VECTOR DATA HANDLING WITH A GIS RASTER-BASED

Jean-Paul Donnay, Benedicte Collette and Marc Binard
University of Liege, SURFACES, Dept. of Geography
7 place du XX aout, 4000 Liege, Belgium

ABSTRACT

The growing use of remote sensed imagery and D.E.M., for instance, in land planning and environmental studies is in favour of G.I.S. raster-based. Nevertheless, the availability and the selectivity of vector formated data force their joint processing. This communication deals with techniques allowing simultaneous handling of vector information and raster imagery, according to raster, vector, or hybrid algorithms. The vector to raster technology used by the current display devices makes the implementation of such procedures easier. Examples of different classes of vector representation are used jointly with remote sensed images to illustrate the items under discussion.

INTRODUCTION

A growing part of digitized spatial information is today available in raster format. So it is with airborne or satellite remote sensed data and also with scanner digitized map. Moreover, the handling of spatially continuous phenomena by a vector grid format is close to a raster structure. Most of the computer systems capable to process raster information by means of specific algorithms are relevant to image-processing. Traditionally, they support some simple vector functions (interactive delineation of training areas or region of interest) but they incorporate no or few typical G.I.S. operations. A vector layer which must be processed in such a system has to be rasterized first.

At the contrary, nearly all geographic information systems which are operational today use vector data structure as well as vector-based algorithms. Some updated products allow vector information overlaying on a back layer constituted by a raster image, with intention of visual interpretation, although a complete compositing of raster and vector data is rarely permitted. A scanned map which must be integrated in a vector-based G.I.S. needs to be vectorized and cleaned-up.

Conversions (rasterization and vectorization) still needed to allow marrying both structures often produce poor quality outome. Each structure possesses its own advantages which must be turned to the best account according to the treatment to be processed. A composite system combining raster and vector data in the same environment, i.e. permitting data input in both formats and executing raster or vector-based algorithms, is able to offer a noticeable enhancement of performance facing common vector-based G.I.S.

From the hardware point of view, improved price/performance made the installation of such an hybrid system feasible (emergence of large amounts of main memory, reduced price of storage media, etc.) (Fallon 1990). Video technology with which CRT are now conceived gives free matricial or vectorial access to graphic memory of computer stations (vector to raster technology) (Schneider 1985). From the thematic point of view, the spatial resolution of second generation satellites allows the use of remote sensed imagery at larger scale than before and geometric corrections (warping and registration) are now efficient enough to permit incorporating of raster images in an hybrid G.I.S.

RASTER-BASED G.I.S.

The G.I.S. facilities offered by combination of both data formats and use of raster and vector-based algorithms will be presented by a few examples. A specific software has been implemented on an IBM PS/2 station (Donnay 1989). The core of the software is dedicated to image-processing and the associated G.I.S. library is raster-based. But the properties of the ancillary data in vector format are preserved to assist the spatial selection and identification of geographical entities. Rasterization is only completed when a raster-based process seems to be more efficient.

When G.I.S. constitutive data layers are available in raster format, thematic selection and intersection operations by overlay are straightforward and can be processed by strictly raster-based algorithm.

