Towards a More Human (Re)design of Digital Spatial Technologies with Emphasis On an Uncertainty-based Cartographic Representation

Kim Lowell¹, Geoffrey Edwards¹, Kim Esbensen²

¹Centre de recherche en géomatique
Université Laval, Pavillon Casault
Ste-Foy, Québec, CANADA G1S 3B7

²Centre for Industrial Research (SINTEF)
Postboks 124, Blindern
N-0314 Oslo, NORWAY

Abstract
Much research in digital spatial technologies has had the implicit goal of eliminating humans from the analysis process. This has not been successful because humans are able to carry out many tasks which are difficult to program (e.g., pattern recognition). It is appropriate to re-examine the role humans should play in spatial information systems. Technology development should be oriented towards combining human and computer processing in such a way that each assists the other by carrying out the tasks that each does best. Ways of handling spatial data based on these ideas are discussed and a new spatial data representation based on uncertainty is presented. Design issues related to integrating human analyses into computer processing are discussed and a remote sensing technique which embodies these ideas is presented. It is concluded that existing spatial data handling needs to be rethought and restructured in order to ensure continual human participation in the analysis and decision-making process.

1: Introduction

In textbook descriptions of spatial analysis performed with the aid of a geographic information system (GIS), the GIS contains a set of polygon-based thematic maps. The maps come from a variety of sources and are developed using a variety of construction techniques. For example, forest maps might be produced through human interpretation of aerial photographs, while other map data may be derived from legal or political information such as county boundaries. Regardless of the source of data or the method of map construction, once input into a GIS, all maps are dealt with as if they had equal certainty. Each line is treated as being equally reliable, and all attributes (e.g., soils class, county designation) are processed as if they are equally accurate and precise.

For both system designers and end-users a number of "nice properties" fall out of this classic vision of GIS. (Herein "system designers" are individuals who design and produce GIS software; "end users" are individuals who actually use the GIS to obtain information about the territory mapped.) First, under the conditions described, all analysis becomes objective -- n different users can ask a question of the GIS and receive the same answer. Second, a single measure of error can be assumed to characterize the entire map region. This is a consequence of the implicit assumption that all features have the same certainty. The assumption of a uniform error has the additional advantage that it can be readily handled in a GIS. Finally, results of queries are expressed as exact values -- not only is a single answer produced to a question, but it is reported as if it were the "right" answer.

As a result of this traditional vision of GIS, an implicit belief common to both system designers and end-users seems to be that a GIS is composed of a set of maps with no error, or a set of maps where error can be inferred relatively easily. Indeed one can find examples of this in GIS research on uncertainty. A number of researchers (for example, [22], [2], and [16]) have used a single estimate of error to describe error at all locations on a map. Even when such researchers have included estimates of spatial autocorrelation in their error studies, they have adopted a single, global estimate of such autocorrelation. In reality, there is no objective reason to suppose that data spatially distributed can be characterized so uniformly.

The cited work on spatial error constitutes an advance in GIS design because the assumption of maps as errorless is questioned. Nonetheless, these studies do not go far enough as they are based on a critical implicit belief: the existing thematic map is an acceptable representation for computer-based GIS. We believe that this belief constitutes a fundamental error in understanding certain types of spatial information, how it is stored and analyzed by human beings, and the role the computer plays in spatial data handling. It also indicates a failure to understand how a map of a given phenomenon is constructed. Simply stated, the acceptance of paper maps as a suitable spatial representation for a digital environment supports a misguided belief that machines can and should duplicate what human beings do, and that the interaction of humans with data analysis and information should be minimized. Instead it is our position that machines should enhance the user environment with increased information, but control throughout the analysis process should remain with the user. This reflects a position that more accurate maps cannot always be created using more accurate technologies. Thus meta-data becomes crucial to assure proper usage.
The same principles also apply to remote sensing (RS) technologies. System designers seem to believe that given enough image resolution (spectral and spatial) and a sophisticated enough algorithm that one can and should duplicate what human beings produce from a similar product. And end-users believe that if RS cannot do this, it has somehow "failed." This again is indicative of a lack of understanding about what a digital sensor is doing, the information it produces, and how the information is analyzed. Having end-users decrease their demands and systems designers decrease their promises would allow RS analysis to be redefined so as to promote an information-enhanced user environment. It would also lead to new analysis tools such as one for image processing which will be discussed subsequently.

