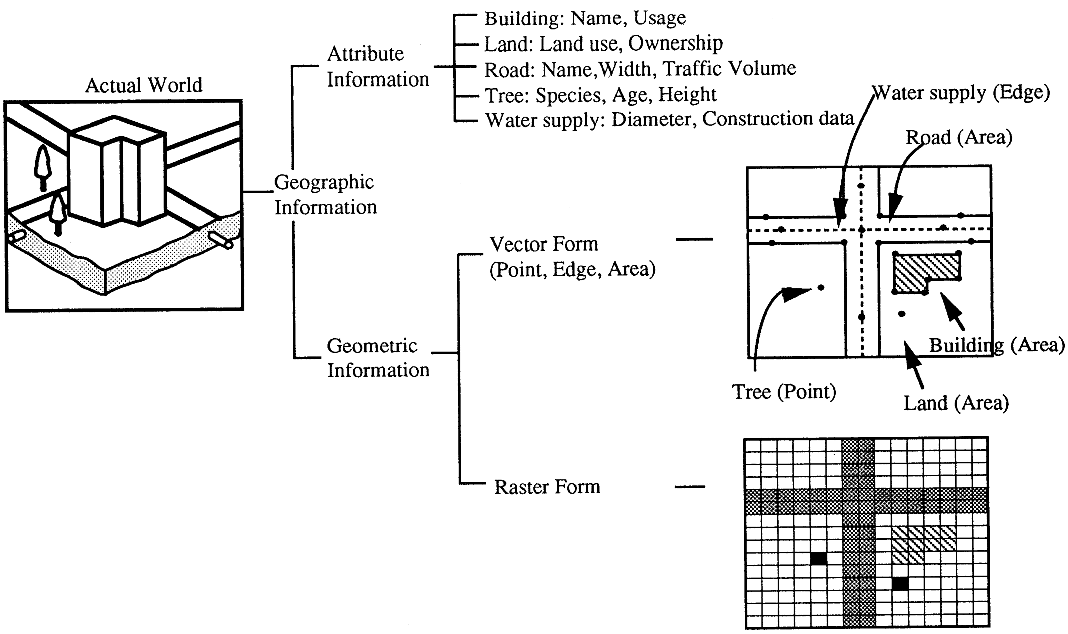


Figure 13.2.1 Representation of geographic information
- Attribute information and geometric information -

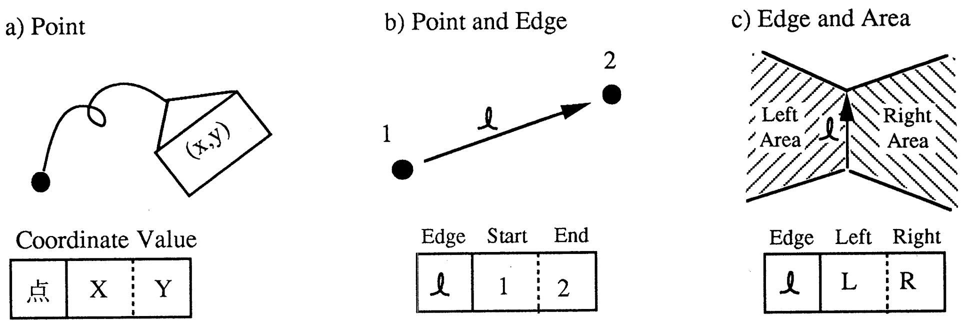


Figure 13.2.2 A basic data structure of vector data
- Representation of topological relations -

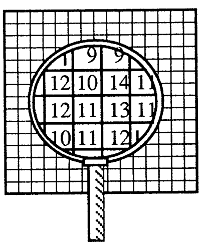


Figure 13.2.3 A basic data structure of raster data

13.3 Data Input and Editing

a. Role of data input and editing

Data acquisition occupies about 80 percent of the total expenditure in GIS. Therefore data input and editing are very important procedures for the use of GIS.

b. Initial data input

Geometric data as well as attribute data are input by the following methods.

(1) Direct data acquisition by land surveying or remote sensing

Vector data can be measured with digital survey equipment such as total stations or analytical photogrammetric plotters. **Raster data** are sometime obtained from remote sensing data.