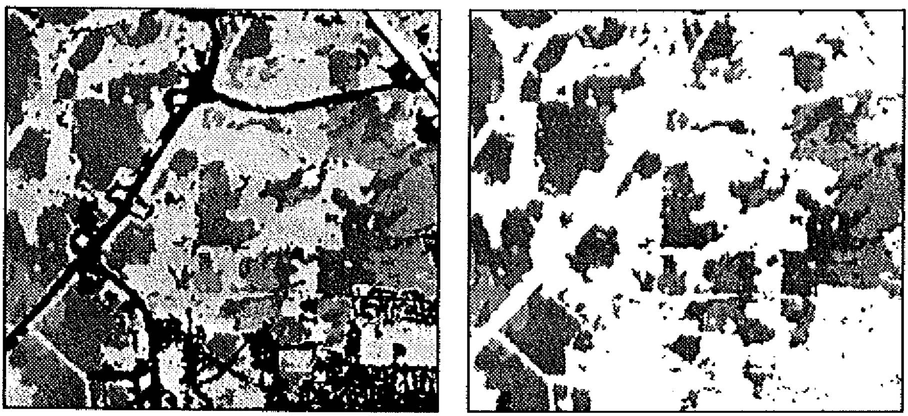
Selection by substitution

Let for example a layer of land uses coming from a supervised classification of multispectral remote sensed data. A thematic selection applied to this image can be regarded as the identification of all the pixels which show a D.N. matching the selected land use criteria. So the selection of the ''forest'' theme has to lead to the constitution of an image where the classified ''hardwood'' and ''coniferous'' pixels are only present. The resulting image is obtained by simple substitution: every pixel not matching with the choosen codes is affected to a D.N. equals to 0. The treatment can be made by the use of a Look Up Table where all values are set to 0, except the ones corresponding to the selected D.N. The following instruction, included in a couple of loops on lines and samples, allows the operation to be processed in only one pass:

```
PutPixel(i,j,LUT[GetPixel(i,j)])
```

where i = Line index & j = Sample index

Figure 1. Thematic Selection = Substitution



Binary masking and boolean operators

Obtaining intersection polygons by overlaying two layers is a typical, although basic level, G.I.S. operation. Every layer can be submitted to a selection, and the procedure can be generalized to more than two sets of map data. When the operation concerns raster format data, it is made really easier and goes through binary mask building.

Two layers and one selection

Let two layers in raster format: the land covers already used and an image of slope aspects coming from a convolution filter over a digital elevation model in raster format. The process consists in pointing out the land uses covering the S-E, S and S-W aspects only. The selection of choosen aspects is done by substitution as early mentionned: codes which do not match the criteria are replaced by 0 and, now, the selected codes are set to 1. The resulting image is consequently a binary mask which multiplicated pixel by pixel by the land uses image gives the expected map.

Two layers and two selections

Keeping the same selection of aspects as in the previous example, the resulting binary mask is conserved. A selection criterion is also applied on the land uses image, pointing out the only ''forest'' pixels. This new selection is carried out by a similar substitution: ''hardwood'' and ''coniferous'' pixels are set to 1, the others being set to 0. So the two selection procedures lead to a couple of binary masks. The final image is produced by the logical AND, operated pixel by pixel on the two binary masks. This layer is also a binary map but if needed, it is easy to recover the constitutive land uses of the ''forest'', i.e. ''hardwood'' and ''coniferous'', by a multiplication, pixel by pixel, of the final binary mask by the original land uses image.

More than two layers

Intersections-overlays with more than two layers follow the same procedure. Every selection applied to one layer leads to a binary mask, while the combination of masks can be considered as boolean operations (AND, OR, XOR, NAND, NOR, NXOR, NOT) between couple of masks. Processing is very fast, especially when compared with analogous procedures running in vector format. Besides, binary images at intermediate or final levels are relatively low memory consuming (bit-mapped images) and are suited for compactness (eg. Run Length Coding).

No selection on the layers

When the overlay procedure goes through no selection on the layers, the final image has to give back all the possible intersected polygons. Here too, the raster structure offers an easier solution. Take back the example dealing with two images: land uses and aspects. The first consisting of 20 uses and the second of 9 categories. A two entries L.U.T. is created determining 20 x 9 intersection possibilities, i.e. 180 distinct codes. Simultaneous reading of both original images provides, pixel by pixel, the two entry indices in the L.U.T. The code value contained in the L.U.T. at the crossing of both indiced line and sample, is affected to the corresponding pixel into the final image.

