UNIVERSITY *of* OTAGO
TE WHARE WĀNANGA O OTĀGO
DUNEDIN NEW ZEALAND

Special Issue:
GeoComputation ’96

A collection of papers authored by
members of the Information Science department
and presented at the 1st International Conference
on GeoComputation, Leeds, United Kingdom

The Information Science Discussion Paper Series

Number 96/25
December 1996
ISSN 1172-6024

University of Otago

Department of Information Science

The Department of Information Science is one of six departments that make up the Division of Commerce at the University of Otago. The department offers courses of study leading to a major in Information Science within the BCom, BA and BSc degrees. In addition to undergraduate teaching, the department is also strongly involved in postgraduate research programmes leading to MCom, MA, MSc and PhD degrees. Research projects in spatial information processing, connectionist-based information systems, software engineering and software development, information engineering and database, software metrics, distributed information systems, multimedia information systems and information systems security are particularly well supported.

The views expressed in this paper are not necessarily those of the department as a whole. The accuracy of the information presented in this paper is the sole responsibility of the authors.

Copyright

Copyright remains with the authors. Permission to copy for research or teaching purposes is granted on the condition that the authors and the Series are given due acknowledgment. Reproduction in any form for purposes other than research or teaching is forbidden unless prior written permission has been obtained from the authors.

Correspondence

This paper represents work to date and may not necessarily form the basis for the authors' final conclusions relating to this topic. It is likely, however, that the paper will appear in some form in a journal or in conference proceedings in the near future. The authors would be pleased to receive correspondence in connection with any of the issues raised in this paper, or for subsequent publication details. Please write directly to the authors at the address provided below. (Details of final journal/conference publication venues for these papers are also provided on the Department's publications web pages: http://www.otago.ac.nz/informationscience/pubs/publications.html). Any other correspondence concerning the Series should be sent to the DPS Coordinator.

Department of Information Science
University of Otago
P O Box 56
Dunedin
NEW ZEALAND

Fax: +64 3 479 8311
email: dps@infoscience.otago.ac.nz
www: http://www.otago.ac.nz/informationscience/

Contents

Editor's Note

The hard copy version of this discussion paper includes the following four additional papers, which are not available in electronic form:

George Benwell: *Spatial databases — Creative future concepts and use*

Sophie Cockcroft: *First experiences in implementing a spatial metadata repository*

Bruce McLennan, Martin Purvis and Christopher Robertson: *Wildlife population analysis with GIS: Conservation management of royal albatross*

Richard Pascoe: *Data sharing using the X.500 directory*

The hard copy version may be requested from the DPS Coordinator, at any of the addresses listed inside the front cover.

Incorporating A New Computational Reasoning Approach to Spatial Modelling

Alec Holt

Spatial Information Research Centre, Department of Information Science,
University of Otago, P.O. Box 56, Dunedin, New Zealand.
Aholt@commerce.otago.ac.nz, Phone : +064 03 479 8301. Fax +064 03 479 8311

Decision support systems, statistics and expert systems were some of the mainstay techniques used for modelling environmental phenomena. Now modelling systems utilise artificial intelligence (AI) techniques for the extra computational analysis they provide. Whilst operating in a toolbox environment and by adopting AI techniques, the geographic information system (GIS) modellers have greater options available for solving problems. This paper outlines a new approach in applying artificial intelligence techniques to solve spatial problems. The approach combines case-based reasoning (CBR) with geographic information systems and allows both techniques to be applied to solve spatial problems. More specifically this paper examines techniques applied to the problem of soil classification. Spatial cases are defined and analysed using the case-based reasoning techniques of retrieve, reuse, revise and retain. Once the structure of cases are defined a case base is compiled. When the case base is of sufficient size, the problem of soil classification is tested using this new approach. The problem is solved by searching the case base for another spatial phenomena similar to that which exists.

Then the knowledge from that searched case is used to formulate an answer to the problem. A comparison of the results obtained by this approach and a traditional method of soil classification is then undertaken. This paper also documents the saving data concept in translating from decision trees to CBR. The logistics of the problems that are characteristic of case-based reasoning systems are discussed, for example, how should the spatial domain of an environmental phenomena be best represented in a case base? What are the constraints of CBR, what data are lost, and what functions are gained? Finally, the following question is posed: "to what real world level can the environment be modelled using GIS and case-based reasoning techniques"?

