GIS - Digital Image Interaction

Mathias J.P.M. Lemmens

Delft University of Technology, Faculty of Geodesy,
Institute of Photogrammetry and Remote Sensing,
Thijsseweg 11, 2629 JA Delft, The Netherlands

Commission IV

Abstract: Images are the most important data source to achieve spatial information. The most convenient representation is the map. Digital representation of both enable a numerical interface. The profit for images is improved analysis and for GIS -the digital counterpart of the map-automatic updating. A basic concept for interfacing image and GIS is developed and a basic theory for knowledge based analysis is presented. The principal ideas are: (1) a GIS is an external knowledge base, although the data is sometimes out of date, and (2) the image analysis should be performed shape-based rather than pixel-based, to incorporate the spatial relationships of pixels and achieve a better accuracy. No shape analysis system is operational at the moment. But the GIS knowledge enables the use of spatial relationships.

Introduction

The growing world population, faced with limited and finite natural resources, stands in urgent need of recent spatial information in order to arrive at a wise arrangement, planning, management and monitoring of the planet earth. One of the most convenient data sources are images taken from airborne or spaceborne platforms. The principal storage medium is the map. Advanced microtechnology enables image data to be processed digitally and spatial information to be stored and manipulated digitally in the form of a geographic information system (GIS). Spatial digital images and GIS are unfrequently interfaced, notwithstanding the fact that more than two decades ago, the considerable possibilities of an integration, especially with regard to remote sensing data, was recognized. Marble and Peuquet (1983) have summarized a couple of reasons why spatial digital data, especially RS data, is rarely employed as direct GIS input. The two main reasons are: (1) the lack of positional accuracy and (2) the lack of thematic accuracy caused by the poor capabilities of the common multispectral (MS) classification methods. Ancillary data, however, can improve substantially both positional and thematic accuracy. An important source of ancillary data is a GIS itself. Because both GIS and image are in digital format, the interface can be exerted in a numerical way. The GIS data can improve the analysis of the spatial digital image data and reversely the image data can improve the GIS-data in the sense of updating in an automatic way. Actually, this approach corresponds literally to the analogue one, in which an interpretator uses a reflecting projector for direct tracing. The analogy is visualized in fig. 1.

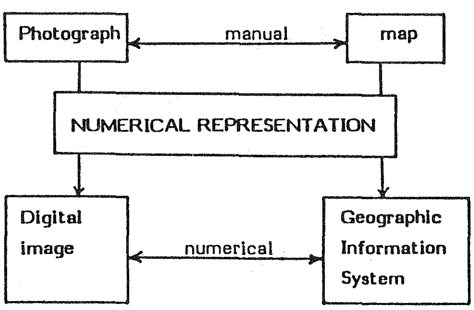


Fig. 1 Analogue-digital correspondence: basic view on GIS-digital image interaction

In the present paper first a theoretical background for information extraction from spatial digital images is developed and next a basic theory to interface these images with Geographic Information Systems (GIS) is presented. By spatial digital images we understand EM-images of the earth, taken from airplanes and satellites, and computer accessible. A basis for integration of remote sensing imagery with GIS is already roughly sketched on a more or less intuitive foundation in (Lemmens, 1987). The purpose of this paper is to generalize this basic theory for all kinds of spatial digital imagery, to refine the formulation and to extend the concepts.

A basic theory for information extraction from spatial imagery

First the information extraction process from EM-images is sketched and next the basic theory for image analysis is developed. The basic theory, presented below, refers to spatial digital images, although the model is valid for all kind of images.

The information extraction process from EM-images

Information extraction about real world objects from images involves four steps (see fig. 2):
- image formation (imaging);
- preprocessing;
- analysis;
- presentation.

Since image formation is performed by physical systems, operating in a physical environment, deviations from the model description are inevitable. Both radiometry and geometry is affected. Radiometric distortions are, e.g. caused by lens aberations, sensor saturation and atmospheric absorption. Atmospheric refraction, lens distortions, and non-rectangularity, unflatness of the image plane and so on, cause geometric deviations. Sources of spatial digital image data are, e.g.:
- digital satellite data, in particular Landsat (MSS and TM) and SPOT;
- scanned aerial and spaceborne metric photographs (black-and-white, colour, infrared);
- scanned multispectral photographs;
- microwave data;
- aerial scanning data, both optical and thermal.

Scanning of analogue images can be performed with such equipment as drumscanner, CCD-camera and video-camera.