Indeed, we believe that system designers and end-users need to reexamine the fundamentals of digital spatial technologies (DSTs -- whether GIS, remote sensing, GPS, photogrammetry or other digital domains). Classic DST design demands too much of system designers and too little of end-users. We propose to demonstrate how this thinking is counterproductive for both GIS and RS, to describe an alternative way of employing computer-based GIS and RS, and to describe initial experiments already conducted. In doing this, it will become evident that there is more common ground between GIS and RS than is currently exploited and that DSTs will be enriched, not weakened, by ensuring human participation in the intermediate stages of analysis.

2: GIS philosophy, history, and background

As noted, the data in a GIS consist of a set of maps and associated databases which are developed from a variety of sources and construction techniques. However, once they are in the GIS, these maps somehow take on an aura of exactness. System designers and end-users both may easily forget that a surveyed township map is fundamentally very different from a soils map in which extensive subjective interpretation of both geometry and attributes is inherent. The literature is full of warnings that a given map is merely one possible representation of space [27] and that classification remains a subjective process [24]. Nonetheless, this is largely forgotten once a map is placed in a GIS despite even relatively recent warnings that the lines in a GIS are not necessarily all the same [26].

To understand this, one must look at maps historically. For many years, thematic maps have been the predominant way that human knowledge about geographical phenomena has been organized given a physical two-dimensional storage medium. Three things have resulted. First, human imagination about alternative ways of representing spatial information has become limited due to a lack of potential alternative media. Second, the suitability of a polygon-based spatial representation for all phenomena has not been considered. The existing representation may be appropriate for objects having definite boundaries in the real world, but fields may be more appropriate for natural resource phenomena whose real world boundaries are transition zones of varying widths. Third, it has been largely forgotten that virtually every map is but one of many possible versions of a given geographic phenomenon. Furthermore, there is no guarantee that the existing map is "the best" version. Indeed, one cannot even really be sure that a "best" version exists for some phenomena. With the advent of electronic maps, it has become possible to create "maps" which consist of multiple versions or representations.

As DSTs have become more widely-available and more sophisticated, yet another error is being made by systems designers and users -- it is assumed that it is only the information explicitly present on the maps that is being analyzed by humans. In fact, when humans interpret maps, subjective assumptions about implicit data quality are made. A pedologist may infer considerable uncertainty for a class such as "steep, stony, ground" which may be a highly variable "catch-all" class. Whereas pedologists would have little confidence in the boundaries and characteristics recorded for this class and will adjust their analyses accordingly, such information is neither available nor exploitable in existing GIS.

We believe that the failure to include uncertainty in the fundamental design of a GIS leads to myriad problems for both system designers and end-users. Note the use the term "uncertainty" instead of "error." "Error" implies that there is some absolute truth from which one is deviating; "uncertainty" acknowledges the inherent impossibility of obtaining some sort of map absolute. We believe that GIS needs to be rethought to make uncertainty an inherent part of a map representation, not an a posteriori add-on as is often done. Furthermore, we need to accept that a machine will not be able to reproduce human beings' interpretive abilities, nor is it necessarily a desirable thing that it should do so. Instead, computers should be used to gather objective information about spatial data, which can be used for subjective human interpretation. To do so, GIS must be redesigned to depict spatial uncertainty in the underlying data structure rather than merely converting a paper map to digital form and using mathematically exact GIS tools on the result.

3: Remote sensing

3.1: Background

Remote sensing image analysis is one of the key data sources for GIS. Just as traditional thinking about maps has led to an acceptance of thematic class maps as the norm in GIS, similar thinking has led to the belief that remote sensing research should be oriented towards producing such maps. Early expectations were that remote sensing would provide GIS with accurate thematic maps via a minimum of human intervention. These unrealistic expectations quickly foundered because of a number of difficulties.
The first major problem encountered is that images are not uniform in appearance or quality, even when recorded with the same digital sensor. Atmospheric effects were discovered to affect images unevenly. Furthermore, the effects of sun and view-angle geometry (i.e., bi-directional reflectance function) mean that images taken at different times of the day, under different geometries, cannot easily be compared.