Acknowledgements:

- Chris Price and the University of Wales for providing the CBR software called CASPIAN.
- Alan Hewitt and Manaaki Whenua -Landcare Research (Dunedin, New Zealand) for providing expert knowledge and data.

INTRODUCTION

Geographic information systems (GIS) are progressing towards systems which incorporate greater geocomputational functions. Three factors provide impetus behind this progression. First, spatial problems are inherently difficult to solve and the spatial information and modelling communities recognise that the lack of analytical and modelling functionality is a major deficiency of current GIS (Fischer & Nijkamp 1993). Second, as GIS databases mature users seek techniques which allow for further analysis (Burrough & Frank 1995). The third, which is a new concept, suggests that with the evolution of mature databases the vendors or data owners will move to develop applications and tools for their clients (Benwell 1996). These factors impel GIS progression towards a toolbox environment.

In addition GIS's progression towards a toolbox environment can be explained by its position as a platform for integrating various databases and systems. Decision support systems, expert systems, neural networks, fuzzy logic and connectionist systems, for example, are some of the many databases and systems which have been successfully coupled with a GIS. AI is also a good integrator and offers another avenue for providing geocomputation features. The various hybrids currently being implemented provide evidence of this. In a GIS-AI hybrid GIS could be used to represent and display the problem and solution while the AI techniques could be used to process the bulk of the problem solving. Some modelling systems use GIS or graphical display embedded in other platforms, including delphi and visual basic applications. An AI hybrid in comparison is also an excellent platform as some have graphical display techniques built into their systems. This is particularly applicable to medical applications.

As a result of these three factors a number of research paths can be taken to provide further geocomputational functions. In observing the diverse subjects of recent GIS conferences, analytical and geocomputation techniques are prolific. This paper identifies overlapping characteristics and diverse functions available in the following list of disciplines; statistics, cognitive science, knowledge acquisition, databases, case-based reasoning (CBR), inductive learning and knowledge discovery. In highlighting these diverse disciplines it will be seen that GIS modellers could progress their GIS systems to greater geocomputation levels by adopting some of the above techniques.

It is suggested that if CBR is incorporated with GIS it will further the geocomputational level of the GIS. CBR offers this GIS-AI Hybrid software reasoning from data, explanation, adaptation, extended generalisation techniques, inference making abilities, constraining a search to the solution template, generation, refinement, validation and maintenance of knowledge bases. These features help in planning, forecasting, diagnosis, design, decision making, problem solving and interpretation.

This paper will focus on these added features CBR offers the spatial reasoning system (SRS) (Holt & Benwell 1995a). With the potential of these added features, CBR aids a GIS and also suggests a new method for modelling spatial data.

Other research that has furthered the GIS progression towards a more geocomputational system include rule and knowledge-based approaches (Webster 1990; Smith & Yiang 1991; Skidmore *et al.* 1991), hybrid connection systems (Kasabov and Trifonov 1993), multiple criteria decision-making methods (Jankowski 1995) and a more innovative research approach where spatial reasoning is used to identify a given situation with other known typical scenarios (Williams 1995). These different analytical approaches are being coupled to form soft computing, for example neural networks with expert systems (Skidmore, *et al.* 1991). Case-based reasoning has been coupled with decision support systems (Burstein and Smith 1994). An interesting and important connection has seen the integration of case-based reasoning and neural networks. This method proposes a co-processing hybrid model for classification by coupling case-based reasoning and neural networks (Malek & Labbi 1995). Interest in hybrid systems is beneficial as effectual new systems for example, combinations such as neuro-fuzzy systems use the strengths of both neural networks and fuzzy systems to provide a more intelligent system. In GIS and modelling communities advances in AI systems through various combinations, bode well for strengthening the analytical capability of a GIS.