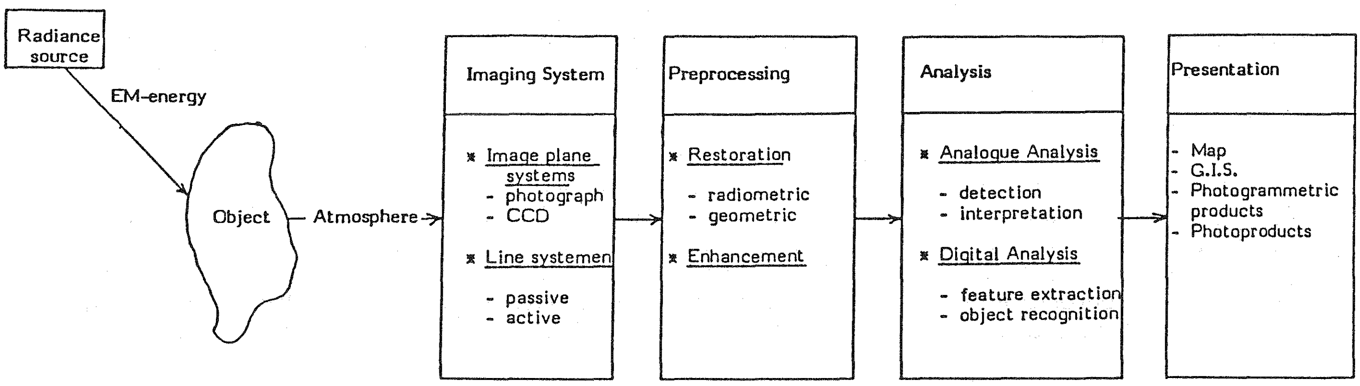


Fig. 2: The information extraction process from EM-imagery

Before analysis commences, a correction, i.e. restoration, is necessary, both radiometric and geometric. Geometric distortions are commonly rectified by using a priori knowledge about the distortions determined by calibration and/or geometric transformation models (GTMs). Image-GIS link requires also a GTM. Often also enhancement, like smoothing and sharpening, i.e. low-pass and high-pass filtering, is needed. Its purpose is to better visualize the information for further manual or computer-assisted analysis. The analysis process consists of two stages: (1) feature extraction (or segmentation) and (2) pattern (or object) recognition. The aim of feature extraction is to split up the image into regions which are homogeneous in some properties, e.g., same grey value, same texture and the same multispectral characteristics. Feature extraction techniques can be subdivided into four broad classes:
- grey value thresholding;
- boundary detection;
- region growing;
- clustering (multispectral images only).

Grey value thresholding segments on the basis of grey value intervals. It works only well in case of few region types. In particular, it is appropriate when one object is situated against a background. The threshold is determined from the bimodal grey value histogram. This situation occurs often in industry , i.e. controlled environment, but for the complex spatial images it isn't practical, although it is applied for road detection in small scale images by (Bajcsy and Tavakoli, 1976). In boundary detection first grey value or texture changes are traced. Grey value changes are detected using gradient operators, leading to an edge image. To trace texture changes, first texture analysis has to take place, e.g. using the texture measures proposed by Haralick (1978); each pixel is assigned with one texture measure. So, the texture image may be viewed as a grey value image and, consequently, gradient operators will trace the texture changes. Although displayed edges form connected structures for the human eye, the computer has to track the edges, that is: boundary following. Boundary detection plays a key role in our basic theory. Region growing starts with a small initial region. Its statistical properties are computed, i.e. mean and variance. Neighbouring pixels, which show sufficient similarity, are added. The statistical properties of the enlarged region are computed, neighbouring pixels are evaluated, and so on. For spatial imagery, with often weak boundaries, it has the disadvantage of missing boundaries. We propose a modification to find regions with the same label in the classified image. Clustering of MS images relies on the spectral separability of objects. Man-made objects show only poorly spectral differences. Using training samples clustering leads to MS classification. Digital spatial images are mainly analyzed by MS classification, although, for some years, texture analysis earned interest.

The computer vision community often refers to radiometric restoration and enhancement as low-level processing: image processing leads again to an image. Feature extraction is often called mid-level vision: from images data to properties of, or relationships between, extracted features. Pattern recognition is often refered as high-level vision: the features from mid-level vision are labeled, according the classes the real world is divided.

Basic theory for image analysis

The process of image analysis is based on a certain amount of stimuli: visual parameters or basic elements of interpretation. These parameters are: (1) tone and colour, (2) texture, (3) shape, (4) size, (5) height and (6) shadows. Often also: (7) pattern, (8) site and (9) association are added to this series and we adopt this addition. They are strongly related to the cartographic expression elements, which is not astonishing, since a map is an artificial image. At present just colour and texture is operational in classifying spatial images. (Note that in our nomenclature colour is not restricted to the visible part of the EM spectrum.) MS classification has a very simple segmentation procedure, since each pixel is viewed as a segment, i.e. just the grey value vector of each pixel is evaluated but the neighbouring pixel information is neglected. Although in texture classification a small neighbourhood of each pixel is involved in the classification, the method is actually, like multispectral classification, a pixel-based, or for short, pixel classification.

It is well known from vision theory that shape is the most important parameter in the human visual system. But, because of its computational complexity, it is never used in remote sensing. In (Lemmens et. al., 1988) some experiments are investigated to recognize roads by shape. The developments in computer sciences invite to look for a scheme to formalize the basic elements of interpretation.

An image is a two-dimensional distribution of greyvalues. It maps the luminance distribution, covering a three-dimensional object space at a certain moment t: J(X,Y,Z,t), onto a two-dimensional image space: I(x,y,t). Ignoring the time parameter:

$$J(X, Y, Z) \longrightarrow I(x, y) \quad ; \quad x, y \in \mathbb{R}$$

A photographic film in the image plane transfers the luminance distribution to a grey value distribution, which is permanently retain:

$$I(x, y) \longrightarrow F(x, y) \quad ; \quad x, y \in \mathbb{R}$$

By placing a sensor, a CCD-line or a CCD-array in the image plane a sampled and quantized, that is, numerical, representation is originated:

$$I(x, y) \longrightarrow G(i, j) \quad ; \quad i, j \in \mathbb{R}$$

G(i,j) is a digital image, i and j are integers, defining a pixel. Each pixel covers a certain area and its number is called grey value. When a photograph is transfered to digital format, using, e.g. a drumscanner or CCD camera, the transformation becomes:

$$F(x, y) \longrightarrow G(i, j)$$

MS images contain several bands, i.e. each pixel consists of a vector of grey values: $G^k(i,j)$,-k=1,..,m. So, a multispectral (i.e. colour) image is a combination of grey value images. Classification is performed by statistical analysis of the grey value vector of each pixel.

