WORKING PAPER 532

A NEW APPROACH TO REGIONAL MANAGEMENT OF UPLAND VEGETATION USING REMOTE SENSING AND A DESKTOP GEOGRAPHICAL INFORMATION SYSTEM

J. HOGG

School of Geography
University of Leeds
Leeds LS2 9JT

B.S. HOYLE

Department of Electrical and Electronic Engineering
University of Leeds
Leeds LS2 9JT

M.N. GAHEGAN

S.A. ROBERTS

Department of Computer Science
University of Leeds
Leeds LS2 9JT

School of Geography
University of Leeds
Leeds LS2 9JT

October 1989

1.0 INTRODUCTION

Remote sensing from aircraft and spacecraft provides a primary source of geographic information for inventory, management and planning of vegetation over extensive, marginal tracts of Western Europe. As the technology of remote sensing has advanced, both the quality and quantity of useful information that can be derived from it for applications in ecology, agriculture, forestry and hydrology have increased. By exploiting the spectral, spatial and temporal variables associated with remote sensors on the SPOT-1 satellite, for example, a broad range of information about the nature, amount and distribution of vegetation, and its seasonal and annual change, can be derived with significant improvements in detail over that derived from earlier satellite imagery. Moreover automatic generation of digital elevation models (DEM's) from stereo-images of SPOT-1 using image correlation is now possible.

In spite of the advances in remote sensing, its routine use for solving problems in environmental resource management in Western Europe has been somewhat limited. While there has been rapid progress in understanding the theory and practice of remote sensing of vegetation, including new theoretical frameworks for relating vegetative indices derived empirically from band ratios to primary productivity, resource managers in government, commerce and industry appear to have shown only moderate interest in routinely using remote sensing data (Kumar and Monteith 1982). This is not to decry the significant scientific achievements which have been made but rather to draw attention to the paradox of scientific achievement yet limited interest from the community of potential users - the managers and planners who are responsible for day-to-day management and planning of vegetation throughout extensive marginal tracts of Western Europe.

There are at least two possible reasons for this paradox. Firstly, the process of abstracting useful information from remote sensing data is very complicated and is still often perceived by resource managers as requiring dedicated and expensive computer systems and highly skilled operators. With the increasing power and sophistication of IBM PC-type computers and the burgeoning range of software packages for image processing, the need for dedicated computer systems is declining (Figure 1). Processing can be carried out on standard desktop computers for all but the most intensive applications (Figure 2). Resource managers are still somewhat sceptical about the benefits of remote sensing however often because of difficulties in deciding what to do with the results of image processing of remote sensing scenes - the classified image. In order to make effective use of classified remote sensing images, they must be integrated with spatial data from maps and aspatial data from field observations and reports. The process of integration must be easy to use and widely available in the offices of the resource managers.

Secondly, the way in which geographic data has traditionally been represented in computers has caused problems in integrating remote sensing and digital map data. A majority of geographic information systems (GIS) use either vector or raster models for representing spatial phenomena. The vector model is the most common method. The basic, logical unit of the vector model corresponds to a line on a map (Peuquet, 1984a). A line is represented by a descriptive heading or attribute and a list of pairs of x and y coordinates to represent location of points along the line. A point is represented by a line of zero length. Vector format is widely used in digital mapping and cartography. In contrast, the raster model represents space as a collection of square, non-overlapping cells of uniform size, called rasters or pixels, each with an associated integer value. The order of data in the digital file is determined by its location. The raster model has become much more important in recent years due to the availability of raster data from satellite remote sensing and other automated devices for data capture such as desk-top scanners and video cameras. Today both vector and raster models are widely used in GIS yet both suffer from major shortcomings (Peuquet 1984a; Maffini, 1987).

The shortcomings manifest themselves in a variety of ways. For example, neither raster nor vector data models provide a high-level view: they are both essentially low-level models. As a result, they are progressively less efficient for basic operations such as searching as data volumes increase. To take another example, most users experience great difficulties when attempting to integrate remote sensing data, typically in raster format, with archival maps, typically in vector format, and attribute data, typically in alpha-numeric format. Some users have become bedevelled by these difficulties while others have preferred to avoid them at all cost. One way of doing this has been by using GIS which rely solely on digital data in vector format.

The way of representing geographic data in computers is a crucial issue in the field of GIS. As Peuquet (1988) argued, it is inextricably linked not only with processes of data storage, manipulation and spatial analysis but also with processes of mathematical modelling of geographic data which are to be applied. It determines to a large extent the overall efficiency and performance of GIS. In future, the success of a GIS for applications in fields such as environmental resource evaluation, urban and regional planning, socio-economic research and retail market forecasting will probably be judged by the capabilities of its analytical core for mathematical modelling of various systems of interest (Figure 3).

Modelling is used increasingly in geographic research to represent our understanding of the structure and processes of complex human and environmental systems that are being studied and to predict their behaviour (Wilson and Bennett,1985). By drawing upon data from a GIS, users will be able to select and perform a range of methods of mathematical modelling, to make various initial assumptions and to study different scenarios. In this context, a GIS will be regarded as a computer-based system to extract value from geographic and other data and to aid in some decision-making process. The way of representing geographic data in such GIS will be an extremely important factor in determining their versatility, performance and efficiency.

In an effort to overcome the shortcomings of traditional data models, hierarchical data models such as quadtrees began to be studied intensively in the early 1980's. Quadtrees provide an integrated model which allows data to be interchanged with vector and raster models and provides high-level views at any selected levels of generalization. While quadtrees are more complex, they offer more elegant, versatile and efficient ways to access, browse, synthesise, analyze, abstract and view spatial data. Quadtrees show great promise for certain types of application involving regional analysis of geographic data and computer modelling of processes in both physical and human geography. While they are not without limitations, they offer a new approach to GIS which overcomes many of the restrictions of raster and vector systems and which is well worth investigating further.

In this paper, the aim is to review past work on quadtrees and their use in GIS for land resources evaluation, to outline the development of an integrated GIS using quadtrees and a relational database management system on a IBM PC-type computer and to discuss briefly its potential for inventory, management and planning of extensive, marginal tracts in Western Europe.

2.0 DIGITAL REPRESENTATION OF IMAGES AS QUADTREES

2.1 Quadtrees

Over the past 10 years hierarchical data structures such as quadtrees and octrees have been studied intensively and there is growing interest in the computer science, engineering and geographic communities in exploring their potential. The original idea underlying the quadtree stems from work by Klinger (1971). The basic motive behind the quadtree is to save storage by compaction of data having identical or similar values in graphic images (Klinger and Dyer,1976). In other words, it is to exploit the spatial homogeneity in images. While savings in storage cannot always be made, where they are there is also savings in execution time arising from compaction. Quadtrees offered other attractions such as ease and efficiency of access and manipulation for certain types of operation (Hunter and Steiglitz 1979). These factors stimulated wider interest in quadtrees in the 1980's when the concept was developed and expanded in various ways by many computer scientists(Schneier, 1981; Samet, 1981,1984; Gargantini, 1982). There is now a large family of quadtrees and octrees. Two recent books by Samet (1989a,1989b) provide a review and introduction to the history, theory and application of hierarchical data structures in diverse fields.

The potential of quadtrees for GIS has been evaluated by several workers. At the Unversity of Maryland, a major research programme into the application of hierarchical data structures to GIS produced a series of annual technical reports covering various issues and providing many algorithms for dealing with quadtrees (Rosenfeld et al. 1982, 1983; Samet 1984,1985 1986). The outcome of this research programme was the Quilt system, a GIS based on quadtrees(Shaffer, Samet & Nelson 1987).

The use of quadtrees for GIS was explored by several others. Peuquet(1984) proposed their use for a knowledge-based GIS. Mark and Lauzon (1984) proposed a scheme of two dimensional run-encoding for representing quadtrees in GIS. Gahegan and Hogg(1986) described a pilot GIS using linear quadtrees and relational database management system. Other GIS based on quadtrees have been described by Callen et al.(1986), Palimaka et al. (1986) and Abel(1988). While there are many different types of quadtrees, the essential concept of hierarchical sub-division of space forms the basis of them all.

2.2 Example of a region quadtree

The concept is perhaps best illustrated by describing the regional quadtree. It is constructed from a $2^n \times 2^n$ binary array of pixels which represent an image. It is based on successive sub-divisions of the array into quadrants. A quadrant that is uniformly black or white is represented by a leaf in the tree; otherwise it is sub-divided further into its four quadrants. Sub-division proceeds until a uniform quadrant or a specified level in the tree is reached.