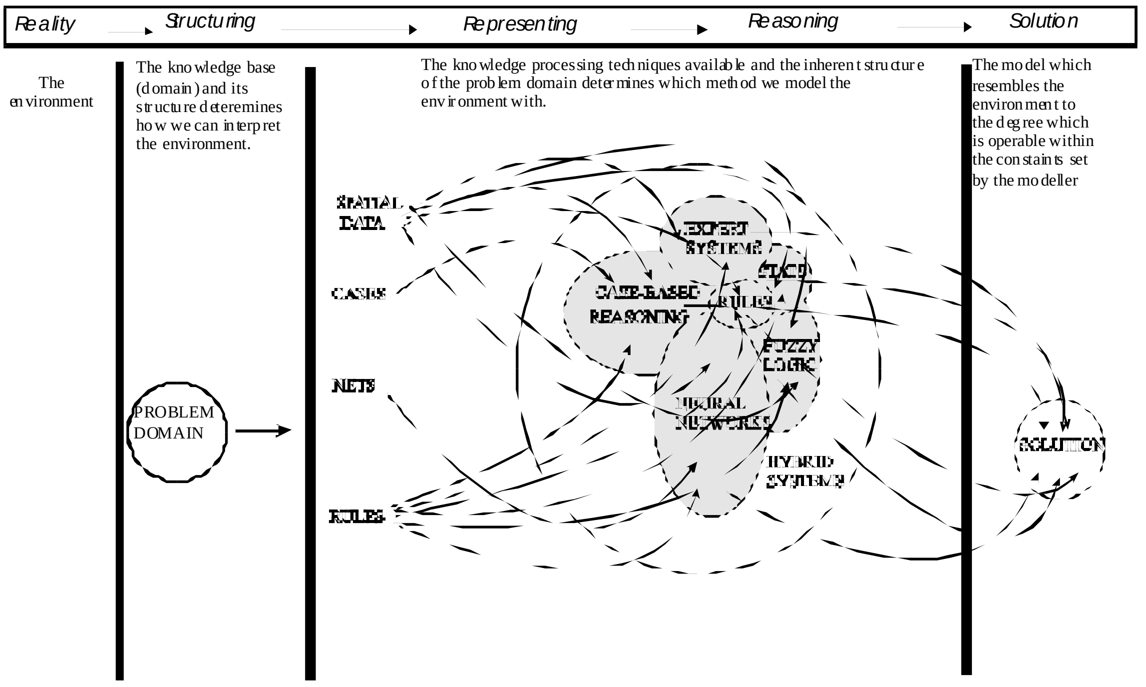
New techniques are illustrated and discussed in this paper. In order to determine “which techniques best suit which applications” an understanding of how the problem data can be best represented, and which of these working units the problem domain allows is necessary. If many working units are allowed the question of which technique should be adopted remains. A difference exists between the tasks AI techniques perform and those which can be best applied to spatial phenomena. Some of the techniques offered by AI techniques may not be required by the spatial modelling community. This paper focuses on CBR as one of these new tasks and documents the saving data concept in moving from decision trees to CBR. Which techniques best suit which problems? The following are limiting factors as specific data can be represented in certain ways and some techniques, such as GIS, in effect force data to be stored in a particular manner. The manner in which data is represented in effect dictates which analysis or reasoning techniques could be used. To reduce the impact of these limiting factors one could employ a large platform base. Having an AI-Hybrid, for example, would allow a number of analysis functions to be applied through one representation technique. A summary table of available techniques, their use, and for what type of spatial phenomena are they particularly suited (some techniques may be available but they are not chosen to be used) is currently under development (Holt in press). This has given rise to the notion of intelligent spatial information systems (Leung 1993; Laurini & Thompson 1992), which have adopted some, if not all, of the following statistical and analytical approaches to be discussed.

AI PROBLEM SOLVING PATHWAYS

Not all AI techniques are being fully utilised in the spatial information systems realm. AI usage for spatial problem solving has tended to be *low level processing, f*or example, in the classification of image patterns, primarily to complete images and to clean noisy data (Openshaw 1993; Kasabov & Trifonov 1993). AI techniques in comparison should be used to provide better decision support and more intelligent modelling systems. These systems could be used to solve spatial dilemmas which current GIS’s fail to do. Problems requiring further analytical processing could be specifically targeted by these systems. The GIS-AI hybrid provides *high level processing* and, therefore, increases the analytical processing ability of a GIS. The ability to reason may produce this higher level of processing. Therefore, it is proposed that a GIS-AI hybrid with the ability to reason should be developed (Holt and Benwell 1995a).