The grey value of each pixel is, in general, strongly correlated with that of the neighbouring pixels, since otherwise the image would contain only white noise and thus no information at all. The kind of spatial variance of the grey values defines the texture. So, texture is a measure for the repetition of grey-values of a neighbourhood. Abrupt grey values and texture changes lead to contours. Contours surround areas, each with their particular shape. So, contours express shape. Grey value and shape are the primary elements of analysis. (Note that this corresponds to the two main approaches of matching of stereo images: signal matching and feature matching.) The other basic elements are derivable from them, using the following functionals:

- combination;
- repetition;
- extensiveness.

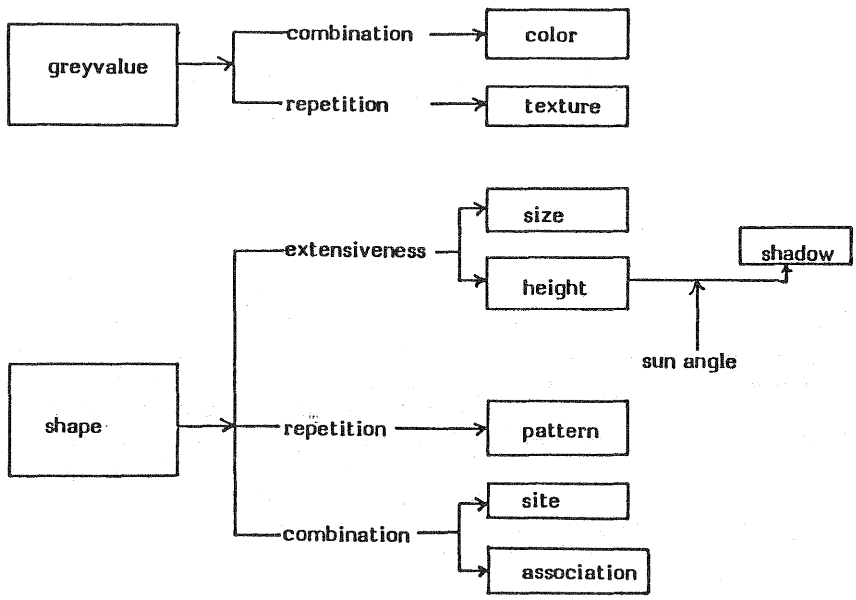


Fig. 3: A model describing the interrelationships between the basic elements of interpretation

For MS images, combination of grey values leads to colour and repetition of grey values to texture. The functional extensiveness has no consequence on grey values. Extensiveness is a measure for the spatial distribution of a shape. In the horizontal plane this measure is size, in the third dimension height. Illumination creates shadow from height. So, shadow gives information about shape and height. Repetition of shape gives pattern. Combination of shapes leads to the elements site and association. Site means that the relative position of the distinct objects give information about these objects. Association means that several shapes of the same object yields information about that object. The above is summarized in fig. 3. An important source of shape is a GIS.

GIS

The GIS phenomenon originated from digital technology. Actually, a GIS is a replacement of the classical map. But for many applications the digital capabilities are far beyond those of their predessors. A map has two functions: (1) data storage and (2) data representation. These tasks are conflicting. The first task demands the storage of all relevant data. But the representation function requires, according to the principal law of cartography, clarity of view. Otherwise the map is unaccessible for the user and consequently contains no information at all. The cartographic requirement, however, may drop too much data beside the edge of map. The present manner of map management is, besides information demands, dictated by the limitations of the early data acquisition equipment, data storage facilities and representation means. Human being is highly accustomed to these historical developed but nevertheless defective concepts, even such that they are gratuitously transfered to the new technologies. But they have their own limitations and properties, not comparable with those of maps.

One of the most important properties of a GIS is that it disconnects the classical data storage and representation task of a map. Its main purpose is the storage of spatial data. For

digital storage and handling, the data have to be encoded and structured. Encoding can be performed in two formats:
- vector format;
- raster format.

In vector format encoding, objects are described by polygons. The connections of vertices define lines, connections of lines define arcs and area. The encoding can be done by:
- storage of the coordinates of the nodes;
- nodes, line segments and polygons are stored in a topological way.

To the nodes, lines and polygon areas, attributes are assigned describing some themes, sometimes called tabular data to distinguish it from vector data.

Raster encoding means actually that a grid is superimposed over the surface of the earth. Each grid element (or elementary cell) is provided with attributes, each of them may be viewed as an overlay. The grid size depends on the application and the available information density, and may vary within one GIS. Raster data demand a large storage volume and a balance between level of detail and ease of implementation. Because there is a direct link between position and theme, raster format is much more efficient then vector format, both in data storage and manipulation, but it is much less accurate and the curbs of an arbitrary raster is rather artificial. Besides, for many data types a raster format isn't adequate. Continuous data, like population density, soil type, air pollution and precipitation are best suited for raster storage. A vector format is more suitable for storage of discontinuous data, such as land use, vegetation cover and ownership.

At present the primary sources for data input are maps. Advanced equipment, like scanners and raster-vector convertors are used to automize GIS formation. As in classical mapping and surveying, the three spatial dimensions are commonly separated into plane and height. Positions in the plane are referenced by Cartesian (x,y)-coordinates or geographical coordinates. The height is stored as a digital elevation model (DEM). The primary data sources are stereo images, mainly recorded from aerial platforms, but the recent SPOT satellite is developed to produce stereo images too. The last years, the photogrammetric community researches intensively the potentials of digital image matching techniques to arrive at an automatic DEM production. High precision automatic tracing of corresponding points is already operational for DEM production of industrial objects. A survey on matching techniques can be found in (Lemmens, 1988).