The second major difficulty is the accuracy of the classification process. Classification accuracies higher than about 80% are difficult to obtain, even when the classifiers are "trained" on a single image. Part of the problem is that the classification process involves adopting a model which relates spectral signatures to meaningful classes. Classification accuracy turns out to be as much a consequence of subjective decisions about what is "meaningful" as it is related to the distinguishability of different spectral signatures. If the human operator is excluded from the classification process, then classification accuracy will suffer.

Hence the production of precise thematic maps with a minimum of human interaction has been found to be unattainable on almost all counts: high accuracy is elusive, removing human interaction inappropriate, and the production of thematic maps error-prone. We believe the original goals of image analysis to be unworkable and in need of rethinking, beginning with the idea of thematic classification itself. The requirement of thematic classification is an arbitrary requirement, based on the use of thematic maps as input to GIS, which in turn is based on standard paper-map traditions of cartography. Thematic maps appear to be a necessary adjunct to human interpretation, a way of simplifying and synthesizing the information present in an image so that a human may draw important conclusions from the summarized information presented. However, simplifications involved in the production of thematic maps mean that a great deal of information is lost. It would be better to preserve this information within a GIS and convert it to a thematic map only when communicating spatial information to humans.

Doing so involves considerable rethinking of how remote sensing image analysis is performed. Today, much image analysis concerns "classification" and is oriented towards producing thematic maps directly from images. As a consequence, there is a paucity of image analysis tools which extract quantitative information directly from the images and make it available for further human analysis. Furthermore, because of the global nature of thematic maps, much of the existing work consists of attempting to produce globally valid estimates of image conditions and little work is being done on more localized image analysis.

One quantitative tool under development at Laval University illustrates this point. The tool consists of software for identifying and mapping individual tree crowns in very high resolution airborne imagery [7], [8], [5]. Counting tree crowns is the kind of mindless task at which the computer excels; it requires considerably less "training" than thematic classification and in fact constitutes a kind of quantitative mapping of forest structure. The technique also does not require a global approach to image analysis. Only tree crowns over a given region need be counted and mapped, unlike traditional thematic classification.

Efforts to extract other biophysical information from satellite imagery, such as the use of the NDVI to estimate primary production, may also constitute what has been named "quantitative remote sensing." Quantitative remote sensing consists of exploiting either empirically observed correlations between image tone or color and biophysical parameters of interest, or of modelling spectral information explicitly based on a more rigorous theoretical approach, such as via the use of forest or agricultural canopy models to predict light response of the canopy to incoming radiation. Again, many researchers have sought to classify images into homogeneous regions first before applying such models, but some lone voices advocate applying the models first and aggregating the results afterwards.

Finally, several researchers ([15], [16], [23] [25]) have pointed out that the results of image classification need not be just a thematic map. In fact, the image classification process generates a set of probabilities for each class for each pixel. Most classification methods simply select the most probable class and assign this to an output pixel. But there is no reason why a set of probability surfaces, one for each class, might not be output from the procedure. In this way, classification is transformed from a nominative process to a quantitative process, where more of the original information is preserved. The results of such a process may then be incorporated into GIS in a more quantitative form, provided GIS are suitably modified to accept such data.

Finally, another approach to image analysis which avoids arbitrary, global classification is the statistical exploration of image content via techniques such as multivariate image analysis (MIA) [9]. These techniques rely on an expert proposing and testing partial models of image components against the spectral content of real images. Classification may occur as a result of such analysis, but the classification now explicitly recognizes the subjective nature of the classification process, rather than hiding behind supposedly objective procedures. Because MIA is a tool that exemplifies a human redesign of DSTs such as we propose, slightly more detail is provided.

3.2: MIA

(The following discussion is based on the ideas presented in [10], [11], [12], [17], [18], [20].)

MIA is a generic approach to analysis of multivariate digital imagery developed in the fields of chemometrics, chemical image analysis and remote sensing. MIA is based on empirical data analysis principles rather than on specific statistical model assumptions. It distinguishes itself from other RS image analytical approaches in that
analysis is always started in the data analytical domain (the "score" space), instead of in the image domain as is generally the existing situation. This alternative starting point has many advantages, primarily that any classification training set in this score domain is more likely to be representative of a category of interest than a training set from traditional image domain classifiers. MIA is designed to cover both the dominant spectral as well as structural characteristics on an image and specifically covers exploratory image analysis and image classification. The main issues which distinguish MIA from traditionalist image analysis paradigms are:

1. Elimination of scene space fixation by MIA's initial score space analysis.
2. Proper consideration of the latent classification information -- i.e., MIA does not produce any classification other than that specifically determined by an analyst.