A region and its binary array image of size $2^4 \times 2^4$ with tree representation for the block decomposition is shown in figure 4. After the sub-division, the image is represented by the quadtree structure.

2.3 Methods of quadtree encoding

Quadtree-encoded images may be represented in at least four different ways:

2.3.1 Using pointers

A quadtree structure that stores the tree structure explicitly is known as a pointer-based quadtree because each non-leaf node has five pointers: one to father and four to sons. Pointers were used in the initial work on quadtrees, to provide a means of aggregating areas, accessing specific locations quickly and displaying an image at varying levels of detail. A major drawback was that the pointers themselves took up a great deal of storage space.

2.3.2 Using treecodes

Oliver and Wiseman(1983) proposed the use of treecodes to represent quadtrees in which the order of tree traversal is implicit. Each node, whether non-terminal or leaf, has a value given in 5 bits. The tree is traversed in depth first order and the 5 bits or flags are interpreted. If a leaf node, the other 4 bits specify its colour. If a non-terminal node, sub-divide it.

2.3.3 Using leaf codes or linear keys

Linear keys represent spatial data as a list of locational codes for the leaf nodes only(Gargantini, 1982). Each leaf node in a quadtree is given a unique key, known as the Morton key, which contains enough information to reconstruct the size and location of the block it represents in the image (Morton 1966). Thus each block in figure 5 possesses a unique linear key which shows the level at which the node was formed and its x,y coordinate address in the image. The linear key is formed by bit-wise interleaving of the x,y coordinates of the image. In general, only the black nodes of an image are stored as linear keys. Variations on this model include associating a colour or edge information with each node and run-encoding of the quadtree.

2.3.4 Using compression and spatial index for linear keys

Gahegan (1989) proposed a much more compact and efficient method of storing and accessing data using linear keys. The method is based on a property of the Morton indexing scheme where the key of any node in a sequence is given by adding the key of the previous node to the size of the previous node

$$\text{KEY } n = \text{KEY } n\text{-}1 \ + \text{SIZE } n\text{-}1$$

Given a list of nodes in Morton sequence, the key itself need not be stored but can be inferred as long as the key of the first node in the list is known (usually zero). Hence all that is required to rebuild an image is the size and colour of a node.

Gehegan (1989) also proposed methods for high-level indexing of images to facilitate search and directly associating the edges of features with the individual nodes of a quadtree (Figure 6). Images of a size 2^{15} x 2^{15} could be accessed using only two bytes to store the colour, edge information and size of a node. The method offers significant improvements in speed and efficiency of using quadtrees.

2.4 LABELLING AND ABSTRACTION OF OBJECTS

Each region within a quadtree-encoded image may be studied individually. Its geometrical characteristics such as its size, length of perimeter, centroid, shape and orientation may be determined. Other characteristics may be computed such as the directional relationships between features(Peuquet and Ci-Xiang, 1988). Such relationships are derived attributes or aspatial characteristics of a region. They are typically derived during one operation for every region in an image.

Other attributes pertaining to a region are equally important in geographic research and include descriptive or statistical information concerning a region. For example, in an agricultural application, this might include land parcel number, ownership, soil type, crop type, fertiliser applications, yield and government subsidies which have been awarded to that parcel. To take another example, a region might equally be a postal sector or district with attributes such as population, retail sales and various performance indicators for locational market analysis. The point is that attribute data needs usually to be attached to a region.

In order to do this, each region in an image is defined as an object and relations are then used to represent each object in a relational database management system (Figure 7). An object can be regarded conceptually as an entity that has a state and a set of operations which can modify or interrogate that state. The representation of state is inaccessible from outside the object. The principle of designing a system as a set of interacting objects offers several advantages from the programming point of view. In our work, this principle has not been adhered to strictly but has been adopted wherever possible to simplify the implementation.

To define an object, the minimum and maximum x-y coordinates of a rectangle surrounding a region are automatically determined during a component labelling function (Figure 8)(Samet 1981). During the same process, the size of the region in leaf nodes or pixels is determined. Each object has the structure shown in Figure 9.

The facilities of relational database management systems can then be used to perform search operations on objects. Typical questions such as: "Where are all the parcels or fields belonging to farmer P?" or "What is at a given location?" can readily be performed. They can be performed on an original quadtree, such as a classified multi-spectral remote sensing image showing land cover types, or, more typically, they can be performed on a derived quadtree which has been formed by overlaying selected features, such as land cover with geology, soils or relief. 'Overlay', in this context, refers to the set operations of union, intersection and complement. The method has an elegance and generality which make it easy to use, efficient and extremely powerful for analyzing geographic data.

3.0 PILOT GIS

Over the past 5 years, two pilot GIS using quadtrees have been developed and evaluated in the University of Leeds:

1. In 1986, a pilot GIS using regional quadtrees and the relational database management system Ingres was developed specifically to evaluate the potential of this approach for regional evaluation of land resources (Hogg,Gahegan & Stuart, 1986; Gahegan & Hogg 1986). It was developed on a VAX 11/780 computer running UNIX. A menu provided prompts for interactive use but there were no facilities for screen graphic display of results on the multi-user university computer. The pilot GIS formed the basis however of a series of studies concerned with agriculture and hydrology, most of which included the use of remote sensing data (Hogg & Stuart 1987; Stuart & Hogg 1987; Hogg, 1988).

2. In 1988, a new pilot GIS using more efficient quadtrees was developed on a Sun 2 computer (Gahegan, 1989). This system was operated using a mouse and a highly ergonomic graphical interface to perform interactive analysis of spatial data

and display results. It demonstrated the potential of the approach for interactive analysis of large quadtrees and stimulated a range of related work (Gahegan and Roberts 1988).

As a result of these pilot studies, a new initiative was taken recently to port the software to IBM PC/AT and PS/2 computers running DOS 3.3 and OS/2 respectively. In order to take advantage of recent developments in database systems and professional software tools, two tool kits were adopted: (1) a database tool-kit; and (2) a windowing environment. The database tool-kit provides a relational database management system with a range of standard facilities for developing applications software. It allows rapid progress in implementing the GIS. The Microsoft WINDOWS / PRESENTATION MANAGER tool-kit was adopted to provide an operating environment which offers several advantages both to the software developer and software users.

4.0 MS WINDOWS / OS/2 PRESENTATION MANAGER

Over the last few years, colour displays for PC computers have improved significantly. Their size, range of colours, level of resolution and overall performance have increased, yet their costs have fallen dramatically. This trend is expected to continue.

Colour gives the user interface a new dimension which can be exploited to convey complex information such as images, maps and graphs. While there are inherent dangers in using strong colours, particularly when it is used to communicate meaning in maps, the advantages of colour displays generally outweigh their limitations.

To exploit the range of different colour displays on desk-top computers, the Microsoft WINDOWS / OS/2 PRESENTATION MANAGER was chosen (Microsoft 1989). It is an operating environment that sits on top of the DOS operating system and provides a graphics-oriented user interface, a multi-tasking capability and hardware independence for the PC family of computers. From the developers point of view, WINDOWS/PRESENTATION MANAGER offers several major advantages, including:

(i) a standard ergonomic user interface
(ii) interprocess communication
(iii) support for display of hardware independent graphics
(iv) concurrent processing of a number of applications
(v) developer-defined menus, complete if required with pop-up submenus;
(vi) support for handling selection and announcement objects
such as dialog boxes and message boxes.

From the users point of view, WINDOWS/PRESENTATION

MANAGER offers an intuitive, interactive interface typically activated using a mouse, with limited use of keyboard. Once a user learns the features of the standard layout of a WINDOWS/PRESENTATION MANAGER display, he can use virtually any of the increasing range of products using this graphical interface.

5.0 TEST SITE AND DATA

A test site has been selected in the Peak District National Park which lies astride the South Yorkshire/ North Derbyshire border in the Pennines (Figure 10). It covers an area of steep moorland and dales surrounding the Ladybower and Derwent Reservoirs. The moorlands rise to over 500 m and are predominantly covered by heather, eroding peat and acid grasslands though there are several coniferous forests on the lower margins. The dales are generally about 200 m above sea level, with extensive areas of improved pasture. Much of the area forms part of water catchments. The Peak District is a popular area for recreation, its widely varying scenery being enjoyed by many visitors.

The main features of the solid geology, soils, rainfall and drainage network were digitized from various maps published by the Ordnance Survey and encoded as quadtrees.