Besides encoding the data structure is important, i.e. how to arrange the data to be (fast) accessible. It is a compromise between computer architecture and user demands. In the course of time several data structures, like hierarchical, relational, network and semantic, are developed, but none are satisfactory. Artificial Intelligence (AI) reveals new hope, but one has to keep in mind that AI is just a combination of search techniques, efficient data storage and symbolic data comparison. The computer languages are optimized for manipulation of symbolic data, but an operational design needs very fast hardware and very much storage.

Once a GIS is created, the user must be able to query it and to ask logical questions. So, manipulation (or analysis) and query facilities are necessary. An output facility puts the demanded information to the user. This is, e.g. achieved by visualizing maps from displays or line-plotters and/or by listings of statistics. Because human being continually changes the view of the earth, spatial information becomes already after a few years obsolete, so a GIS needs an input facility to be refreshed. The main sources are images.

So, a GIS efficiently stores, retrieves, manipulates and displays spatial data (Marble and Peuquet, 1983) and it is made up of four interrelated subsystems:
- input;
- storage;
- manipulation and analysis;
- output or representation.

The above subdivision is generally accepted in literature, but bears often other names. It corresponds also to the four stages of information extraction from images. This is not astonishing, since each information extraction process can be divided into the present four stages.

Let's formalize the above a little more. The fundamental spatial data types are: (1) points, (2) lines and (3) areas. Often a set of connected lines, that is an arc, is added as fourth type.

GIS data consists of: (1) positional features and (2) thematic (or semantic) features. To each spatial data type a set of positional descriptives P, consisting of the attributes p_i, i=1,..k, is assigned, describing some positional properties of the datatype and a set of thematic descriptives T, consisting of the attributes t_i, i=1,..,l, describing some thematic properties. There may exist a set of rules to transform the one representation into the other.

The positional attributes of a point will be its spatial coordinates and attributes indicating the measuring precision σ_m. The elementary positional property of a line is its length (l), also provided with a precision measure, derivable from those of the coordinates. Positional properties of an area are, e.g. area, perimeter and elongatedness, each assigned with a precision measure derivable from those of the coordinates.

The three data types are also assigned with a set of thematic attributes, e.g. land use and cover type. The thematic information has a limited accuracy too, which affects the geometric accuracy, e.g. the boundaries of a didge may be recognized unprecise, although the measuring precision may be high. So both positional and thematic precisionσ_t determine the final geometric precision:

$$\sigma_g^2 = \sigma_m^2 + \sigma_t^2$$

The thematic precision is difficult to quantify. We divide the set of thematic properties into:
- administrative or land management properties;
- physical properties.

Administrative properties are, e.g. ownership, selling value, and building rights. Of course, they aren't derivable from image data. Moreover, images don't contain all of the physical information, e.g., crop cover and soil type, but just a subset. The remote sensing community has experienced very much, the last two decades, to trace the physical information obtainable from spatial imagery and still, these investigations are going on.

Concerning the spatial changes, forcing GIS update; with a look at fig. 4, representing some polygons from a vector format GIS (for convenience just single valued) the following situations may occur:
- both positional and thematic features remain unchanged;
- positional features remain unchanged, the thematic features change;
- positional features change, thematic feature remain unchanged;
- both positional and thematic features change.

The thematic changes may refer to all three elementary data types: nodes, lines and areas. Positional changes are caused by vanishing boundaries and/or new created boundaries.

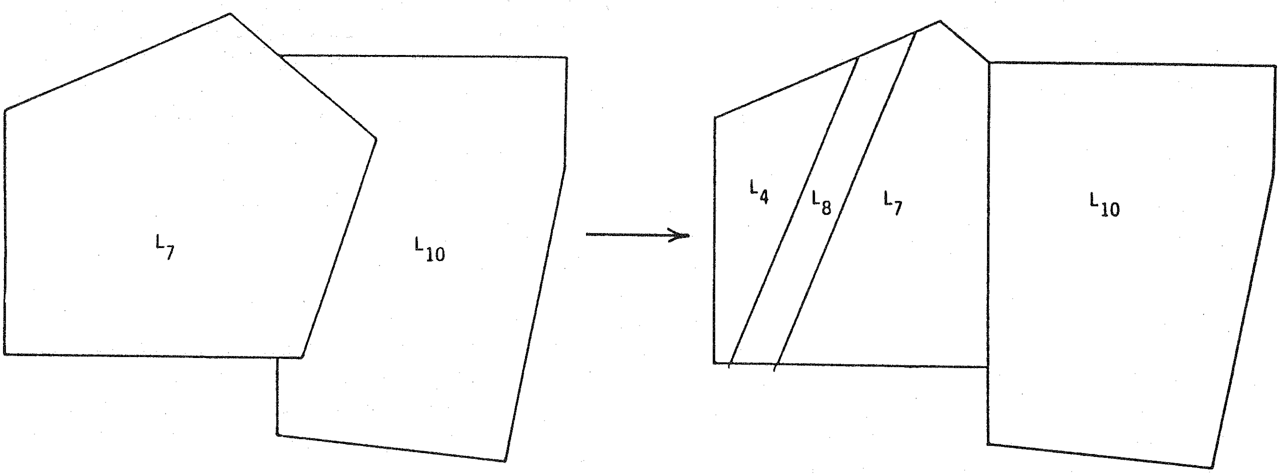


Fig. 4 Some changes in thematical and positional information of the GIS

The next section shows how the GIS is interfaced with digital images to detect them and to use the GIS knowledge for a more accurate analysis of the images.