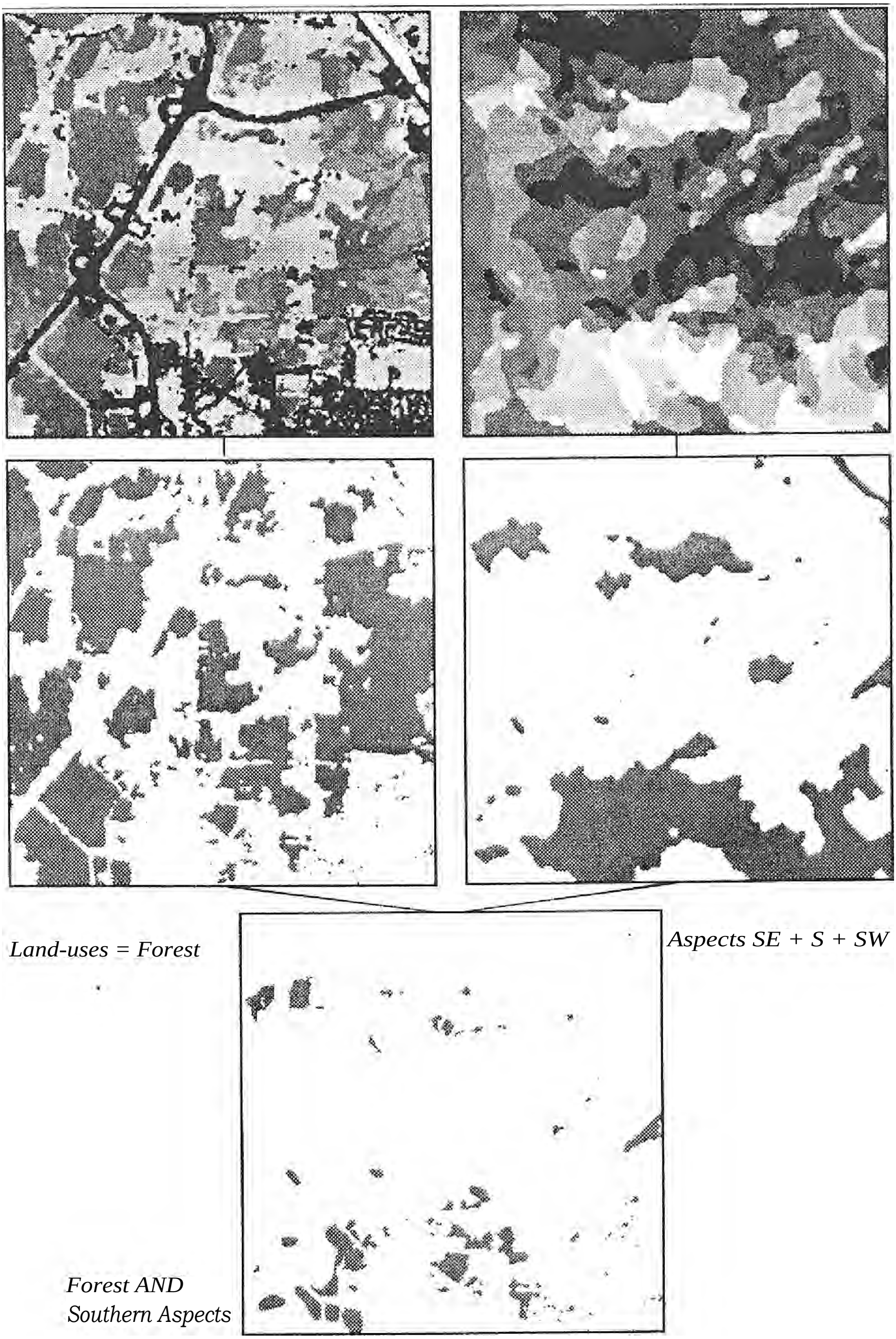


Figure 2. Two loyers and two selections

The different possibilities given by layers overlay, here restricted to a couple to allow clearer showing, are illustrated by the following figure:

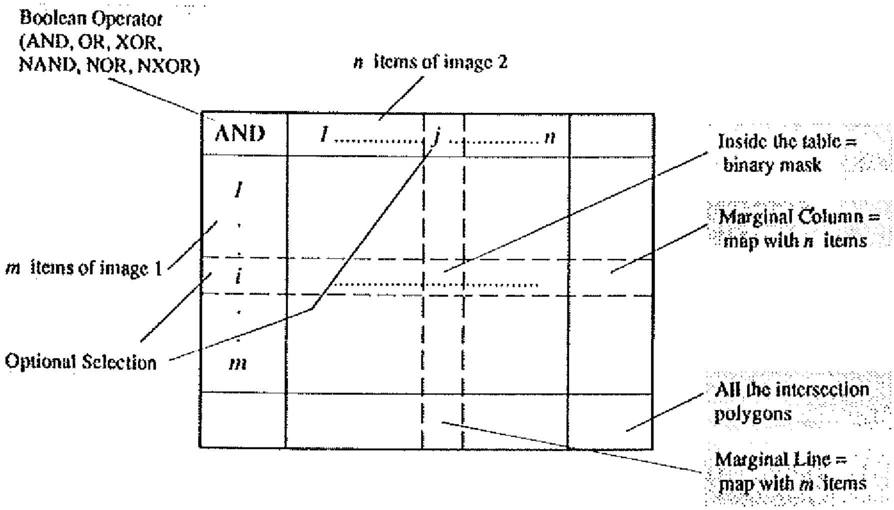


Figure 3. Intersection polygons between two raster images (Every cell in the table leads to a mask or a map)

Extraction_of_information
Obtaining and storing a map resulting of intersection of different layers makes only one aim of overlay operation. Evaluating absolute and relative surfaces of new polygons is also often requested and must be given under the form of a reporting or of a file feeding a data bank. In raster mode, every pixel covers a constant surface determined by the map scale. The surfaces estimation boils simply down to a frequencies counting of the D.N. present inside the image. So when the final image is get according to pixel by pixel multiplication of the original image by a binary mask, D.N. counting into the original image is conditionned by the presence of a unitary pixel at the corresponding place into the mask.

VECTOR STRUCTURE HANDLING
One can see that the raster mode is particularilly well dedicated when executing the basic G.I.S. operations. Nevertheless, if the selection criteria applied to the different layers are of spatial type, the raster structure is less efficient. The order of the data elements as stored in an image is dictated by their geographical position, so that a certain amount of spatial interrelations can be easily found among the pixels. But identification of basic topological concepts (point, line and area) is not obvious because topological relations between geographical entities are not present in the raster structure. All these characteristics, on the contrary, are endured without difficulty by an appropriate vectorial structure.

The conservation of vectorial structure, of its topological properties and the possibility of a vector overlay on a raster image permit to solve efficiently the spatial selection problems. It is worth noting those overlay possibilities are given by all video technology CRT, working in vector to raster mode. The graphic superposition of vectorial informations on a raster image causes no practical problems: both dimensions of the picture, lines and samples, are associated with the cartesian axes of the vectorial plane, Y and X respectively. The vectorial drawing primitives, Move, Draw and all secundary functions, are regulated in the image plane by the same algorithm as the one used in vector mode (Bresenham, for example): the algorithm determines the list of the pixels belonging to the drawing as it fixes the successive movements of the motors of a plotter for instance.

Areal element

The first example used to illustrate the complementarity of the two graphic modes consists in individualizing a region of interest on a raster image according to a boundary conserved in a vector format. The first step is the drawing of vectorial informations superposed to the raster image. As the drawing is made in vector mode, it can be treated in all the ways given by this graphic mode (generalization, smoothing, use of colour or symbolized lines, etc.). If several boundaries are overdisplayed and if there is only one to be used to delimit the R.O.I., the selection can be made interactively according to a classical vector algorithm (plub-line algorithm). Once the border selected, the raster image and the other vectorial data are erased. The only remaining boundary undergoes then a filling treatment with a uniform colour, according to a raster or vector algorithm. The resulting figure is in fact a binary mask which multiplied by the original raster image, pixel by pixel, will show the desired R.O.I.