It will suffice here to present briefly the features of MIA most relevant to the concepts presented in this paper. The first step in MIA is to perform a principal components transform on the image. The user is then presented with a two-dimensional plot of any two principal components desired -- typically PC1 and PC2. To interpret this information, the analyst looks for clusters of observations on this score plot -- not on the image. The analyst identifies a portion of the scoreplot observations which appear to be clustered, and may then see the spatial location of the scoreplot observations in the image. Similarly, these observations may be "brushed" onto score plots of alternative principal components to determine if additional class discrimination is possible by examining another score plot. Note that there is no expectation whatsoever that the analyst should strive for a "best," or a "correct" class definition. On the contrary, MIA experience shows that usually it takes a number of such interdependent score space/scene space class refinements before the classification is acceptable. Note further that it is the domain-specific expert knowledge of what the image represents that determines a useful classification.

MIA is thus a highly analyst-driven approach. All the expert knowledge can and should be used in this interpretation stage before a final classification is produced. Furthermore MIA works in a strict sequential fashion for multiple class definitions. Only one interpretable class at a time is permitted -- never simultaneous multiple classifications. MIA also has the advantage of depicting the dominating image classes (i.e., classes consisting of large fractions of the total number of image pixels) in a damped, reciprocal fashion. That is, such classes are delineated in highly coherent groups in the score space and MIA will emphasize subtle class characteristics irrespective of their scene/image space distributions. Conversely, many subtle and/or spatially dispersed classes run the risk of being "swamped" by the major pixel classes in the image space. MIA specifically allows for individual, sequential class delineation and interpretation, often with the subtle, spatially dispersed classes being identified first. This is in contrast with existing image analysis philosophies in which an algorithm does not provide the analyst with total graphic access to both the structural data and the conventional scene/image domains at intermediate stages of analysis.

4: New map structures for GIS

The new ideas emerging from studies of uncertainty in maps and from quantitative image analysis and fuzzy classification in remote sensing appear to be converging on a single vision of spatial data and digital spatial technologies (DSTs). The characteristics of this vision include the following:

1) DSTs should not attempt to duplicate what a human being does well given a map, digital image, and/or photograph. Computers are good at manipulating numbers; humans are good at interpreting spatial information. Let the computer produce summary information that is difficult for humans to obtain -- e.g., spatial autocorrelation coefficients -- and let the humans do the interpretation itself.

2) The computer should facilitate the tenet "analyze, then classify" as opposed to the currently employed philosophy "classify, then analyze." Classification simplifies a system by discarding information, and therefore increases the error in subsequent analysis. A goal should be to keep data in as "raw" a form as possible in a computer, and only classify them when requested to do so by a user.

3) Computer-based spatial systems should enhance the information available to users and let the user make final decisions. This is in contrast to the existing philosophy that a user should be able to provide a system with all information necessary for a given analysis a priori, push a button, and receive the "optimal" or "correct" solution. Instead when classification is conducted, it should provide a wide range of information about different class probabilities to which a pixel or a region may be assigned and give the user more power to choose and modify the resulting classification at intermediate stages (as MIA does).

4) DSTs should recognize that a given map is only one possible version of a fuzzier "reality," and DSTs should explicitly recognize, model and display the fuzzy model underlying any given map version. This is true regardless of the particular technology used, whether raster or vector, image analysis or cartographic representation.

Indeed, our work and the work of others in the field complements our understanding of spatial uncertainty between nominative (themetic) and interval (variables defined on a continuous scale) spatial fields. Interval fields (as opposed to polygons with fixed boundaries) are characterized by surfaces of transition zones, within which no (or few) sharp boundaries may be present, and none explicitly present in the data structure. Thematic fields are characterized by homogenous regions with well-defined boundaries. Usually themetic or nominative fields are...
classified from interval fields. Studies in boundary uncertainty show, however, that each thematic map may be treated as one possible version of a set of thematic maps, each with somewhat different boundaries and hence as a single realization of a thematic map with fuzzy boundaries. The latter is very similar in structure to an interval field and, in principle, one may be converted into the other and back given appropriate additional information (such as the scale at which spatial autocorrelation is present).