(2) Digitization of existing maps (see Figure 13.3.1)

Existing maps can be digitized with a **scanner** or tablet **digitizer**. Raster data are obtained from a scanner while vector data are measured by a digitizer. In GIS, raster data and vector data are frequently converted to vector data and raster data respectively, which are called **raster/vector conversion** and **vector/raster conversion** respectively.

c. Editing

Editing is needed to correct, supplement and add to the initial input data through interactive communication on a graphic display using the following procedures.

(1) to input manually or interactively those complicated attributes which are not effectively digitized in the initial input stage.

(2) to correct errors of input data or to supplement with other data.

d. Problems in Data Input and Editing

There are two main problems.

(1) Manual operations

It is difficult to automate data input and editing because of unremovable noise and incomplete original maps, which result in a large amount of manual work with resultant inefficiencies in time and cost.

(2) Unreliability of input data

As the input involve many kinds of errors, mistakes and misregistration because of the manual input, further effort should be applied to obtain data high quality and reliability.

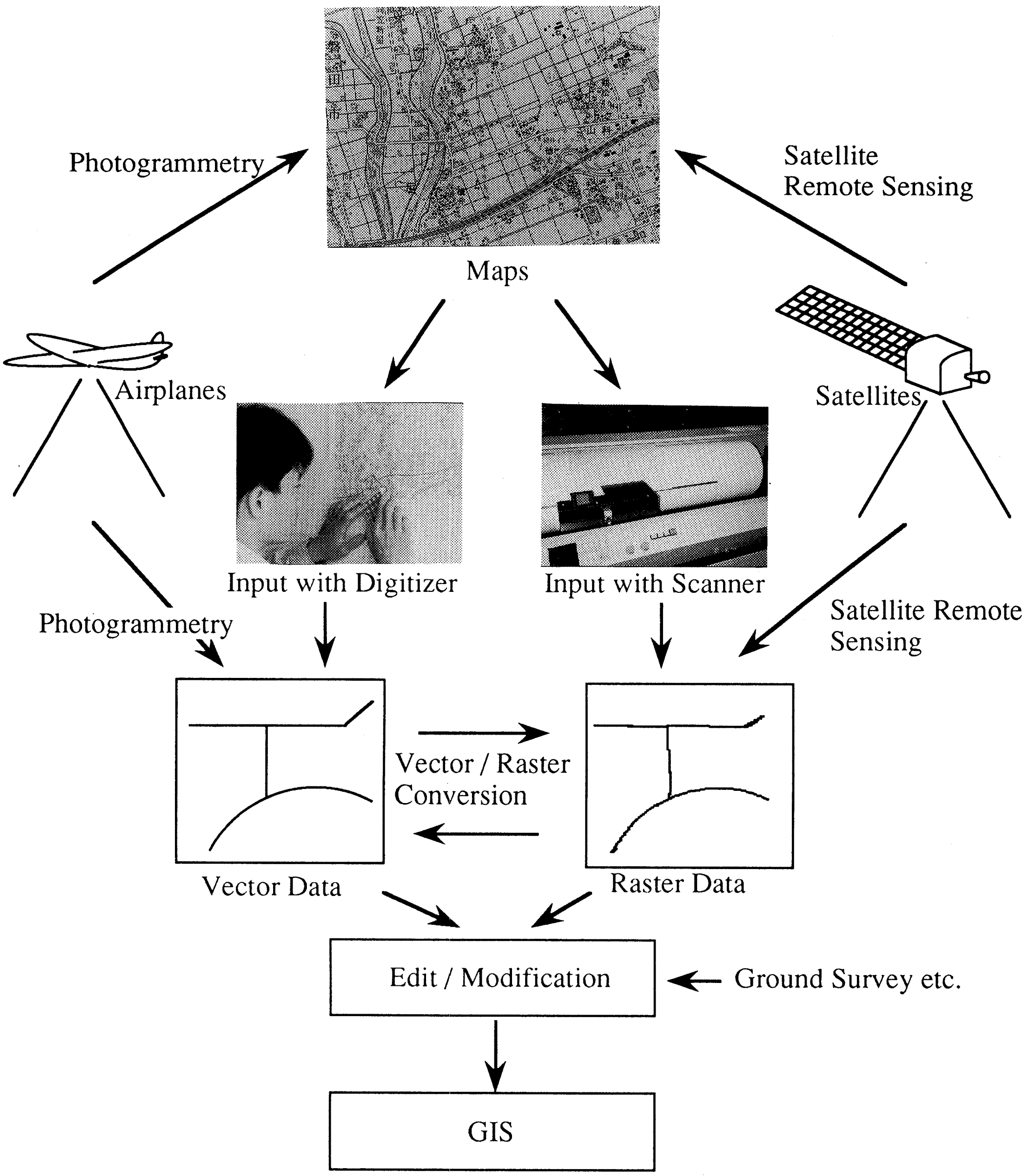


Figure 13.3.1 A process of input / update of geographic data

13.4 Spatial Query

a. Types of spatial query

Spatial query is a search of the data to satisfy a given condition. There are two types of spatial query.

(1) **Query of attribute data**

A spatial distribution or an area will be searched with respect to a given attribute of interest.

(2) **Query of geometric data**