A Landsat Thematic Mapper scene acquired over the North of England on 26 April 1984 was selected because it was free from clouds and at this time of year differences in upland vegetation manifest themselves well. Part of the scene covering the Peak District National Park was resampled to the National Grid of the Ordnance Survey and a square image of 1024 x 1024 pixels with a ground resolution of 25 m was produced. Spectral bands 3, 4, 5, 7 and 4/3 ratio were used to produce eleven categories of land cover by maximum likelihood classification. Subsequently, some of the classes were combined for this study to show only major types of vegetation. A subset of 512 x 512 pixels covering the area around Ladybower and Derwent Reservoirs has been used for illustrations in this paper.

6.0 OPERATIONS

There are at least five phases in evaluation and management of human and physical resources with any region. They are: (1) inventory; (2) monitoring; (3) analysis; (4) modelling and prediction (5) decision-making.

(1) Inventory is concerned with the collection, recording and classification of data about a specific subject of interest such as agricultural crops, farm animals or population.

(2) Monitoring is concerned with detection and recording of change over a time interval.

(3) Analysis is concerned with studying the capabilities and constraints imposed by the physical,

economic and social environment on variables such as crop growth, industrial investment and employment and the classification of these.

(4) Modelling is concerned with describing of some system of interest or object of study, identifying its components and the ways in which they interact and representing them statistically or mathemetically. Models may be concerned with physical relationships between a crop and its environment, such as a crop growth and yield model, or with economic and social aspects of the rural economy, such as changes in population and employment. The goal is usually to achieve a working model which can form the basis of predictions about key issues affecting the management of particular resources in various ways. Such a model will allow users to pose 'What if?' questions using different values for variables in a models such as different growth rates or crop prices. It should allow the effects of different assumptions on the outcome to be viewed and constraints on variables to be introduced to produced a desired outcome.

(5) The final phase concerns formulation of a policy and its implementation as a management plan for the physical or human resources being studied.

6.1 Spatial frameworks

Many modelling methods have been developed on the basis of an exclusive and exhaustive zoning system on a map.

The zones provide a fixed, spatial framework. They typically correspond to census, postal or other administrative regions for which there is census or other suitable data for modelling. The zones may vary in size and shape but usually have relatively simple boundaries.

The alternative is to have a spatial framework which corresponds to features on a map, such as rock formations, soil or vegetative types. In this situation, the boundaries of zones or regions may be highly convoluted and there is usually difficulty in getting data pertaining to such regions. For example, classifications of vegetation derived from multi-spectral remote sensing produce regions with relatively uniform spectral characteristics but often with wide variations in size and shape. Where such regions relate to large fields of crops, derived vegetation indices and photosynthetic absorbed radiation may be obtained and used with known data about that crop to model its growth and estimate its yield. Where such regions relate to natural vegetation on the other hand there is greater difficulty in obtaining collatoral data which is necessary for modelling.

The role of IBM PC GIS in the above task falls broadly into three categories:

(1) Management information systems.

The information collected during the inventory and monitoring phases forms the basis of a management information system. This provides instant information about the nature, extent and condition of resources within a given study area and gives answers to basic questions from resource managers such as rainfall during the previous week at a given site, the type of crop growing in a particular field or all land belonging to one owner. Essentially the information system provides access to the data that was collected and fed into the GIS. In the example drawn from the Peak District, a classified Landsat TM image shows different types of vegetation cover and is integrated with soils, geology, climate and topography to provide the basis for a management information system (Plate 1). By adding boundaries of water catchments, the vegetation cover within a catchment can be disaggretated and used for modelling rates of water run-off through the catchment.

(2) Manipulation and analysis

The information collected during the inventory and monitoring phases may be analyzed to produce new information about spatial and aspatial relationships in the original data. Simple analysis provides answers to distance or neighest neighbour queries such as: "Where is the nearest source of water for cattle in a particular field?" More complex queries requiring overlay or set operations might include: "Find all locations where there are shale rocks, steep slopes and heavy rainfall?" This might identify for example all locations where there is a need for contour ploughing to avoid excessive soil erosion. Similarly, queries of aspatial data in the relational database produce new information. Thus "Find all fields of winter wheat belonging to a specified farmer on loam soils?" To answer this question, a search of all fields or objects belonging to the farmer might be made first and a new image would be created to show those which contained wheat. This image would then be intersected with with loam soils. Thus, by performing various types of analysis on the original data, new or derived data may be created in the GIS.

(3) Modelling and Decision-making

Environmental systems are generally complicated, with many interacting processes which are rarely fully understood. Even where there is a detailed understanding, there are usually difficulties in obtaining all the relevant data and so one has to work with whatever data are available. The methods that are increasingly being applied to handle this complexity are systems analysis and mathematical modelling. Wilson and Bennett(1985) argue that the conceptual basis of much geographical modelling can be considered broadly under three headings, namely, accounting, optimization and dynamics. While the mathematical analysis underlying these approaches becomes increasingly complex the deeper one delves, involving the most advanced mathematics in dynamical systems theory, a great deal can be achieved in practice using elementary mathematics on a IBM PC computer.

To illustrate this, consider an accounts model using a fixed framework of zones (census districts, postal sectors, etc). The variables for this can readly be obtained from within a IBM PC GIS which uses quadtrees and a relational database. This is because quadtrees permit object abstraction. For each zone, its spatial variables can be derived including area of each zone, land use by sectors, centroid and nearest neighbour. Its aspatial variables would have to be read into the relational database management system. They might include census, employment and other socio-economic data. The model accounts for variables pertaining to each zone at the beginning and end of a time interval. Such models may be applied to economic, population, social, ecological or hydrological accounts. They are basically concerned with flows of goods, people, nutients or water from one zone to another within a time interval. Represented as a series of matrices, such accounting models form one way forward for IBM PC GIS (Plate 2).

6.0 DISCUSSION

If we are to exploit the new capabilities for collection and processing of geographic information, then we must focus attention more and more on developing improved strategies for its use and management in GIS. As Dale and McLaughlin (1988) argue, such strategies are concerned with the development and organization of resources to achieve sets of objectives. These objectives show a general tendency to become more ambitious and complex and to span an increasingly broad range of applications. They include the need to develop a range of methods of mathematical modelling to help managers and planners to decide what course of action they should take.

There are limits to the type of modelling that can be achieved on IBM PC computers. Computationally intensive types of modelling will no doubt continue to be carried out on mainframe computers. Greater efforts are now being made however to include capabilities for modelling and decision-making on IBM PC computers. This is being done either by exploiting new means of communication between mainframe and IBM PC or PS/2 computers or by transferring all or parts of modelling software to personal computers. At the same time, there is growing interest in the development and use of expert systems for GIS (Ripple and Ulshoefer, 1987).

The new GIS is capable of representing an area of 32768 x 32768 pixels on a 16-bit IBM PC/AT computer. A window into this image is mapped to the dimensions on the colour display screen which is being used. It will initially provide an efficient spatial and aspatial information system for inventory of land resources, market research and urban and regional planning. As the system is developed it will be tailored to specific applications involving regional analysis and modelling.

A critical issue in setting up a GIS is the choice of size for leaf nodes. The leaf node represents the smallest area which can be handled by the system. It must be chosen carefully by a user to dove-tail into the range of applications for which the system will be used. By choosing a small pixel size, say 2 m x 2 m on the ground, then the size of the maximum area covered is specified. The test data has a pixel size of 30 m x 30 m, corresponding to the level of resolution available from Landsat Thematic Mapper data. There is of course no reason why data with different levels of spatial resolution cannot be used together providing they are chosen with care.

A feature of quadtrees is that they cover square blocks. If the area to be studied fits within one block at the level of resolution which is needed, then only one quadtree database is required. If the area to be studied extends beyond one block, then separate quadtree databases will have to be built, each one adjacent to the other and forming a grid. Square blocks can be formed to provide coverage of countries, continents or even the entire world. Mark and Lauzon (1985) discussed one approach for building GIS using quadtrees at continental or global scales while Tobler and Chen (1986) proposed another using authalic coordinates.

There are several major issues which remain to be addressed such as the questions of accuracy and uncertainty in GIS. There will no doubt be progress in specifying and coping with these but there are major problems to be overcome. For example, the intersection of two regions, such as heavy rainfall and sandstone rocks, may produce a new image but how significant accurate is it given the differing levels of uncertainty attached to rainfall and rock boundaries. Some means of dealing with issues like these need to be developed.

7.0 CONCLUSIONS

There is increasing recognition that geographic information has value and needs to be carefully managed to maximize its potential benefits. These benefits can only be made available to users if we devise new efficient strategies for handling geographic information and for using it to solve problems in various fields. Such strategies must include efficient methods for storing, searching, browsing, windowing, editing, scanning and displaying spatial data. They must also include efficient methods for abstracting objects and performing various types of mathematical modelling for applications in both in human and physical geography. Even with relatively simple mathematical models, a great deal of valuable information can be derived from synthesising geographic information. This information must be presented to users in a form which helps them to decide what the best course of action might be to achieve a given set of objectives.