Hence the kind of GIS we propose consists of either sets of probability or certainty surfaces (certainty surfaces are similar to probability surfaces, except that the strict coupling to probability theory has been relaxed and the values represent a relative scale of qualitative "certainty" rather than a true probability measure), or of thematic maps with sets of perturbed boundaries. Each set of surfaces is consistent with either one or several single-valued "maps" in the old sense (i.e., a paper map). The process of producing these maps should involve explicit recognition of the human role during the interpretation process, rather than hiding behind a screen of pretended "objectivity." Analysis procedures should be implemented which properly recognize the central importance of human interpreters and which benefit explicitly from their presence.

5: Uncertainty-based spatial representations

5.1: Background

The idea of incorporating uncertainty and fuzzy sets [31] into a GIS is not new. Indeed, others have proposed the idea for a number of years (see, for example, [1], [28], [29]). However, what is new here is an implementation of these ideas to demonstrate how geographical analysis can change and be more robust than existing analysis.

Whereas existing GIS use one map "laye1/theme/cov,, for each phenomenon being mapped, our proposed GIS will require one layer for each map class. Suppose that one has a soils map containing 10 classes. The "theme" Soils will be represented by 10 layers with each location in the data base having a value between 0 and 1 showing its likelihood of having a given class at that location. Perhaps the easiest way to think about this is with the raster data structure in which space is divided into a regular tessellation. In the proposed GIS, each raster cell would have an associated vector containing 10 values which sum to 1.0. This is akin to the fuzzy membership values (FMVs) described by [31] for fuzzy set theory; a surface of FMVs is referred to here as a fuzzy surface.

Note that though the raster data structure is put forth for purposes of visualization, this is not likely to be an optimal solution. Rasters suffer from the same problem as the vector data structure for the purposes outlined -- moving a small distance will not necessarily show a change in FMV as is the case in the real world. An acceptable alternative must be developed; preliminary work on the Voronoi diagram as a spatial data structure suggests that this may eventually be a suitable alternative [21].

5.2: Obtaining FMVs

One of the most difficult parts of spatial uncertainty work is estimating FMVs. Four approaches are discussed here.

Perhaps the most intuitive of these is the probabilities obtained in remote sensing before classification is done. In classification, some statistical process is conducted to estimate the probability of each pixel being a given category of interest. These probabilities are rarely examined, however, and a user is given only a final classification based on the probabilities. (One example of the examination of these may be seen in [30].) However, these may be easily used in the proposed GIS.

A second approach has been advanced by [22]. These researchers worked with a soils map in which initial uncertainty information was available in the map legend. That is, one soil type was described as "80% A with 20% inclusions of B." Given this information and making an assumption about the amount of spatial autocorrelation within this type, the original map was "perturbed" a number of times to better estimate the expectation that a given location was actually A or B. A similar approach (map perturbation) has also been examined elsewhere (see, for example, [13]).

A third approach has been advanced by [5] and [6]. In this approach, as with map perturbations, one assumes that any given map is merely one possible realization of a given phenomenon. This is akin to statistical sampling in which one wishes to have information on a population. However, obtaining population information is too costly so one samples the population and then makes inferences about the population from the sample. In the same way, one realization is but a sample of the population and additional samples increase the strength of the inferences made about the population -- ground-truth in this case. However, instead of the map perturbation approach of making assumptions about the nature of uncertainty and perturbing the map accordingly, one measures the uncertainty directly from multiple interpretations of the phenomenon. That is, n realizations of the map are produced by n cartographers giving a spatial sample of size n for the phenomenon. The boundaries on the n maps are then parameterized statistically in much the way that any sample is -- i.e., by the mean and variance. Work on synthetic images has proven fruitful and additional research is planned.