With a given geometric condition for example location, shape or intersection, all data that satisfy the condition will be searched. In the case of a vector data form, to search an area which includes a given point, and to find all line segments which intersect a given line would be a typical query of geometric data.

In the case of raster form of data, it will be easier to search any attribute and geometric data based on a given grid.

Figure 13.4.1 shows an example of a query of attribute data in the aster form, in which the areas with slope gradient of greater than 30 degrees are located.

Figure 13.4.2 shows an example of a query of geometric data in which the area was searched that includes a point, as given by a cursor.

b. Data Structure for High Speed Query

It is important to develop a data structure which allows for high speed query, because the data volume is usually very huge.

For example, in order to search all points which are included in an area, it is necessary to check many points whether those points are included in the area or not.

Tree structure and **block structure** are typical data structures used to save time of query. Figure 13.4.3 shows the block structure for solving a point-in-polygon problem, where only the block that includes a polygon should be checked and searched instead of all other blocks.

The **Quadtree** structure has been proposed and used not only for high speed query but also for data compression.

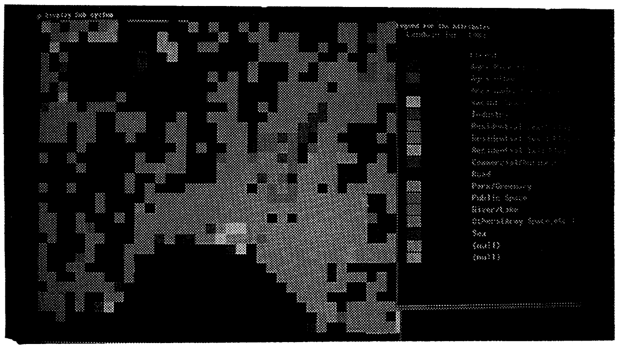


Figure 13.4.1 An example of query of attribute information
(Raster data)

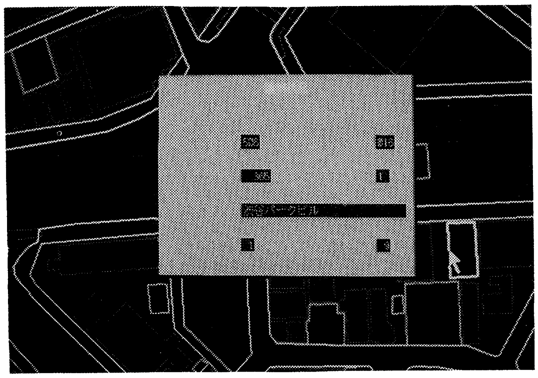


Figure 13.4.2 An example of query of geometric information
(Vector data)