AI techniques available include, fuzzy logic (FL), neural networks (NN), case-based reasoning (CBR), genetic algorithms (GA), statistical models, knowledge and rule-based, fragmentation indexes and hybrid connection systems. A variety of pathways are therefore available to solve complex problems. An approximate map of the pathways of computational methods available for analysing spatial data has been drawn (figure 1). More than one AI pathway (some have overlapping functionality) could be used to solve a problem as each solves the problem differently according to their encapsulated functionality. One such pathway, which indicates how to solve complex spatial problems using a proposed GIS-CBR hybrid, is indicated in figure 1. In the course of the evolution of the spatial analytical toolbox it may be possible that a different GIS-hybrid is formed. Some AI techniques can increase their level of their performance if they are combined with other AI-techniques to provide hybrid systems. CBR-NN, FL-NN, for example, become more effective when combined.

Figure 1. Problem solving pathways



This may provide a comprehensive GIS where connection hybrids may become integrated with GIS. The GIS-CBR hybrid, which will be used to further the comprehensive nature of GIS, will be discussed. Two examples using the reasoning of a CBR to manipulate spatial data will be illustrated. The examples use CBR to evaluate test sites with previous spatial sites and amalgamate the spatial similarities of the test and previous sites to provide decision support to solve the spatial dilemma.

CASE-BASED REASONING

CBR is a general paradigm for reasoning from experience. It assumes a memory model for representing, indexing and organising past cases and a process model for retrieving and modifying old cases and assimilating new ones (Kolodner 1993).

It is important to define a case as they form the basic elements of CBR systems.

> *A case is a contextualised piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of a reasoner (*Kolodner 1993:13).

A case is a problem-solution pair. This emphasises the problem solving mechanism of CBR using the problem-solution pair to solve a similar problem. The two components of the pair are input and stored cases. *Input cases* are descriptions of specific problem situations. *Stored cases* encapsulate previous specific problem situations with solutions and outcomes. *Stored cases* contain a lesson and a specific context in which the lesson is applied. The context is used to determine when the lesson may apply again. These input cases are used to find other similar situations in a spatial manner (Leake 1995; Aha 1994).

Cases are examples which have occurred in reality or problems that have occurred and been solved (success and failures) by a problem solving mechanism. (Althoff *et al.* 1994)

The major components of a case include;

1. **Problem/situation description**: the state of the real world at the time the case was happening and, if appropriate, what problem needed to be solved at that time.
2. **Solution**: the stated or derived solution to the problem specified in the description or the reaction to its situation.
3. **Outcome**: the resulting state of the world when the solution was carried out.
4. **Extensions**: the context (justification) which links to other cases and the failures encountered.

(Kolodner 1993; Althoff *et al.* 1994)

These components of a case are the cogs of the case-based reasoning-cycle, or the solution and outcome components which make it possible to reuse, revise and retain cases.

More specifically, case-based reasoning is defined as;

> *a cyclical artificial intelligence problem solving paradigm that stresses reuse of solutions to similar problems, where solutions are maintained in a carefully indexed memory* (Aha 1994:3).

The above definition suggests that the components of CBR are representation, indexing and storing of cases for problem solving by retrieving, adapting, explaining, critiquing and the interpreting of previous situations. This process is used to create a solution to a problem using previous information. It is suggested that these components be added to GIS to complement its analytical functionality to build a *spatial reasoning system*. The proposed *spatial reasoning system* is designed to test the hypothesis that case-based *reasoning can be used to complement spatial analysis in GIS.*

Tasks that all CBR methods undertake include identifying the current problem situation, finding a past case similar to the new one, using that case to suggest a solution to the current problem, evaluating the proposed solution and updating the system by learning from experience.