Finally, [25] has considered information on uncertainty which is implicit in a forest map. For example, the boundary between a clearcut and a mature forest can be assumed to be fairly precise -- i.e., multiple interpreters will put the line in roughly the same place. However, the line between forests of density 40% and 60%, respectively, will be much less precise. Therefore a forest
map of aggregate classes is "distilled" to lines of high certainty only -- i.e., between obviously different types -- and points which are the polygon cores of similar types. Thus instead of maintaining the lines between similar types, one makes the assumption that it is actually the polygon centre which has the greatest certainty of being the type mapped. Once one has distilled the map to contain only features of high certainty, interpolation is conducted to estimate the certainty of any given location being a particular type.

Regardless of from where the information comes, one will have a GIS composed of fuzzy surfaces on which the likelihood of having a given type for a given factor will have been estimated for all locations. The fuzzy operators of union ("OR") and intersection ("AND") whose Boolean equivalents are used in existing GIS for spatial analysis have been described ([31]). Thus one may still ask of the GIS questions such as "What soil type will I find here?" and "Show all the areas having a combination of Forest type 3 and Soil Type C." Instead of responding with a Boolean map, however, the uncertainty-based GIS will respond by showing membership grades (0 to 1) relative to the query posed. Users are then free to decide subjectively what level of uncertainty is acceptable given the problem statement.

6: Discussion

We believe that our philosophy relative to DSTs and uncertainty make our approach unique. Note that we specifically employ the term "uncertainty" rather than "error" because the latter implies the existence of an absolute obtainable truth against which estimates can be compared. Conversely "uncertainty" recognizes the fundamental subjectivity of the cartographic process.

Other researchers have also considered the problem of uncertainty, but only as an *a posteriori* add-on. We believe that if one wants computers to be of optimum use to humans, one must base DSTs on uncertainty as it is this information that humans subjectively and implicitly interpret and not only the geographic information as represented in maps of fixed points and polygons. Thus a GIS design should be based on uncertainty, and RS analysis should provide a GIS information on uncertainty rather than a final classified map. This is also a more desirable approach for analysis purposes as it is clear that the more one simplifies before conducting geographic analysis, the less meaningful will be the results of an analysis.

We also differ with others in the field by believing that it is not desirable to have as an ultimate goal the removal of the human being from the spatial data handling process. The inappropriate nature of this goal is exemplified by algorithmic attempts to eliminate sliver polygons that result from the overlay of two maps (see for example [32]). Sliver polygons are thought to originate because of map inaccuracy -- the difference between the map and the real world. In the case of surveyed boundaries, this is perhaps true. But in the case of a soils and forest map, this cannot possibly be the case as there is no such thing as an identifiable "truth." Thus much programming effort has been expended to find a way to eliminate sliver polygons whereas a human being would recognize them as indicators of spatial uncertainty. Thus it is more desirable to quantify uncertainty and use this information in a GIS rather than eliminate it simply to produce aesthetically pleasing, programming-compatible objects. An uncertainty-based approach to DSTs implicitly does this.

There are a number of advantages inherent in an uncertainty based spatial data approach. Continuous phenomena are maintained as continuous rather than as being composed of homogeneous polygons defined by widthless boundaries. It also takes advantage of the unique abilities of the computer and human being. The computer can do what it does best -- store, retrieve, and represent data -- and humans can do what they do best -- interpret spatial data to support human decision-making. An uncertainty approach also decreases the problem of information loss as the uncertainty-based surfaces are closer to the real-world situation than is a conventional polygon-based thematic map. Less obviously, the uncertainty-based approach provides for a better integration of satellite imagery into a GIS. Briefly, satellite images are classified in large measure to be compatible with GIS. But if a GIS no longer demands classified thematic maps, RS no longer must make the leap from complex summary statistics to highly simplified final map [3].

There are also negative factors associated with the use of an uncertainty-based representation. Related to this is that an uncertainty-based spatial representation will require more participation by an end user than is currently asked. This, however, is actually seen as a strength of the approach as it gives a user more control over analysis than is presently supported. A very real problem, however, is the amount of data storage required and/or the development of a suitable data structure. In the present study a raster representation was employed; problems with this representation were noted and an alternative needs to be found. The final problem is one that cannot be easily resolved: uncertainty-based spatial representations require a fundamental change in what humans consider to be "a map." Because the two-dimensional thematic map is so entrenched in geographic thinking and spatial analysis it will be difficult to change this concept rapidly.