The Interface

Pixel classification assigns to each pixel a label independent of its neighbourhood. Just training samples are used as external knowledge. But external knowledge consists of several elements (Taylor et. al., 1986):
- models of the imaging process;
- knowledge about the types of structures in the scene and their relationships;
- other data sets: (general knowledge about the scene, GIS, maps);
- expert system.

A GIS is a global knowledge base, which enables the incorporation of shape in the analysis process by a loophole.

The interface process can be divided into three stages: (1) geometric interface, (2) positional interface and (3) thematic interface. The positional link is just of interest for the vector format GIS, since the thematic interface is only indirect -over the positional link- coupled with the position. The themes of raster data are directly coupled with position, so they don't need a positional interface.

We will first treat the geometric interface, next the image-raster format GIS interface and finally the image-vector format GIS interface.

Geometric Interface

The common approach corrects first the image data for all known interior and exterior distortions, e.g. attitude and position deviations and next computes the parameters of the geometric transformation model (GTM), using control points. Common GTM's are: similarity, affine and second order polynomial transformations. Next the image and GIS must be referenced to the same system. There are several possibilities:

- transform GIS and image to an exterior reference system;
- transform the image to the GIS reference system;
- transform the GIS to the image reference system.

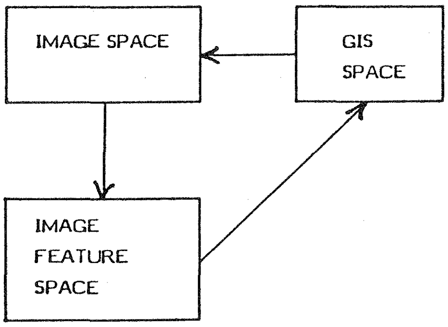


Fig. 5

Because a GIS is already referenced to an earth system the long way about of transforming both systems, can, in general, be avoided. The transformation of the entire image to the GIS requires a costly resampling. More appropriate is to transform just the detected image features, because of the tremendous data reduction, i.e. backproject the GIS features into the image (see fig. 5).

Image - Raster Format GIS Interface

There is a direct link between image and GIS (see fig. 6). The image is supposed to be pixel classified. For convenience, let there be a 1-1 correspondence between image and GIS, that is: both elementary cells have the same size. (Generally the image cells will be smaller then the GIS cells, but a non 1-1 correspondence doesn't effect the method noticable.)

Fig. 6

The most simple update execution is to assign to each GIS cell the label of the corresponding image cell. However, this inevitably leads to misclassifications. First of all the likelihood of changes isn't used, e.g. a change of grass-land to full-grown oaks, in a few years, is extremely unlikely. A further improvement is to take the neighbouring cells into account. Since each cell in a rectangular tesselation has eight neighbours, topological information could be appropriate to remove misclassifications. But, what to do if just one pixel in the 3x3 neighbourhood is identical to the midpixel? So, a purely topological approach is just suited to remove isolated pixels, i.e. pepper and salt noise removal. Moreover, it is a deterministic approach, whilst, pixel classification is a statistical pattern recognition technique, explicitly shown in maximum likelihood MS classification. However after label assignation the probability values are dropped. Using neighbourhood information is well known in digital image processing as relaxation. Relaxation techniques are iterative procedures to incorporate neighbourhood information to support or diminish the likelihood of the presence of a phenomenon at a location. There are many ways to implement them. We propose the following scheme for pixel classified images.

Assume that a certain pixel has on basis of its spectral or textural characteristics probabilities P_i, $i = 1,..,m$, to belong to label L_i. Note that $\mathrm{sum}(P_i) \neq 1$. Without loss of generality we may rearrange P_i such that $P_1 > P_2 ,>, P_m$. Suppose that just P_j, $j=1,..,k$, $k< m$, exceed the probability threshold P_t, under beneath no reliable label assignation can be carried out. (If all $P_i < P_t$, the pixel achieves the label 'unknown'.) Commonly, the label with the largest probability is taken and all other information is dropped, although it may happen that P_1 and P_2 and even P_3 don't differ much. To decide which label is the realistic one, neighbouring information has to be incorporated. To come to an outline, just P_1 and P_2 are supposed to exceed P_t. Depending on the label probabilities in a 3x3 neighourhood, P_1 and P_2 are augmented with a constant K, remain unchanged or are decreased with K. The choise of K depends on the required accuracy and convergence rate of the iterations, e.g. k=0.1, or 0.2. Of course, when a probability reaches its limit, i.e. 0 or 1, the limit is not exceeded.

In fig.7. some evident cases are listed for illustration. The numbers (e.g. 4 and 7) refer to labels. The first row of 3x3 neighbourhoods shows labels of largest probability, i.e. P_1. The second row lists the labels for P_2. A dash means that the probability is below the threshold.

	case 1	case 2	case 3	case 4
	4 4 4	7 7 7	7 7 4	7 4 4
P_1	4 7 4	7 7 7	7 7 4	7 4 4
	4 4 4	7 7 7	7 7 4	7 7 4
	6 - -	5 - 3	3 6 5	- 7 7
P_2	- 4 -	- 4 -	4 4 -	- 7 7
	- 3 9	3 - 7	3 4 -	- - 7

Fig. 7

In case 1, the midpixel of P_1, is surrounded by deviating pixels. No support is given by P_2 that the label is something else than a misclassification. So, P_1 is decreased by K, but P_2 is augmented, because it supports the neighbourhood labels of P_1.