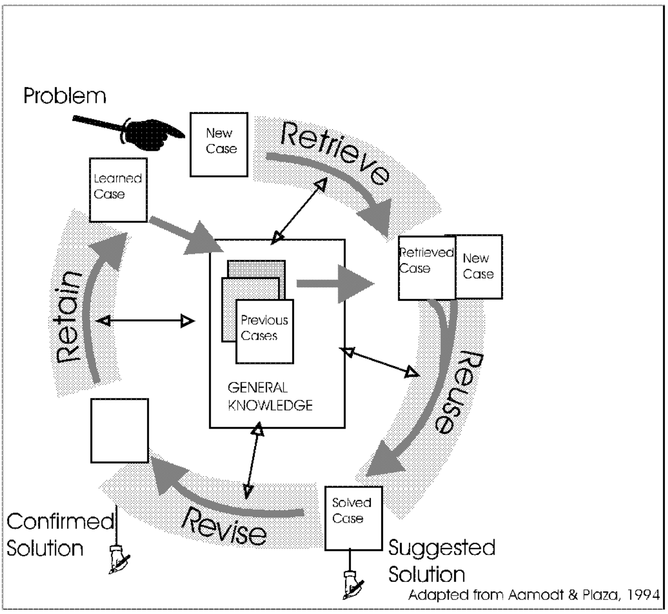
THE CASE-BASED REASONING-CYCLE

A CBR cycle may be described by the following four processes see figure 2 (Aamodt & Plaza, 1994; Kolodner 1993; Aha 1994; Leake 1995);

1. **retrieve** the most similar case(s),
2. **reuse** the information and knowledge in that case to solve the problem,
3. **revise** the proposed solution,

4. **retain** the parts of this experience likely to be useful for future problem solving.

Figure 2 CBR Cycle



Adapted from Aamodt & Plaza, 1994

CBR is a relatively new tool for solving spatial problems. There have been previous applications of CBR to solve spatial phenomena and environmental problems (Branting & Hastings 1994; Berger 1994; Jones & Roydhouse 1993; Kolodner 1993; Lekkas *et al.* 1994).

- Berger (1994) uses CBR to solve a medical problem using spatial data. The application is called *ROENTGEN* and is a CBR system that aids in planning radiation therapy for new patients based on geometrically similar previous patients.
- Branting and Hastings (1994) have developed a system called *CARMA* that uses model based reasoning and CBR to combat rangeland pests.

- Jones and Roydhouse (1993) produced a system called *MetVUW Workbench* which has been used for the retrieval and the display of historical meteorological data.
- Lekkas *et al.* (1994) developed a system called *AIRQUAP* which has been used to predict air pollution levels.
- Kolodner (1993) designed systems called ARCHIE. This is a system which aids an architect to design a new building. The cases represent knowledge about previous designs of buildings with similar specifications and situations.

THE SPATIAL REASONING SYSTEM (SRS)

Some definitions for reasoning in the GIS community include spatial cognition and the representation of knowledge (Hernandez 1993; Williams 1995). Frank (1996) defined reasoning as “the conceptualization of situations as space”. For the purpose of this research r*easoning* in this paper means the ability to reason; learning; thinking and the ability to draw on conclusions from facts (Holt 1996).

In problem solving a GIS uses raw data, not processed data, as there is no cycle and no facility to retain the solution. Therefore, there is no reuse of a previous solution or the process taken to derive that solution. This paper proposes a GIS-AI Hybrid called the spatial reasoning system. CBR offers this GIS-AI hybrid software an ability to reason, explain, adapt, extended generalisation techniques, inference making abilities, constraining a search to the solution template, generate, refine, validate and maintain knowledge bases. These features help in planning forecasting, diagnosis, design, decision making, problem solving and interpretation.

The Spatial Reasoning System (SRS) will eventually be used;

1. As a problem solving tool which has the ability to reuse previous similar spatial problems and their solutions to solve a current problem (Holt & Benwell 1996).
2. *As a problem solving tool which has the ability to reuse previous similar spatial problems and their solutions to solve a current problem,* with the added function of using a graphical interface to enter criteria.
3. As an exploratory spatial data analysis technique for data mining/trawling and pattern searching/matching (Holt in press).
4. As a new method to represent and store spatial data. Storing data as spatial cases, equivalent to object oriented languages, but having the added benefit of learning features.

With respect to the SRS this paper will describe the first option above. It is suggested that initially a GIS-CBR hybrid could be used to help in spatial decision making, spatial problem solving and spatial interpretation. This would produce a more intelligent GIS system as CBR could be used in the following ways;

1. The GIS-CBR hybrid is used *to facilitate searches* and answer the following questions. *Are there any other spatial phenomenon such as this? Identify extreme*

areas, evidence of trends, patterns or other variations. Is there clustering? If so, what attributes are associated with that phenomenon? In finding a similar spatial pattern a GIS is needed to display and store the data. CBR provides the functionality to find a similar pattern and, more importantly, to analyse its properties. These properties would extend from the obvious spatial pattern to other attributes associated with that spatial pattern. This type of functionality could be used for classification or in solving more complex problems using previous experiences.