Related to this is that an uncertainty approach may not be appropriate for certain phenomena. The authors come from a natural resources background in which polygons do not exist in the real world, though such phenomena have been mapped as such on paper maps historically. In other disciplines, phenomena are based on real world polygons. One need only consider political sub divisions of land to see one such phenomena. In such a case, a polygon-based model of space is clearly appropriate as each political unit -- a township, a county, etc. -- is subject to uniform laws and is usually considered in its entirety. For such
phenomena, there is no uncertainty associated with the
digital data base, although error may be present due to
digitizing, conversion of maps from one cartographic
projection to another, or the cartographic construction
process.

Finally, the central premise of this paper is that spatial
analysis should be left in the hands of users and that
computers should only provide an information-enhanced
user environment. In saying this, we consider "spatial
analysis" to be the interpretation of data and the ultimate
decision making based on such information whereas
others may consider "spatial analysis" to be techniques for
summarizing and gathering information about maps.
Regardless of the name, we strongly agree that the
computer can do the latter much more efficiently than
humans, and that there is a great need for the development
of new tools to extract information from maps. However,
we continue to believe that computers should merely make
such information available to users -- including
uncertainty information -- and leave its
analysis/interpretation to humans with their capacity to
incorporate implicit information into the decision-making
process.

7: Conclusions

We have presented a fundamentally different approach
to the storage, retrieval, and analysis of spatial information
in a computer than existing methods and have given
concrete examples of how this can be employed in both
GIS and RS. The adoption of the concepts discussed will
cause a major refocusing on what can be expected of
system designers and end-users; it will also benefit both
groups greatly. System designers will not have to feel that
they have "failed" if an implemented algorithm fails to
reproduce what a human being can produce. End-users
will gain more control over the analytical process at all
stages of the process. The ultimate benefit will accrue,
however, to all of those involved in computer-based
geography -- these individuals will be liberated from the
erroneous ideas that a map representation of "truth" can be
produced for all phenomena and/or that existing maps are
the best possible representations of various phenomena.

8: Literature Cited

Cartographica 19(2), Monograph 28, pp. 27-32.
Error in Categorical Maps. Unpublished Ph.D. Dissertation,
Univ. of Bristol, 261 p.
GIS: fundamental questions and new approaches.
Proceedings: 16th Canadian Symposium of Remote
Sensing, Sherbrooke, Quebec, pp. 873-878.
with fuzzy boundaries in geographic information systems.
Proceedings: 6th Symposium on Spatial Data Handling,
September, Edinburgh, Scotland, pp. 223-239.
variability in forestry data bases. Proceedings: IUFRO
Conference on the Spatial Accuracy of Natural Resource
Data Bases (in press), May, Williamsburg, Virginia.
uncertainty estimator for photo-interpretation based on a
model of the perceptual process. Photo. Eng. and Rem.
Sens. (in press).
density surfaces derived from high resolution digital image
analysis. Proceedings: GIS'93, Vancouver, February,
Volume 2, pp. 947-955.
Inventory Information: Extraction from High Resolution
Airborne Digital Images. Proceedings: 16th Canadian
Symposium of Remote Sensing, June, Sherbrooke, Quebec,
pp. 443-448.
image analysis (MIA). Chemometrics and the Intelligent
Laboratory 7:67-86.
multivariate image regression (MIR). Chemometrics and
Intelligent Laboratory Systems 14:357-374.
between higher-order data array configuration and problem
formulation in multivariate data analysis. Journal of
Multivariate image analysis (MIA) of hyper-spectral
satellite data. Proceedings: 8th Thematic Conference on
Geologic Remote Sensing, April, Denver, pp. 215-228.
Monte Carlo simulation. International Journal of
58:245-252.
membership of land cover classes in the suburban zone.
representation of vegetation continua from remotely sensed
data: an example from lowland heath. Photo. Eng. and Rem.
Sens. 58:221-225.
imagery: principal components regression for modelling
and prediction purposes. Journal of Chemometrics 5:97-111.
Image analysis in chemistry - II: multivariate image
information useful in chemistry? Journal of
Chemometrics 3:95-98.
K., 1989. Principal component analysis on multivariate
images. Chemometrics and Intelligent Laboratory Systems
5:209-220.
Voronoi approach. Proc.: Canadian Conference on GIS,
March, Ottawa, pp. 419-431.
and test of an error model for categorical data.