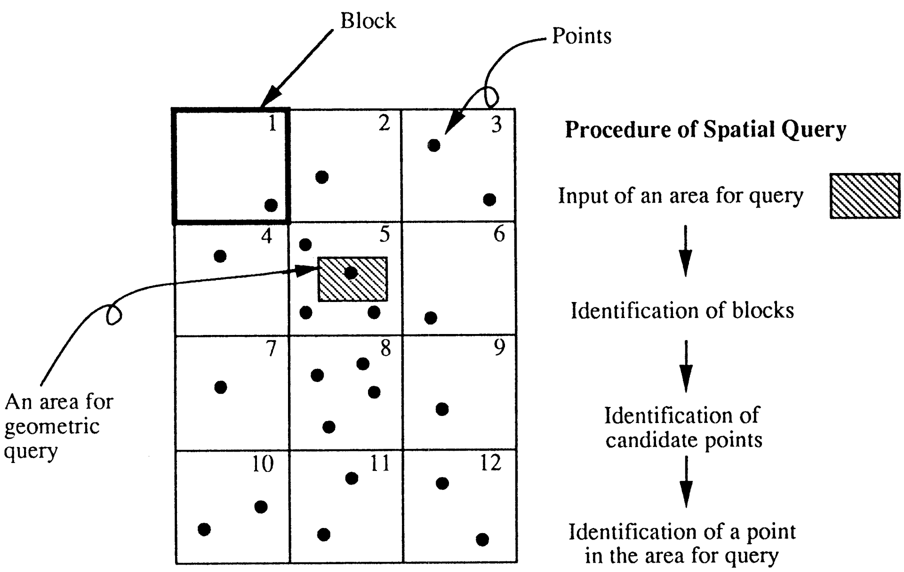


Figure 13.4.3 An example of data structure for faster geometric query
(Block structure)

13.5 Spatial Analysis

a. Concept of Spatial Analysis

Spatial analysis is used to produce additional geographic information using existing information or to enhance the spatial structure or relationship between geographic information. Many techniques have been proposed, as follows.

b. Production of Additional Geographic Information

The following three techniques are very often used in GIS.

(1) **Overlay** technique (see Figure 13.5.1)

Various geographic data comprised of multiple layers are overlaid with logical operations including logical addition or logical multiplication. For example, a hazard risk area of soil erosion can be estimated by overlaying deforested and slope gradient maps in a mountainous area.

(2) **Buffering** technique (see Figure 13.5.2)

Buffering is to find an area the within a certain distance from a given point or a line. For example noise polluted areas will be extracted by buffering an area within 30 meter distance from a trunk road.

(3) **Volonoi tessellation**

An area may be divided in a group of "influential areas" termed Volonoi tessellation, that can be formed by bisectors between spatially distributed points. For example, a school zone can be drawn by Volonoi tessellation between differently located schools.

c. Statistical Analysis for Spatial Structure

Spatial auto-correlation is one of the statistical techniques to find the spatial structure of geographic information. Spatial auto-correlation is a correlation factor between two differently located events. High accuracy spatial interpolation can be executed with a lower density of samples in the case of high spatial auto-correlation.

d. Combined Technique

Figure 13.5.3 shows an example of a combined technique using remote sensing, buffering and overlay. In this example, the land use change ratio is tabulated with respect to accessibility to a railway station and land use zoning.