In Case 2, the situation is clear; the midpixel is surrounded by identical labels. P_1 is augmented with K, P_2 is decreased.

In case 3, the midpixel could also belong to label 4. This is even supported by P_2. But we can't make a fair decision, so both P_1 and P_2 remain unchanged.

In Case 4, P_1 is also a don't care situation, but P_2 supports label 7. So, P2 is augmented. If after relaxation P_2 is larger then P_1, label 7 is assigned to the midpixel.

The above is just meant as example. It is't operational.

Image - Vector Format GIS Interface

In our model the primary basic element of interpretation is shape. A vector format GIS already contains (old) shapes in the form of polygons. We have just to check whether these boundaries are still there, whether new boundaries are created and whether thematical changes appear, i.e. postitional and thematic changes.

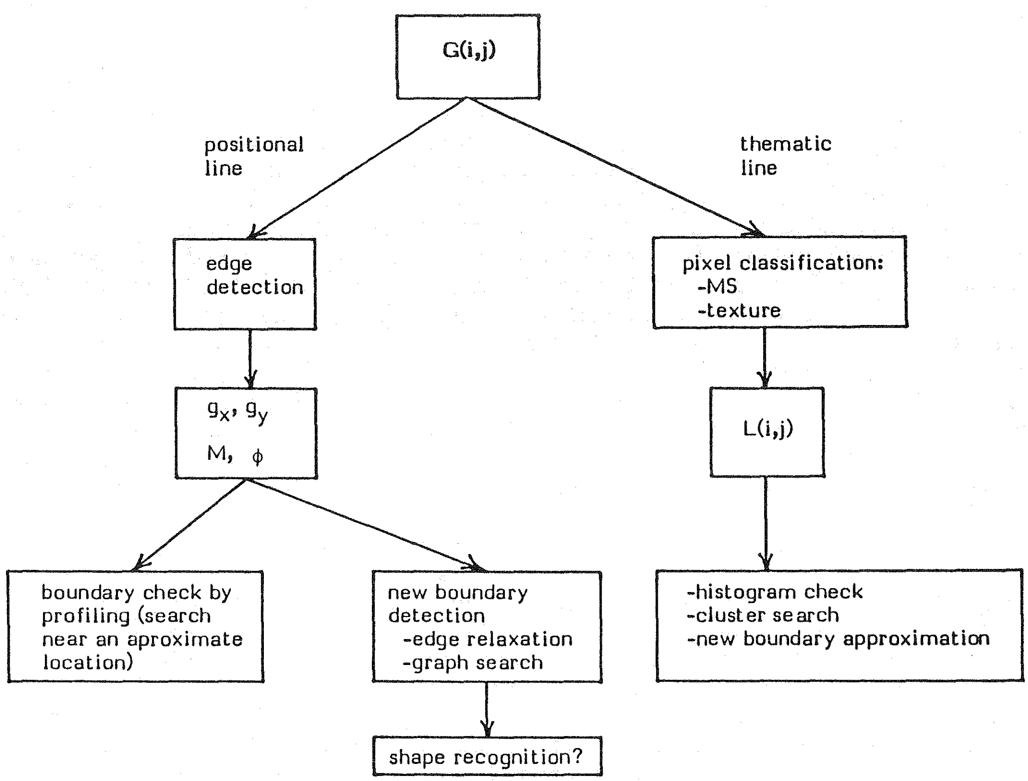


Fig. 8

Thematic Interface

The image G(i,j) is first pixel classified, leading to the classified image L(i,j). Noise causes misclassification, i.e. a minor group of pixels with deviating labels. So, use the GIS polygon as seat-belt to incorporate neigbourhood information. For each label the pixels are counted and stored in a frequency histogram. If the percentage of pixels with the same label exceeds a certain threshold, e.g. 90%, the label is assigned to the whole polygon. It may also occur that two or more labels have a high percentage. That indicates that the parcel is splitted up: a new boundary is created. A clustering method, e.g. region growing, is able to check L(i,j) change. The boundary between the regions can be roughly located in the label image. This location may be used as approximation for refined boundary detection methods, like edge relaxation and graph search, to locate the precise boundary in the grey value image. Edge relaxation and graph search are used in (Lemmens et. al. 1988) for road recognition in large scale images. Labels with a small percentage will be caused by noise and ignored. The above is sketched in fig. 8, following the thematic line and treated in more detail in (Lemmens and Verheij, 1988). The adjustment of upper and lower bounds depends much on precision. The following features affect the precision:

- the quality of the vertices of the GIS polygons;
- the quality of the image data;
- the quality of the control points and kind of GTM;
- the filter operations in low- and mid-level processing;
- the quality of classification.

Positional Interface

The positional interface is carried out using the GIS lines, defined by the coordinates of the node points. So, for the present purpose boundary detection is adopted as primary feature extraction method. It consists of (see introduction): (1) edge detection and (2) boundary following.

Edges are individual elements, marking grey value (or texture) changes. In Fig. 9 masks of some common edge detectors are listed. From the gradient images g_x and g_y, the edge strength or magnitude M and the edge orientation Φ are computed: $\Phi = \Theta + \pi/2$, with Θ the gradient direction: $\Theta = \mathrm{atan}(g_y/g_x)$. They are called magnitude image and direction image, respectively.

	g_x	g_y
normal gradient	$\begin{bmatrix} -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
Roberts operator	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Prewitt operator	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$
Sobel operator	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$

Fig. 9 some common edge detectors

$$M = (g_x^2 + g_y^2)^{\frac{1}{2}}$$

To save computation time, M can be approximate by:

$$M' = |g_x| + |g_y|$$

or

$$M'' = \max(g_x, g_y)$$

$$\Phi = \mathrm{atan}(g_y / g_x) + \pi/2$$

Several boundary detection methods can be distinguished:

- contour following;
- search near an approximate location;
- Hough transformation;
- edge relaxation;
- graph searching combined with dynamic programming.