By the end of this process, rasterization occured but it is important to note that it is not rasterization of all vector elements but only a binary mask building in raster format, compatible with the treatments of efficient overdisplays allowed by this mode. The structure of the boundaries has not been modified and conserved all its properties.

Linear element

The second example will show the technique on the base of a linear element. The application consists in delimiting a corridor of a given width around the vectorial linear element. The first step is identical to the one in the previous example: vector mode drawing of the overdisplayed elements and occasional selection of the desired line. The building of the corridor uses Bresenham's algorithm slightly modified: on every pixel selected by the algorithm on the line path, a circle is centred, diameter of which corresponds to the width of the corridor, and which is filled by a uniform colour. The so obtained picture is a binary mask that can be used by the raster-based G.I.S.

A variante to this application consists in determining a ring around a border of an area. Two binary masks are made: the one isolating a corridor around the boundary, considered as a linear element, the other covering the area like in the first example. This last mask is inverted (boolean operator NOT) and a logical

AND operation, pixel by pixel, with the first mask gives the wanted final image under the form of a new binary mask.

The techniques illustrated by those examples can also be applied to ponctual entities with potential masking of a selected neighbourhood, delimitated by any shape or any distance-decay function. Moreover, it is possible to process several vector elements in only one pass, whatever is their type (areal, linear or ponctual).