The use of hierarchical data structures like quadtrees provide a neat way on integrating vector and raster data within an integrated GIS. It allows more effective use of remote sensing imagery and map data than many of the current generation of desktop GIS. By providing facilities for aspatial or attribute data within the relational database, the new GIS offers a way of combining geographic data of different types and from diverse sources into a unified GIS.

The strategy behind one approach using an IBM personal computer, the Microsoft WINDOWS / PRESENTATION MANAGER environment and a highly efficient quadtree encoding and integral database has been outlined. While work on the development and evaluation of this approach continues, results show that it provides an integrated GIS which will allow easy integration of raster and vector data thus allowing remote sensing data to be combined with digital map data. Moreover, attribute or aspatial data can be held in a relational database management system which is integrally linked to spatial data encoded as quadtrees. The integrated GIS on a IBM PC computer will be suitable for use on a range of computers using the Intel family of processors, will provide multi-tasking capabilities and will have a intuitive, ergonomic interface for applications in a broad range of fields. Current plans are to extend the existing project to provide modelling and executive decision-support facilities within an advanced product running under OS/2.

8.0 REFERENCES

Abel, D. 1988 **SIRO-DBMS User's Manual.** TR-HA-88-1, CSIRO Division of Information Technology, Canberra, 156pp.

Callen M., James,I. Mason,D.C. & Quarmby N. 1986 A test-bed for experiments on hierarchical data models in integrated geographic information systems. In:Diaz B M & Bell S B M (ed) **Spatial Data Processing using Tesseral Methods**, NERC, Swindon, UK. pp193-212.

Dale, P.F. & McLaughlin J.D. 1988. **Land information management.** Clarendon Press, Oxford. pp266.

Gahegan, M. & Hogg, J. 1986 A pilot geographical information system based on linear quadtrees and a relational database for regional analysis. In: Diaz,B.M. & Bell,S.B.M. **Spatial Data Processing using Tesseral Methods**, NERC, Swindon,UK.pp213-231.

Gahegan,M.N. & Roberts, S.A. 1988 An intelligent, object-oriented Geographical information system. **Int. J. Geographical Information Systems**, Vol.2. No.

Gahegan, M. 1989. The efficient use of quadtrees in a geographic information system. **International Journal of Geographic Information Systems**. Vol. 3, No. 3, July-Sept, pp 201-214.

Gargantini, I. 1982 An effective way to represent quadtrees. **Communications of ACM** 25(12):905-910.

Hogg, J. 1988. Modelling Land Resources Within a Pilot Geographical Information System. **Proc. IGARSS '88 Symposium**, Edinburgh, Scotland, 13-16 Sept. pp 101-105.

Hogg, J., Gahegan, M. & Stuart, N. 1986 Evaluation of regional land resources using geographic information systems based on linear quadtrees. **Proc. of the Seventh Int. Symposium on Remote Sensing for Resources Development and Environmental Management**, ISPRS Comm VII, Enschede,25-29 August,pp917-925.

Hogg,J. & Stuart,N. 1987 Resource analysis using remote sensing and an object-oriented geographical information system. **Proc. of the 13th Annual Conference of the Remote Sensing Society**, Nottingham, England. p79-93.

Hunter G M & Steiglitz K 1979 Linear transformation of pictures represented by quadtrees. **Computer Graphics and Image Processing** 10, p286-289.

Klinger, A. 1971 Patterns and search in Statistics. In: **Optimizing Methods in Statistics.** Ed. J.S. Rustagi, Academic Press, New York. pp 303-337.

Klinger A & Dyer C R 1976 Experiments on picture representation using regular decomposition. **Computer Graphics and Image Processing** 5: 68-105.

Kumar, M. & Monteith, J.L. 1982 Remote Sensing of Crop Growth. In: "**Plants in the daylight spectrum**". Ed. H. Smith, Academic Press.

Maffini, G. 1987. Raster versus Vector Data Encoding and Handling: a Commentary. **Photogrammetric Engineering and Remote Sensing.** Vol. 53, No. 10, October, pp1397-1398.

Morton, G.M. 1966 "A Computer Oriented Geodetic Data Base: a new Technique for File Sequencing" **Unpublished Report**, IBM Canada.

Oliver M A & Wiseman N E 1983 Operations on quadtree encoded images. **The Computer Journal** Vol 26, No 1,p83-91

Palimaka,J. Halustchak,O & Walker, W. 1986 Integration of a spatial and relational database within a geographical information system. **Technical Papers, ACSM-ASPRS Annual Convention**, Vol. 3, Washington, DC pp

131-140.

Peuquet, D.J. 1984a A conceptual framework and comparison of spatial data models. **Cartographica** 21, 66-113.

Peuquet, D.J. 1984b Data structures for a knowledge-based geographic information system. **Proceedings of the International Symposium on Spatial Data Handling,** Zurich. Vol 11, p372-391

Peuquet D J 1987 An algorithm to determine the directional relationship between arbitrarily-shaped polygons on the plane. **Pattern Recognition** Vol 20 pp 65-74.

Peuquet, D.J. 1988. Representations of Geographic Space: Toward a Conceptual Synthesis. **Annals of the Association of Amercian Geographers,** 78(3):375-394.

Ripple, W.J. & Ulshoefer,V.S. 1987. Expert systems and Spatial Data Models for Efficient Geographic Data Handling. **Photogrammetric Engineering and Remote Sensing.** Vol. 53, No. 10, October, pp1431-1433.

Rosenfeld, A. Samet, H. Shaffer, C.A. & Webber, R.E. 1982 **Application of hierarchical structures to geographical information systems.** University of Maryland TR-1197.

Rosenfeld, A. Samet, H. Shaffer, C.A. & Webber, R.E. 1983 **Application of hierarchical structures to geographical information systems,phase II.** University of Maryland TR-1327.

Samet, H. 1931 Connected component labelling using quadtrees. **Journal of the ACM** 28(3):487-501.

Samet, H. 1984. The quadtree and related hierarchical data structures. **ACM Computing Surveys** 16,187-260.

Samet, H. 1989. **Design and Analysis of Spatial Data Structures.** Addison-Wesley, New York.

Samet, H. 1989. **Applications of Spatial Data Structures : Computer Graphics, Image Processing and G.I.S..** Addison-Wesley, New York.

Samet, H., Rosenfeld, A. Shaffer, C.A. Nelson, R C & Y-G Huang 1984 **Application of hierarchical structures to geographical information systems,phase III.** University of Maryland TR-1457.

Samet, H., Rosenfeld, A. Shaffer, C.A. Nelson, R C, Y-G Huang & Fujimura, K. 1985 **Application of hierarchical structures to geographical information systems, phase IV.** University of Maryland TR-1578.

Schneier, M. 1981 Two hierarchical linear feature representations:edge pyramids and edge quadtrees. **Computer Graphics and Image Processing** 17(3):211-224

Simonett, D.S. 1988. Considerations on integrating remote sensing and geographic information systems. In: Mounsey, H. & Tomlinson, R. **Building Databases for Global Science.** Taylor and Francis, London, pp105-128.

Stuart, N. & Hogg,J. 1987. Quadtree Spatial Spectrum Analysis for Landsat TM Scenes Stored in a Pilot GIS. **Proc. of the 13th Annual Conference of the Remote Sensing Society**, Nottingham, England. p618-627.

Wilson, A.G. & Bennett, R.J. 1985. **Mathematical methods in human geography and planning.** John Wiley, Chichester. p 411.

This paper was presented at a meeting held at the Joint Research Centre of the Commission of the European Communities, Institute for Remote Sensing Applications, ISPRA, 5-7 September 1989, organised under the auspices of the 'Land use monitoring less favoured areas' programme.