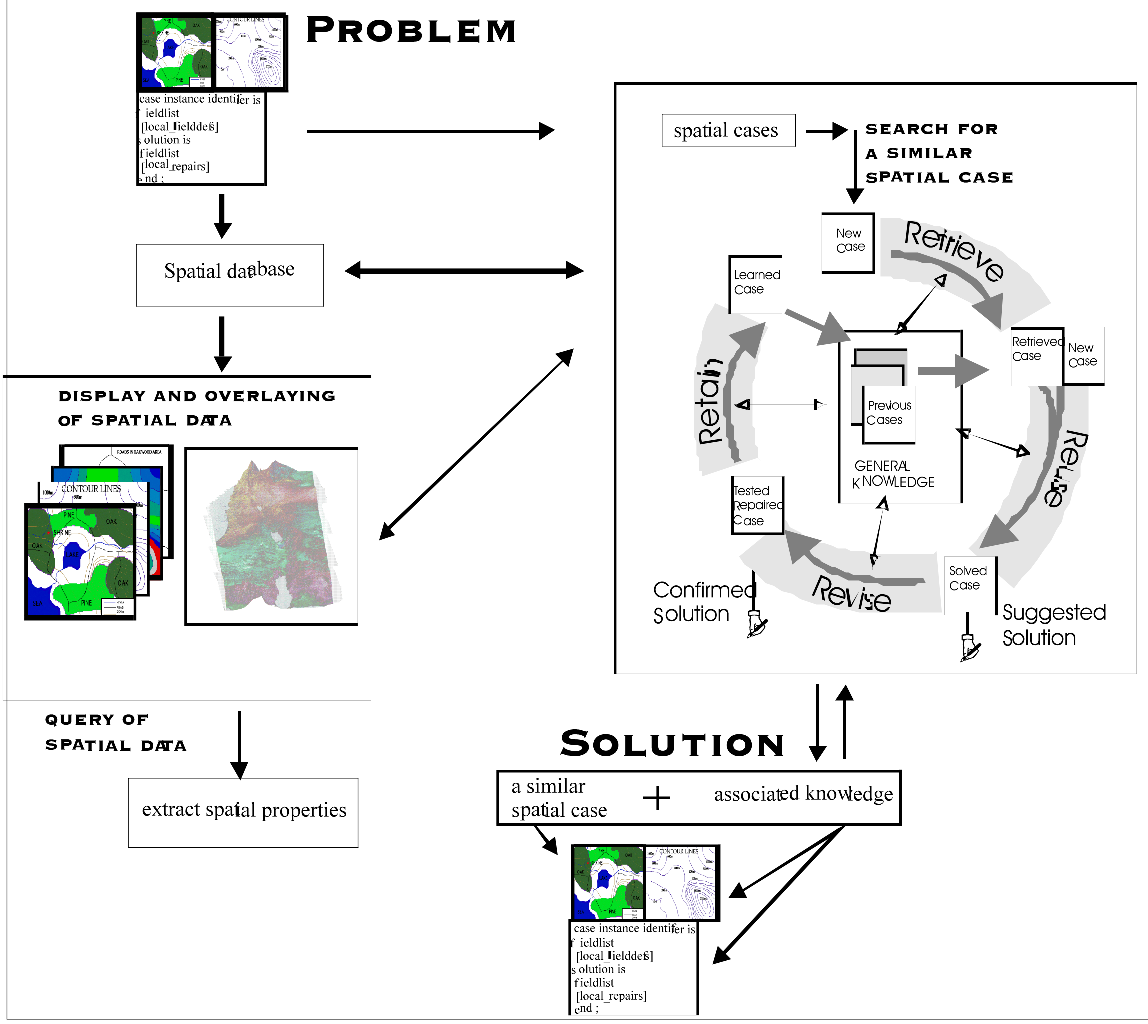
2. To make simulation possible. This is useful for the estimation and prediction of spatial phenomena including the *display of spatial-case distributional properties.*

3. Providing new opportunities in spatial analysis via information retrieval and pattern recognition. The following questions may be answered; Is there evidence of clustering in respect of specified sources or possible causes? What spatial associations exist between cases? Would a GIS-CBR model describe spatial relationships better?