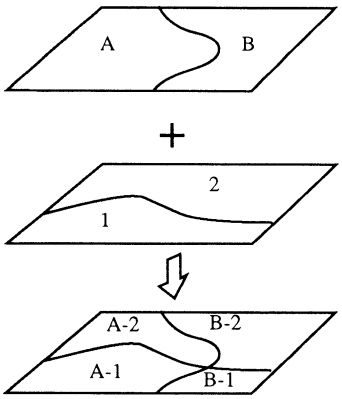


Figure 13.5.1 A concept of overlay

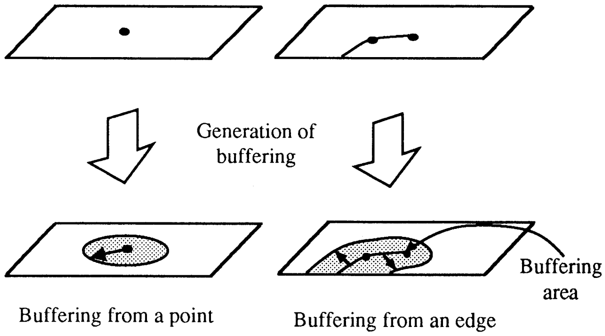
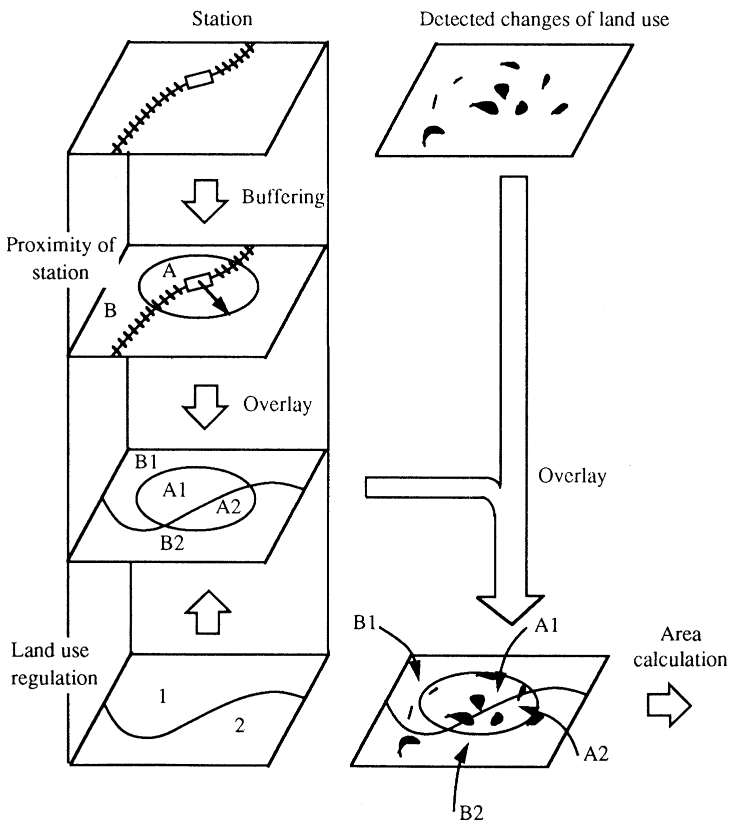


Figure 13.5.2 A concept of buffering



Ratio of land use change by zones

Characteristics of zone		Ratio of change in area
Proximity to station	Land use regulation	
A	1	30%
A	2	10%
B	1	10%
B	2	5%

Figure 13.5.3 An examples of a process of land use change analysis

13.6 Use of Remote Sensing Data in GIS

Remote sensing data after geometric correction, can be overlaid on other geographic data in a raster form. In GIS, there are two uses of use of remote sensing data; as classified data and as image data.

a. Use of classified data

Land cover maps or vegetation maps classified from remote sensing data can be overlaid onto other geographic data, which enables analysis for environmental **monitoring** and its change.

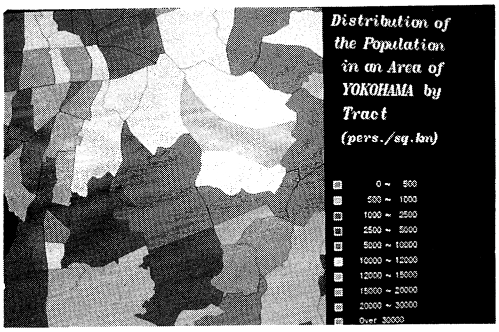
Figure 13.6.1 shows a case study in which statistical data with lower spatial resolution are reallocated with a higher spatial resolution using the fact that the remotely sensed data have higher resolution than the statistical data.

b. Use of image data

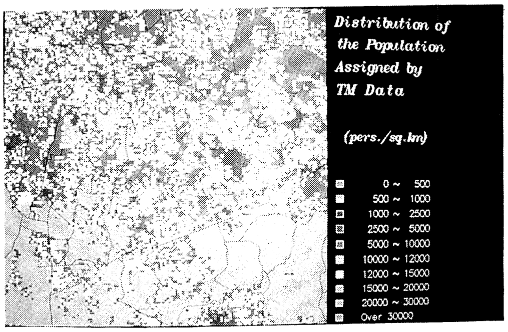
Remote sensing data will be classified or analyzed with other geographic data to obtain a higher accuracy of classification. Figure 13.6.2 shows a comparison between two results of classification without the use of map data and with the use of map data. If ground height and slope gradient are given as map data, rice fields, for eg., can be checked and located only in flat and low land areas. Forest areas and mangrove area are also classified with less errors if map data are combined with remote sensing data.