For the present investigation search near an approximate location, edge relaxation and graph searching are of importance.

One interfacing possibility is linking after boundary detection and vectorisation of the image. Next the image shapes are matched against the GIS polygons to find corresponding ones (see fig. 10);

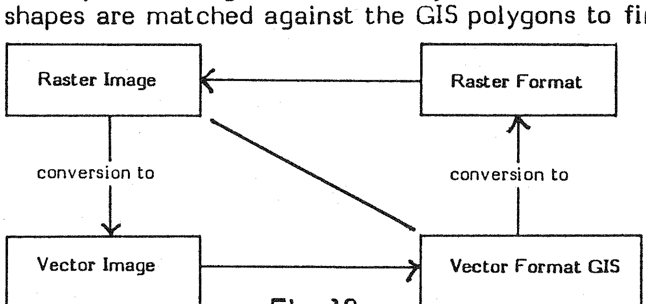


Fig. 10

But shape matching drops the positional information, is computationally quite difficult and thus inefficient. Moreover, the GIS information is employed in a rather late stage. Because the location of the GIS lines are known in the image up to a certain accuracy from the geometric transformation, their present existence can be checked, using profiles perpendicular to the line (see fig. 11). They are evaluated with respect to grey values changes, which indicate the existence of the boundary. The evaluation can be carried out in the original grey value image, like is done by (Groch, 1981) to trace roads in small scale images, but also in de edge image. The last has the advantage, that besides magnitude also direction can be used. M is the leading parameter, but in case of doubt, i.e. two pronounced maxima are present in the cross-section, the direction may arbiter.

This is incorporated by a costfunction: $C_i = M_i - \alpha \, \Delta\Phi_i$

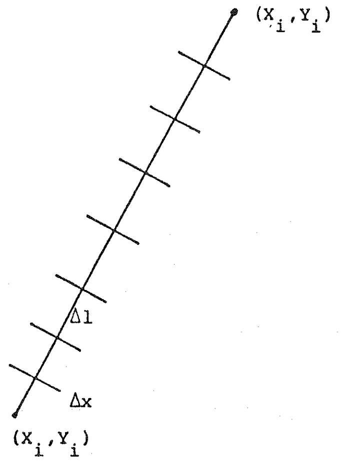


Fig. 11: Search near an approximate location
Δl defines the length of profile
Δx defines the interval between two profiles.

With: $\Delta\Phi_i = |\Phi_l - \Phi_e|$

α: a weighting factor determining the relative influence of M_i and Φ_i

Φ_l: the direction of the GIS line

Φ_e: edge direction

Suppose a profile is m pixels (Δl) long (the length depends on the precision of the interface) than pixel k will be part of the boundary if:

$$C_k = \max(C_i), \quad i = 1,..,m \; ; \; C_k > C_t$$

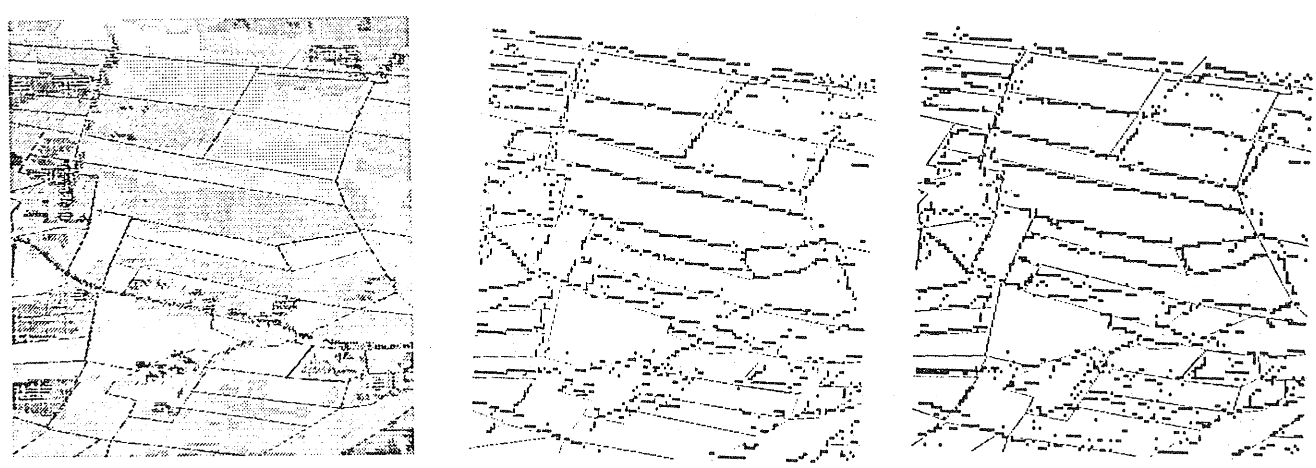
with C_t a certain threshold defining a lower bound. For computational convenience and storage save, both M and Φ can be rescaled to the range 0 - 255.