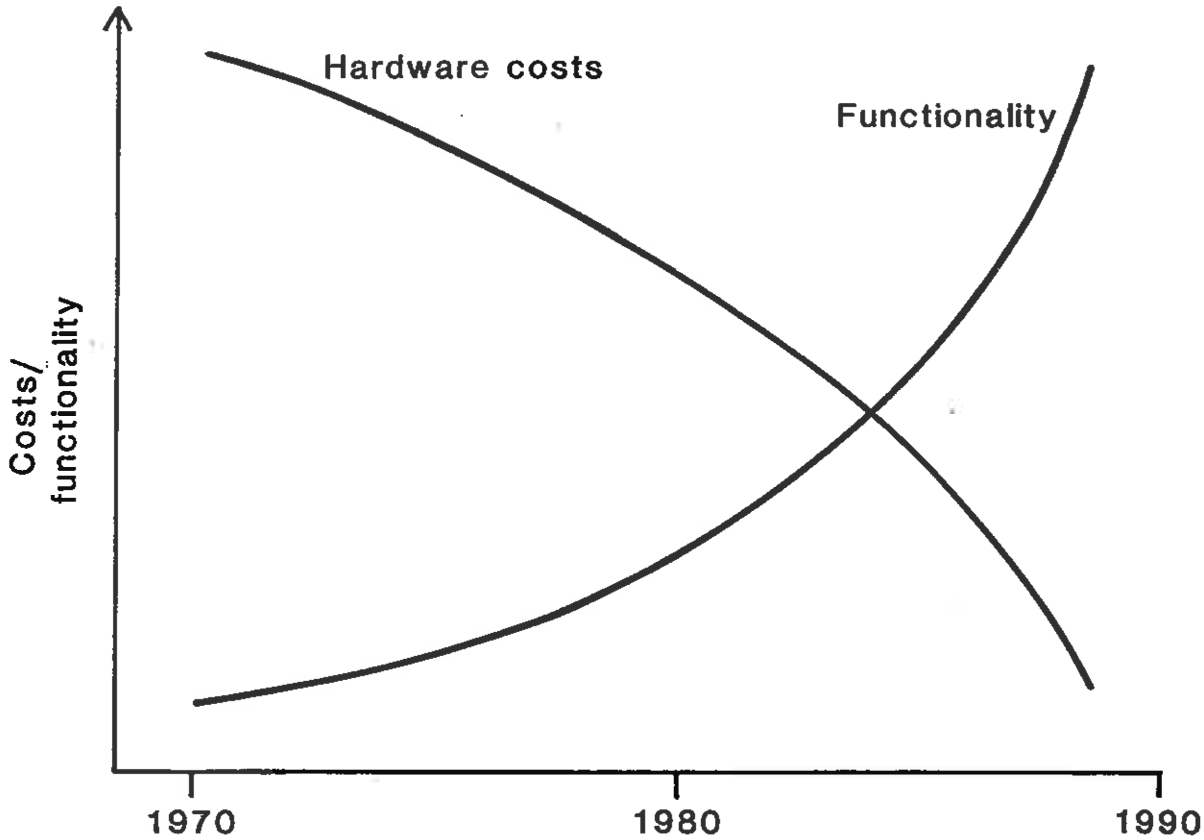


Fig. 1 : Functionality of GIS

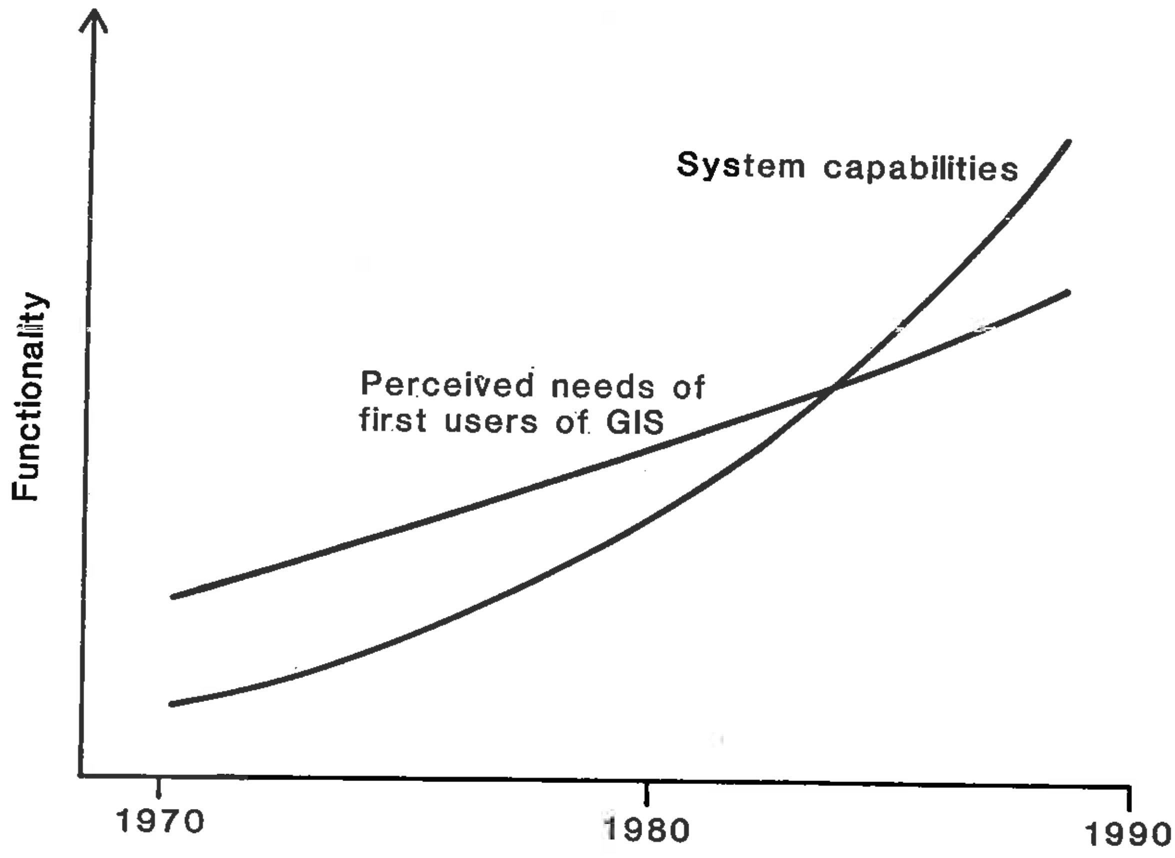


Fig. 2 : Perceived needs of first generation users of GIS

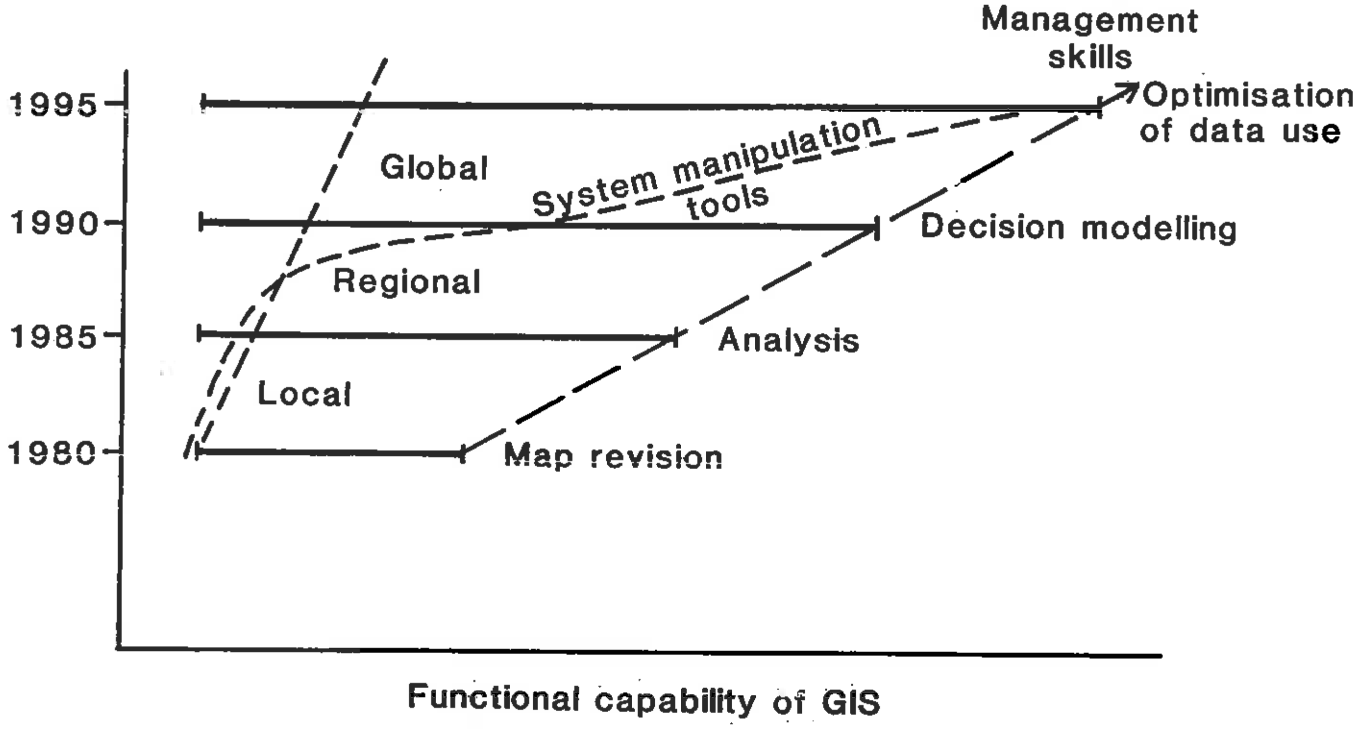
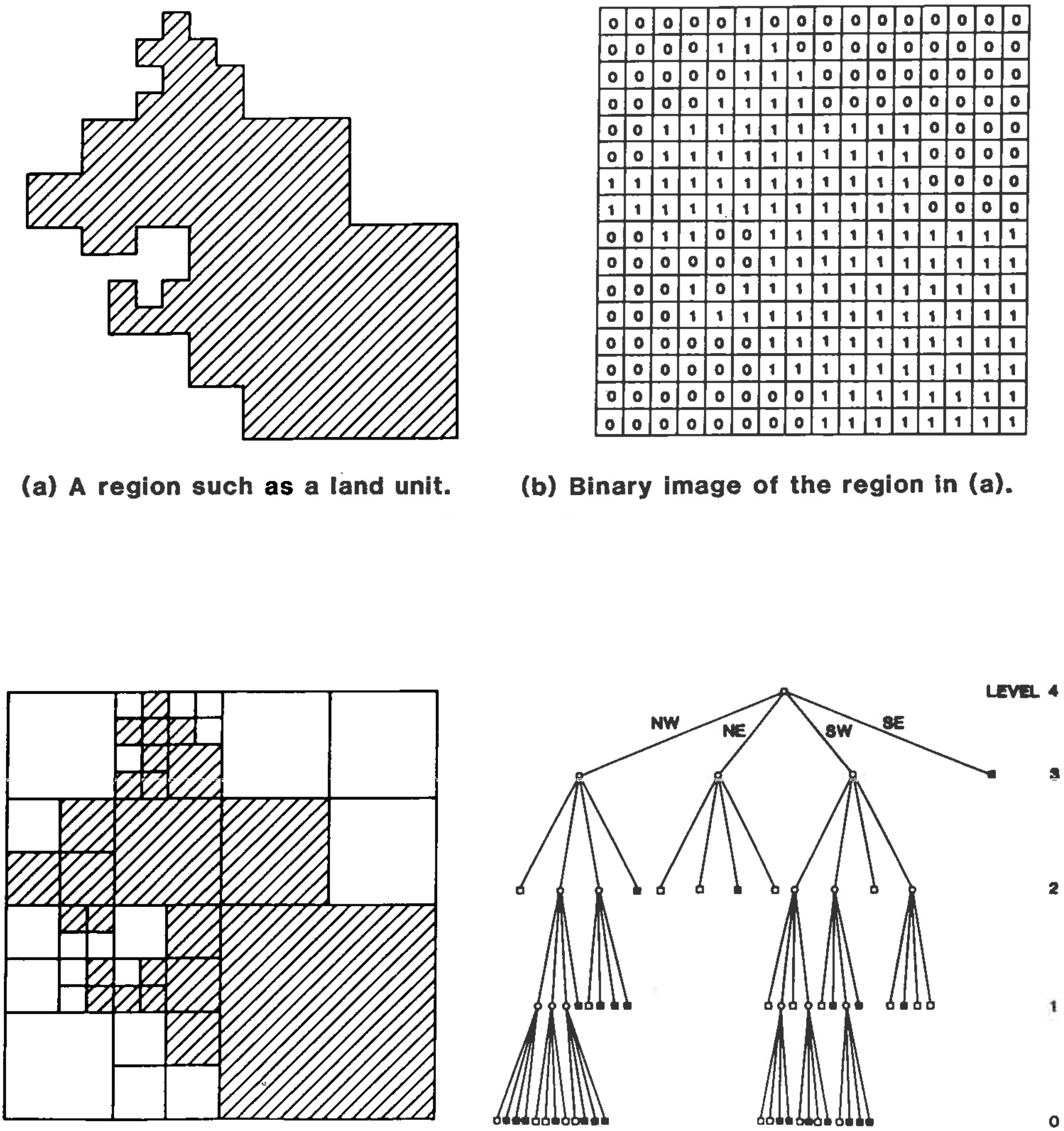


Fig. 3 : Functional capability of GIS



(a) A region such as a land unit.

(b) Binary image of the region in (a).

(c) Block decomposition of region in (a).

(d) Quadtree representation of region in (a) showing black and white leaf nodes at each level.