Image data are sometimes also used as image maps, with an overlay of political boundaries, roads, railways etc. Such an image map can be successfully used for visual interpretation.

If a digital elevation model (DEM) is used with remote sensing data, shading corrections in mountainous areas can be made by dividing by cos q (where q : angle between sun light and the normal to the sloping surface).



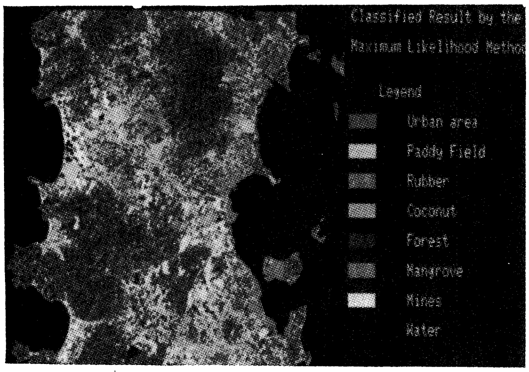
a) Existing population data by zones



b) Estimated distribution of population

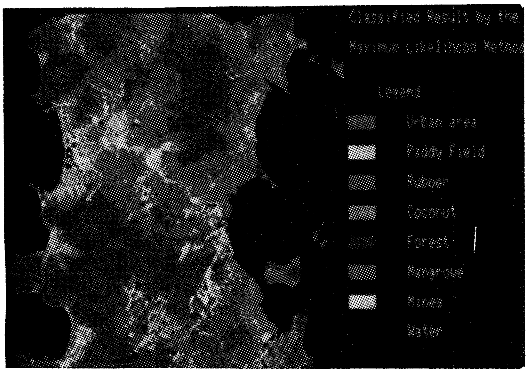
with a land cover map derived from LANDSAT TM data

Figure 13.6.1 An estimation of spatial distribution of population using LANDSAT TM data



a) Classification without using map data

Erroneous classification
(Paddy field on steep slopes)



b) Classification with the use of map data

The erroneous classification on steep slopes is corrected by using a knowledge the no paddy fields are on steep slopes

Figure 13.6.2 Land use classification with auxiliary use of map data

13.7 Errors and Fuzziness of Geographic Data and their Influences on GIS Products

a. Errors and Fuzziness of Geographic Data

Table 13.7.1 shows various errors in geographic data with respect to error sources. Of those errors, errors due to the data input method can be avoided by a proper control and check system, while errors due to measurement methods are difficult to avoid completely. It is necessary for users to evaluate the data errors and their influences by **sensitivity** analysis. Quality control of geographic data in also very essential in GIS.

b. Influences of Errors and Fuzziness

Influences of errors and fuzziness are explained in the following two examples.

(1) Influence on spatial query
Consider a case to check any underground pipe which may be damaged by excavation at a road construction site, as shown in Figure 13.7.1. If only the geometric relationship is checked, there will be no problem as shown in Figure 13.7.1 (a). However, one should consider the uncertainly or error of excavation as well as the pipe, which will make possible the problem as shown in Figure 13.7.1 (b).

(2) Influence on spatial analysis
Consider a case to select suitable land for rice paddy fields by overlaying a slope gradient map, soil map and irrigation areas in Jogjakarta, Indonesia. The area of suitable land will change depending it that there is uncertainly along the boundary of the overlaid areas. If there is 120 meter width of uncertainly along the boundary, the area of suitable land will reduce by about 50 per cent. Thus this uncertainly or error should always be considered.