An Example: Fig. 12 shows an image recorded by an optical scanner. GIS boundaries are backprojected into the images. Fig. 12a shows the images preprocessed by a conditional average smoothing. Fig. 12b shows the result of profiling with costfunction:

$$C_i = \max(M_i) \; ; \quad C_t = 10$$

Fig. 12c shows also a result of profiling, but now with costfunction: (M_i and Φ_i rescaled):

$$C_i = \max(M_i - \alpha \, \Delta\Phi_i) \; ; \quad \alpha = 0.1, \quad C_t = 10$$



a b c

Fig. 12

Since each GIS line is provided with an attribute, the kind of boundary one may expect can be indicated. In fig. 13 some kind of edges are listed. This knowledge may be used with advantage to evaluate the profiles. They define the initial width (dx, or dx' in the edge image) of the boundary lines and the difference between the greyvalues (or texture values) (dg, or dg' in the edge image) (see fig. 14). The method reaches high flexibility by allowing variable dx, dg, Δl, and Δx values, depending on the progress of line search (see fig. 11)

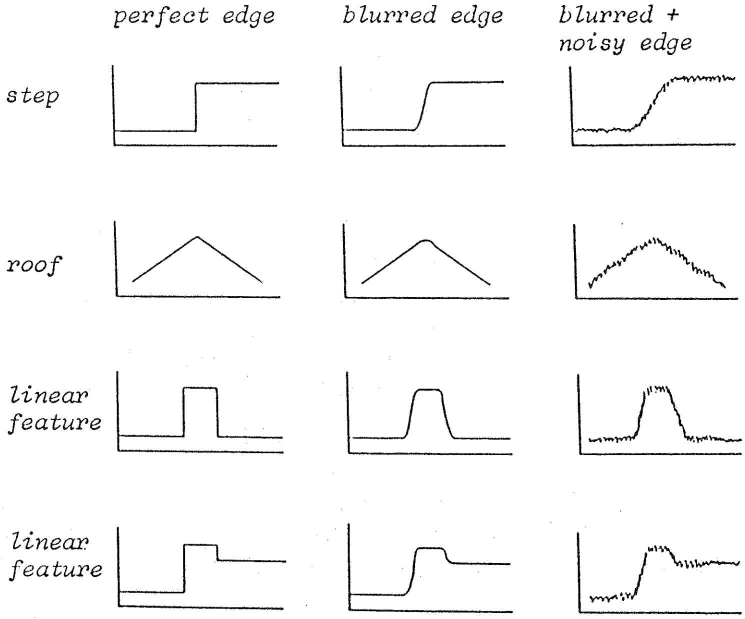


Fig. 13 Some common edge types

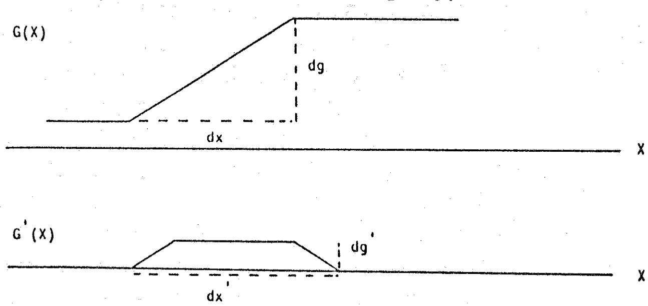


Fig. 14 Grey value distribution and gradient across an edge

Conclusions

A theory for interfacing GIS and digital spatial image for image analysis improvement and automatized GIS update is presented. Both the thematic and positional GIS information is used. The GIS information can be interface at several levels with the image (see fig. 8). How the thematic line and the positional line have to be linkend to arrive at an efficient procedure, has to be subject of thoroughly investigation. Ohter important research areas are:
improvement of the pixel classification using GIS neighbourhood information, precision analysis both of the positional as thematic data, efficient methods determination for profiling, and intrinsic geometry of feature extraction, i.e. how accurate are boundaries located by the digital operators

References

Bajcsy, R., Tavakoli, M., 1976, Computer recognition of roads from satellite pictures, IEEE Trans. on Systems, Man and Cybernetics, Vol. 6, no. 9, pp 623-637.
Groch, W.D., 1982, Extraction of line shaped objects from aerial images using a special operator to analyze the profiles of functions, Computer Graphics and Image Processing, vol. 18, pp. 347-358
Haralick, R.M., 1978, Statistical and structural approaches to texture, Proc. 4th Intern. Joint Conference on Pattern Recognition, Kyoto, Japan, pp 45-69.
Jensen, J.R., 1986, Digital Image Processing, a Remote Sensing Perspective, Prentice Hall, pp. 379.
Lemmens, M.J.P.M., 1987, Towards a model for (semi-) automatic analysis of digital images, Joint Colloquium of the ISPRS working groups II/1, II/2 and II/6, London, Sept. 21-23, pp. 15.
Lemmens, M.J.P.M., 1988, A Survey on Matching Techniques, Archives of the 16th ISPRS Congress, comm. V, Kyoto, Japan.
Lemmens, M.J.P.M., Verhey, K.F., 1988, Automatically GIS updating from classified satellite images using GIS Knowledge, Archives of the 16th ISPRS Congress, Comm. IV, Kyoto, Japan.
Lemmens, M.J.P.M., Bruel, E.W., Fennis, F., 1988, Linear feature extraction and road recognition from large scale digital aerial images, Archives of the 16th ISPRS Congress comm. VII, Kyoto, Japan.
Marble, D.F., Peuquet, D.J., 1983, Geographic Information Systems and Remote Sensing, Manual of Remote Sensing, 2 edt., Vol.1, Chapter 22, American Society of Photogrammetry.
Taylor, A., Cross, A., Hogg, D.C., Mason, D.C., 1986, Knowledge-based interpretation of remotely sensed images, Image and Vision Computating, vol. 4, no. 2, pp. 67-83.