Fig. 4 : Principle of regional quadtrees

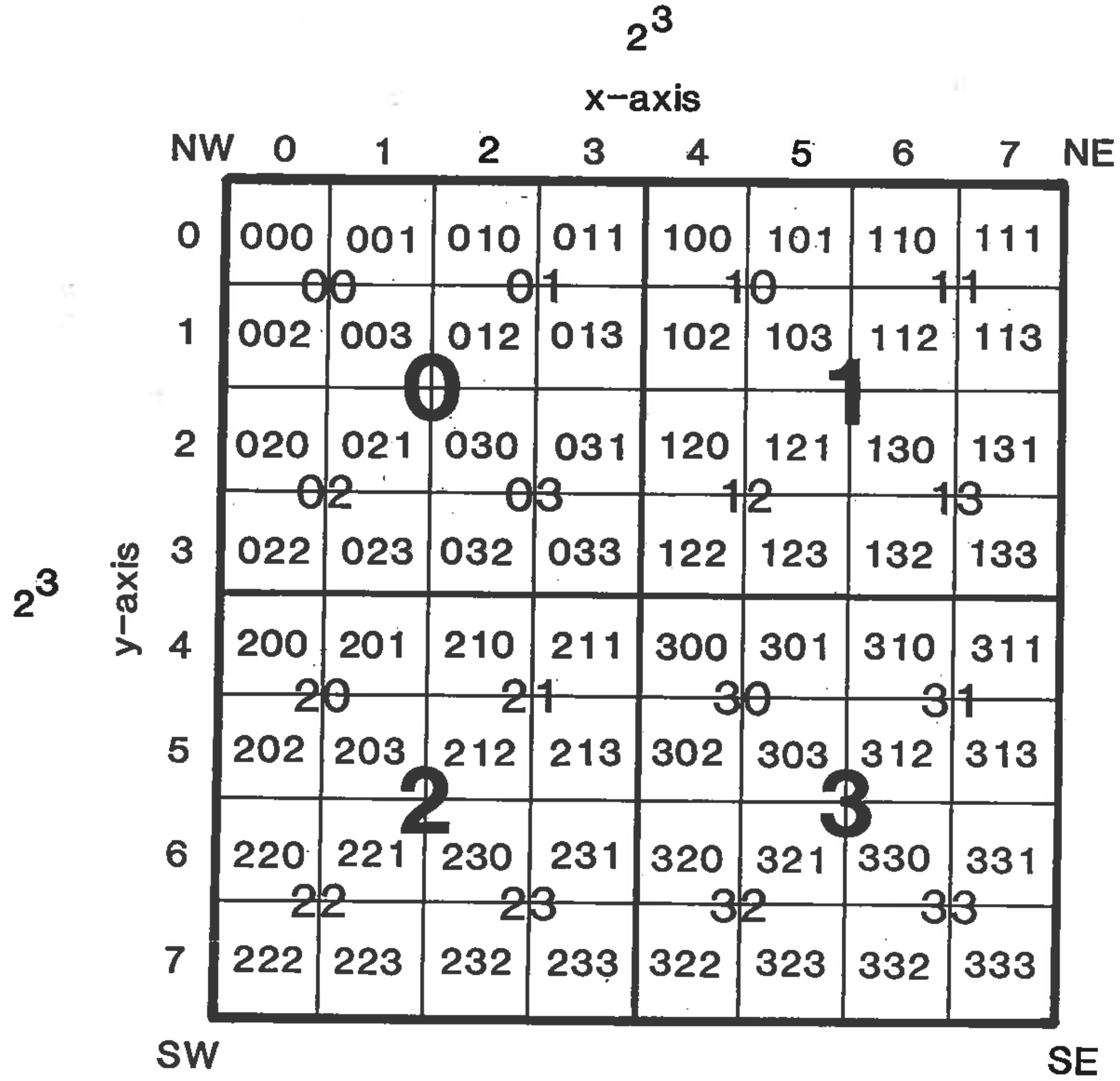


Fig. 5 : Spatial addressing using Morton scheme of bit-interleaving

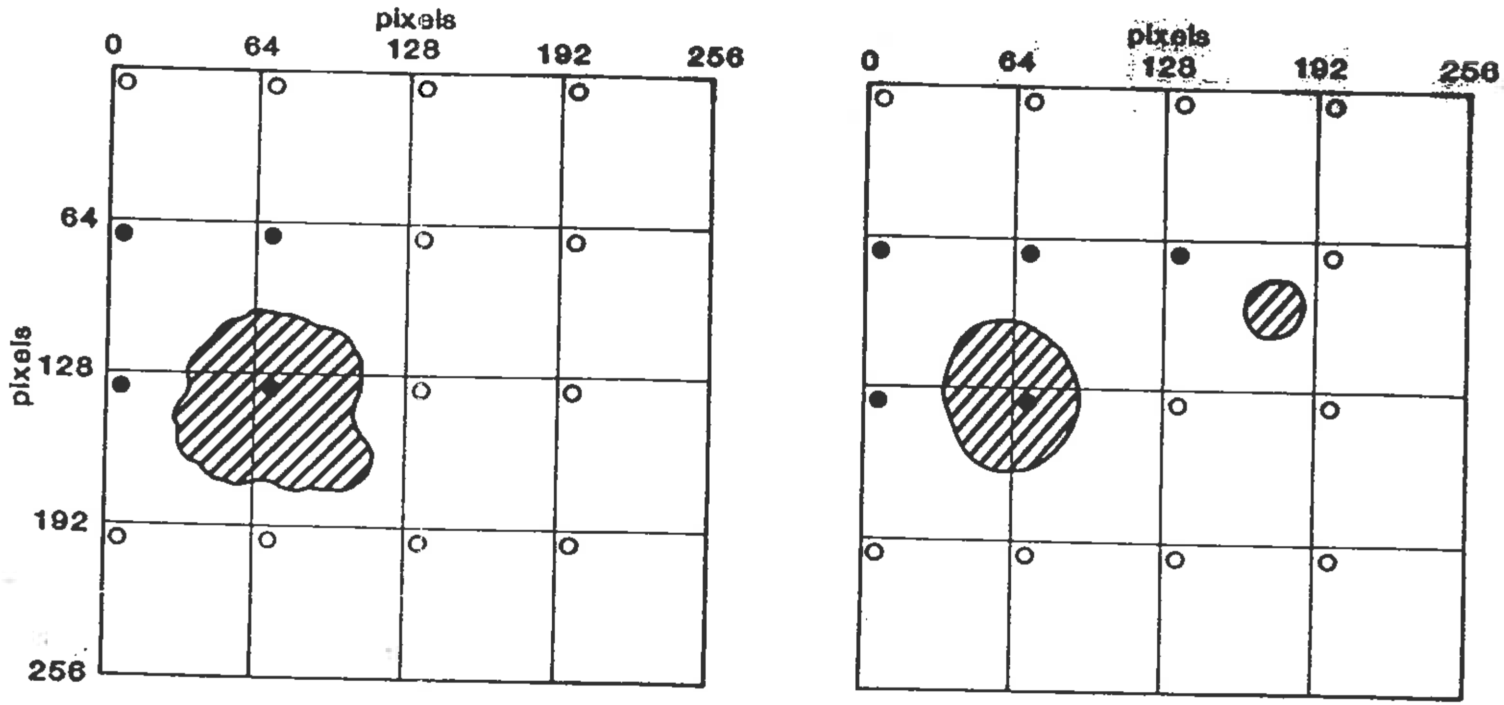


Fig. 6 :

High-level indexing of spatial images using optimised quadtrees (Gahegan, 1989)