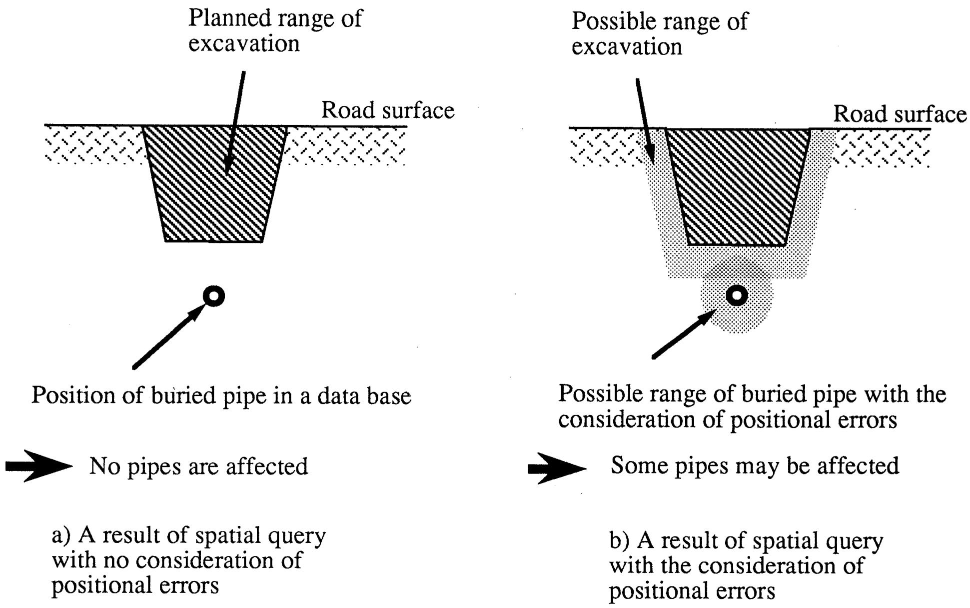
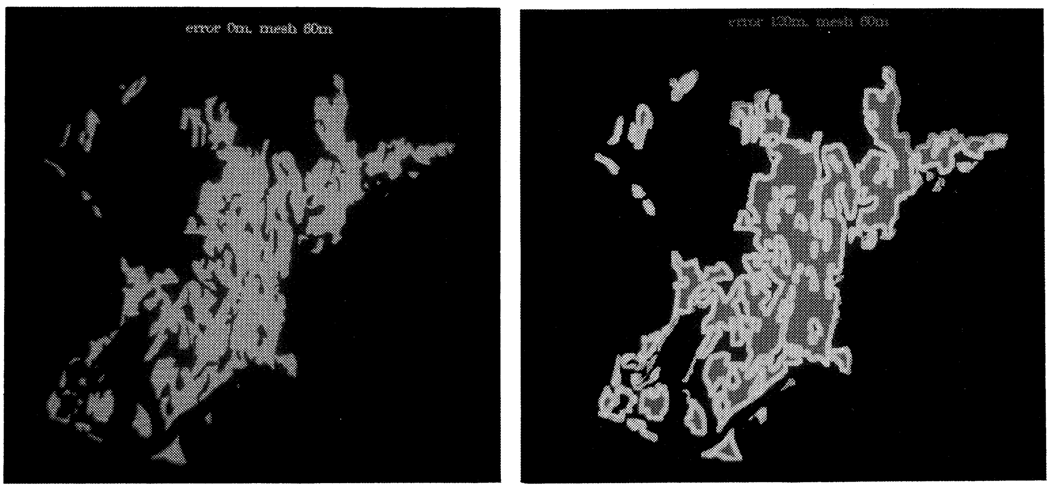


Figure 13.7.1 An example of influence of positional errors on spatial query



a) with the consideration of positional errors
result of land suitability analysis for paddy field development through an overlay of soil map, slope map, and i irrigation map (1:100,000)

b) with the consideration of positional errors
A possible influence is estimated of positional errors in the map data.
It is assumed that there are " transition bands" along the boundaries of classed of the map, whose width is about one millimeter on the maps.
More than 50% of the area suitable for paddy field are affected.

Figure 13.7.2 An example of influence of positional errors on spatial analysis