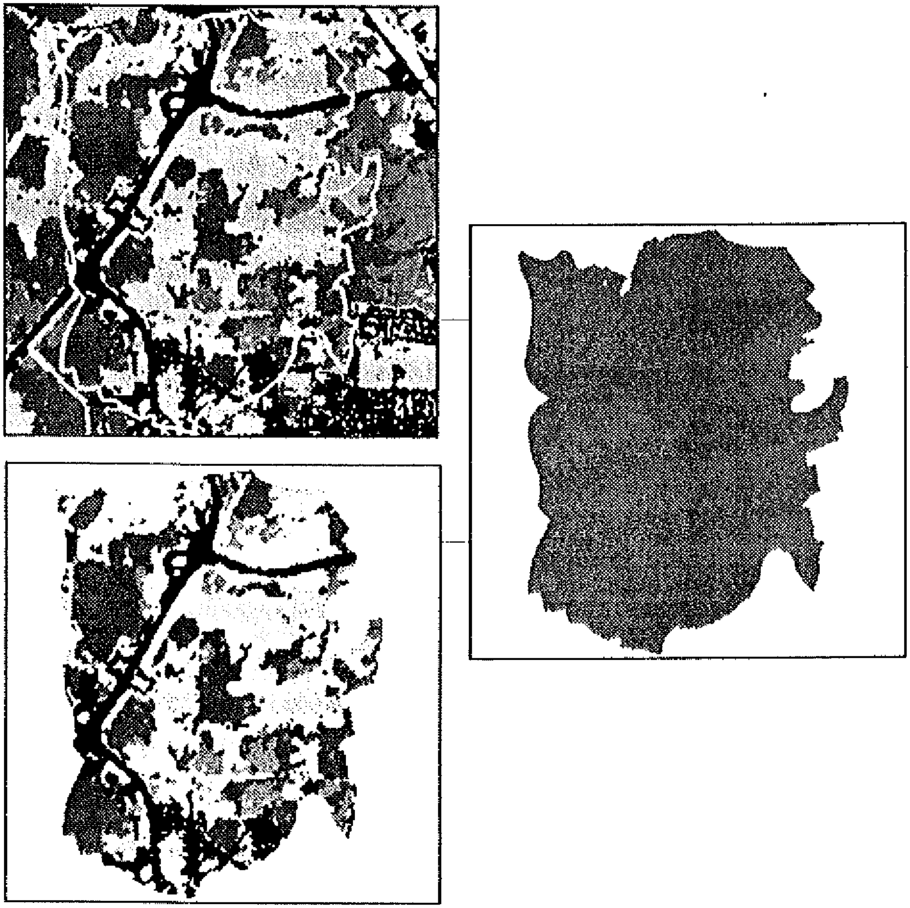


Figure 4. Vector formated area masking a R.O.I. on a raster image

Spatially continuous phenomenon

A last example of complementarity between raster and vector structures is related to phenomena with spatially continuous distribution, in opposition to ponctual, linear or areal discrete entities. So it is with natural phenomena but also with outcomes from spatial models (potential, trend surface, etc.). At first, the data are obtained from measurements at irregularly or regularly spaced points or from a set of contour lines. The continuous variation of the phenomenon can be produced by interpolation over space. The vector format is well adapted for structuring data points or lines. The raster format, thanks to the fact that it brings a discretization of space (finite number of pixels over the whole space under consideration) suits very well to the continuous surface handling.

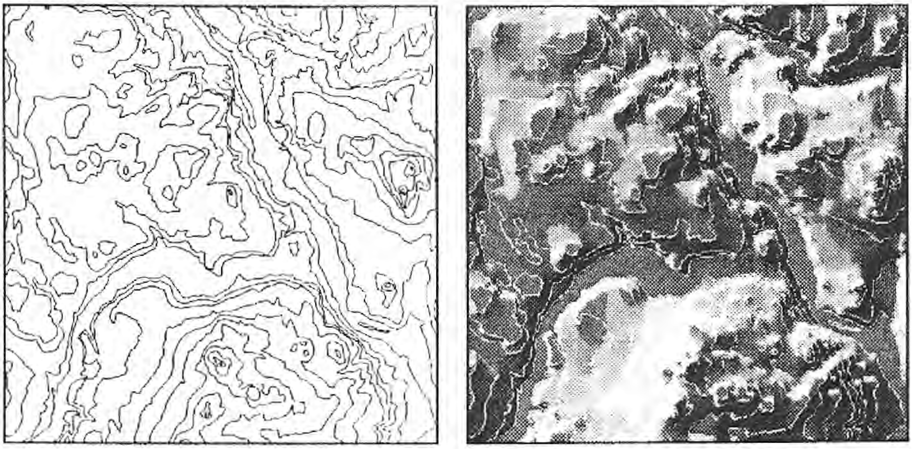


Figure 5. Shaded D.E.M. processed in raster mode from vector formated contour lines

The case of digital elevation model is illustrative. Scattered ponctual elevations or contour lines are recorded in a light vector structure, whereas the interpolation process is carried out for every image pixel in raster format. In this particular example, the derived products of the D.E.M. (gradients, aspects, shading, etc.) are easily get by means of convolution filters over the raster D.E.M. (Burrough, 1986). The applications of hybrid G.I.S., allowing vector input and raster output, are also quite numerous in the field of formal modelling and spatial simulation (Weber locational model, Voronoï polygons, etc.).

CONCLUSION

To conclude, it seems that the two graphical modes raster and vector, which often have been presented as incompatible, must be regarded as complementary. G.I.S. have not to exclude one mode on behalf of the other. Permitting the simultaneous taking into account of both formats, keeping the properties of each other and allowing the most efficient algorithm according to the process, constitute the recommanded way for future G.I.S. Moreover the raster structure, until now essentially reserved to image-processing only, has proved to be very suitable for supporting basic level G.I.S. operations.

REFERENCES

Burrough, P.A. (1986), Principles of Geographical Information Systems for Land Resources Assessment, Clarendon Press:

Oxford.
Donnay, J.P., (1989), GIPSY: Geographical Information & Processing System, Final Report for IBM ACIS Study Contract.
Fallon, K. (1990), The Perfect Couple: New methods of marrying raster and vector data enhance AEC drawing management, Computer Graphics World, 13-2, pp. 77-80.
Schneider, E. (1985), IBM Enhanced/Professional Graphics Hardware-Software-Positioning, IBM International Systems Centers, Document Number GC24-1720.