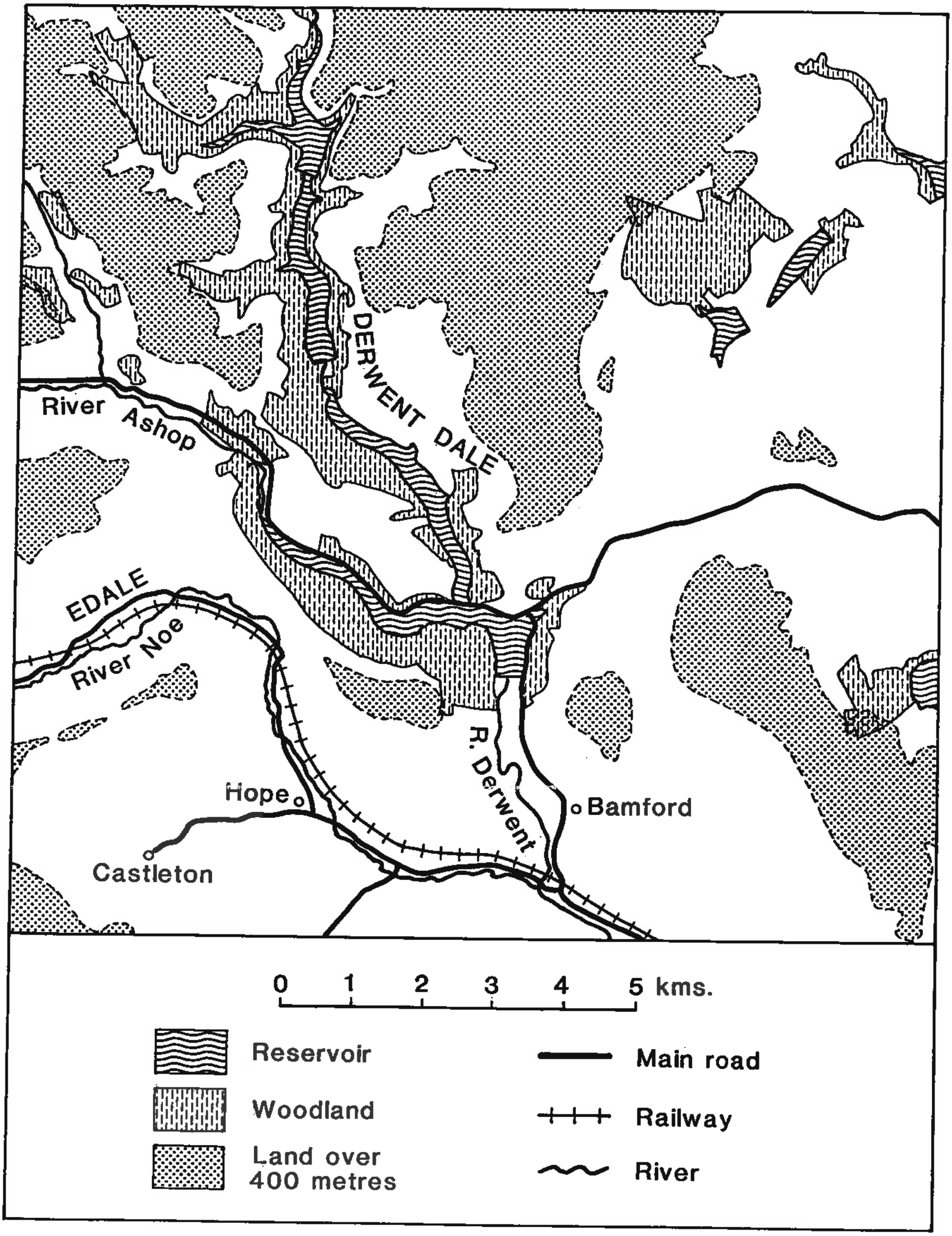


Fig.7 : **Topographic features of the test area**

Elevation 300 m

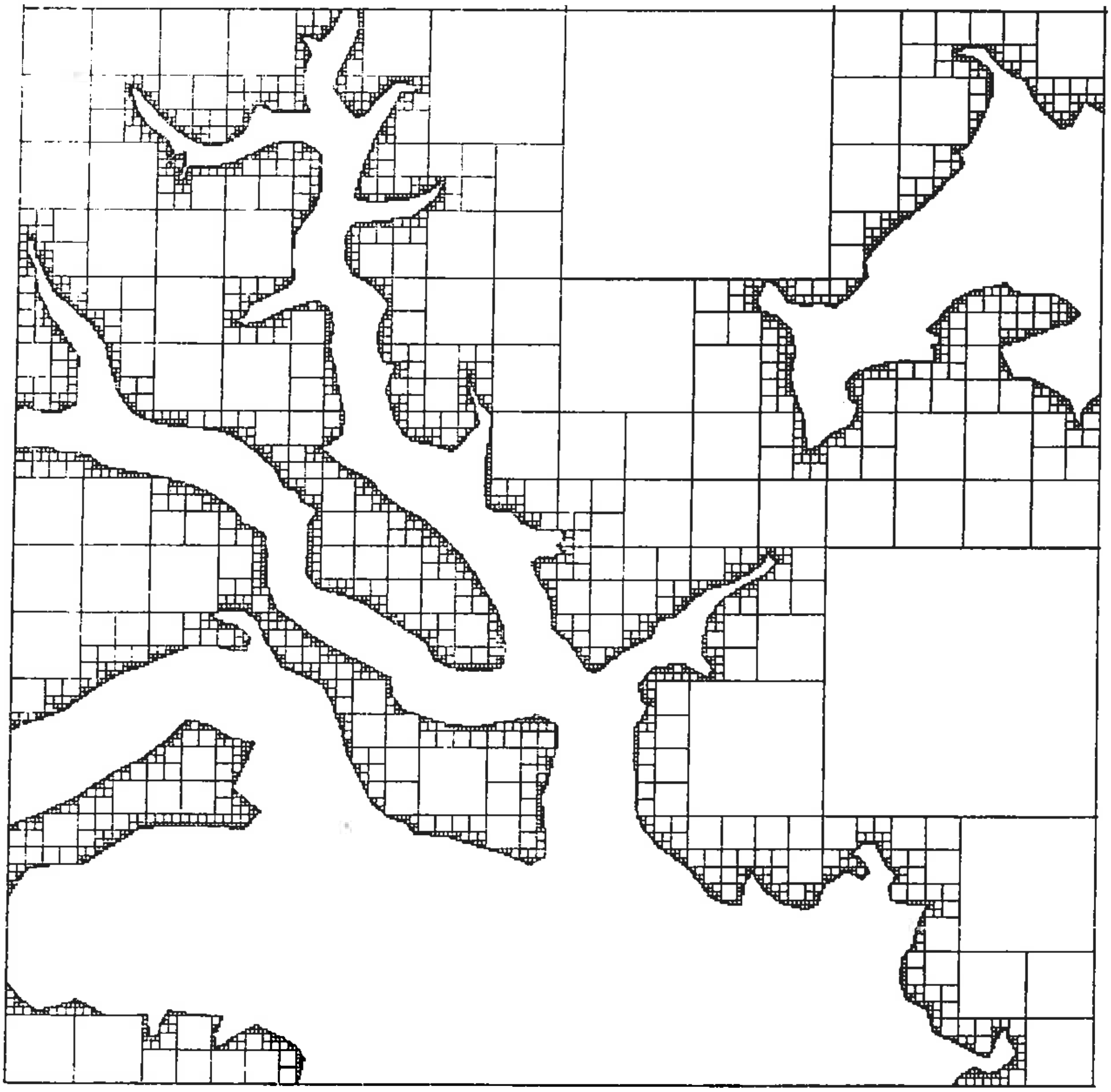


Fig. 8 : Quadtree representation of 300 m contour

Elevation 300 m

Component labelling

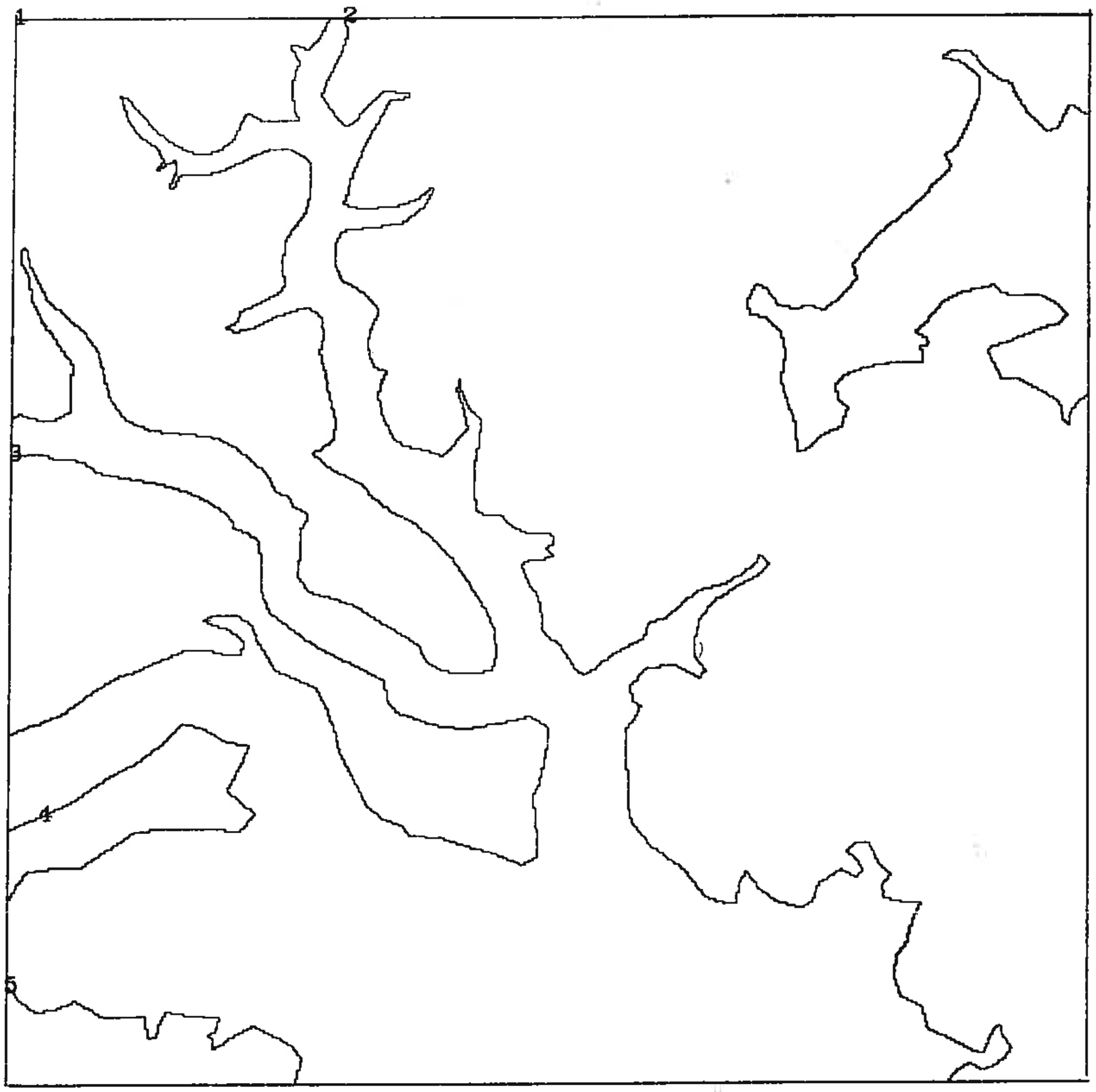


Fig. 9 : Component labelling of 300 m contour

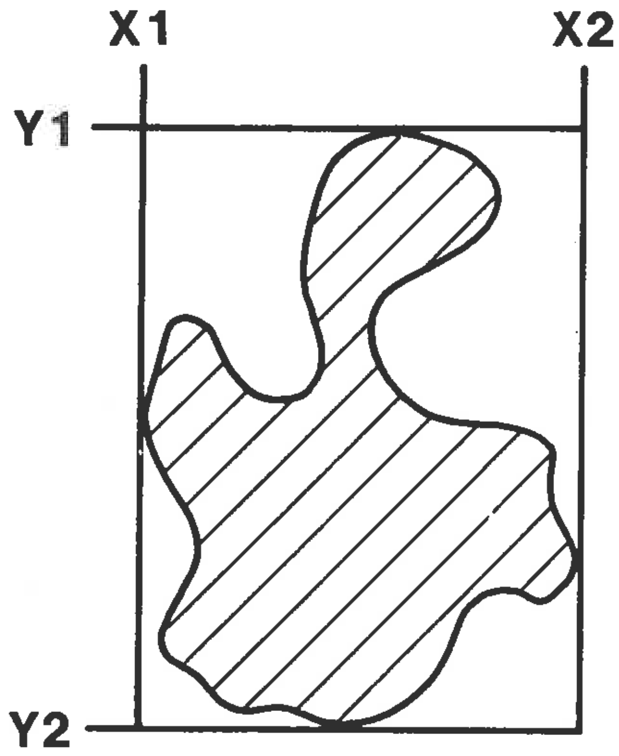


Figure 10. The minimum and maximum X and Y coordinates to define a rectangle that neatly surrounds an object

height300 relation

label	x1	x2	y1	y2	size
1	0	232	0	315	30984
2	146	512	0	512	115657
3	0	257	209	407	17693
4	0	117	339	423	4055
5	0	140	465	512	4301

Fig. 11 : Height 300 relation showing each object represented as a record in a relational database