4. The GIS-CBR hybrid is used *to facilitate queries* and answer the following questions; Which spatial phenomena have the following criteria? What attributes are associated with a spatial phenomena with these criteria?

The path taken to solve a problem using the GIS-CBR hybrid is shown in figure 3.

Figure 3 A Spatial Reasoning System



These criteria have spatial properties and the benefit of using a GIS for selecting slope, height and aspect include its ease of interpreting, manipulating and representing spatial data. A fully integrated GIS-CBR hybrid would have the ability to enter spatial data directly from digital maps and digital terrain models into a CBR. Once the data are entered the select action searches for a similar case, which is then displayed with any associated attributes. CBR provides the unique function of allowing further information related to the similar case to be used. These data can be saved as new cases if a decision is made based on previous cases. This function indicates CBR’s learning ability. This model has been tested and has provided satisfactory results (Holt & Benwell 1995a,b). The example of ZONATION, which uses soil to portray the spatial reasoning concept will be outlined. This application provides an interesting focus as soils have implicit spatial distribution properties.

ZONATION

As well as displaying CBR techniques in a SRS, it is also suggested that knowledge can be saved in a toolbox environment, for example, in this paper the knowledge saving transition is from decision-trees to CBR.

More specifically this research applies a combination of techniques to the problem of soil classification. The logistics of the problems that are characteristic of case-based reasoning systems are discussed. *ZONATION* employs Irvin *et al.* (1995) and Hewitt’s (1995) philosophy of using landforms to aid in the classification of soil series. An experiment is conducted based on data derived from Hewitt's series of tests.

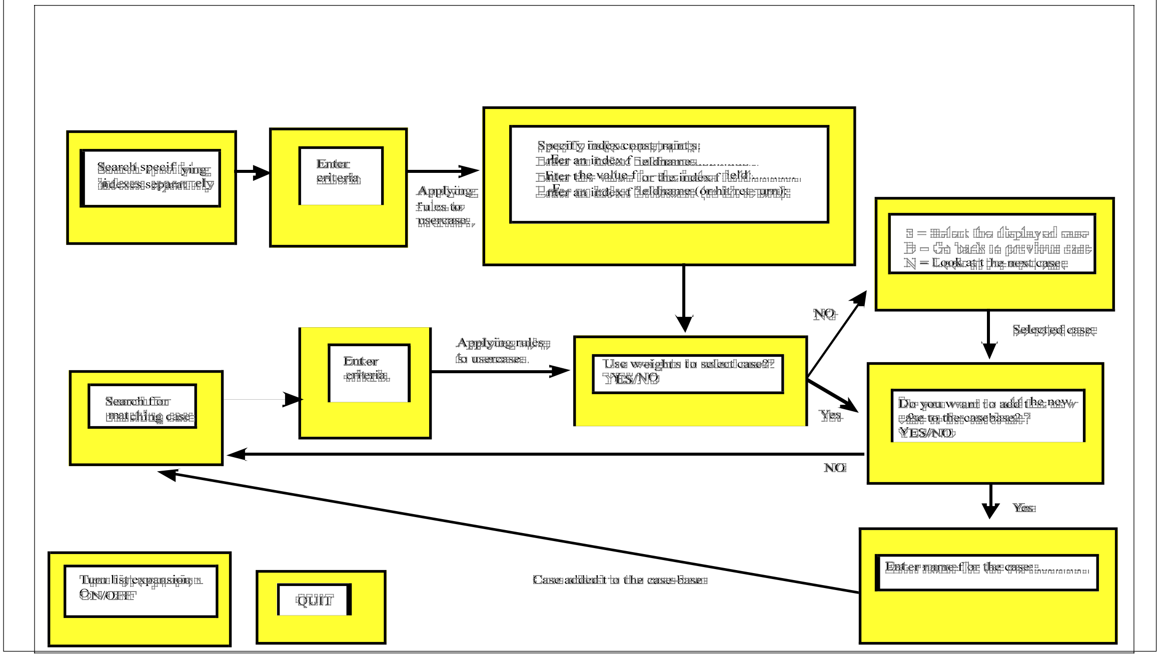
ZONATION employs a method for soil classification which utilises spatial information system techniques to classify individual pixels of a digital terrain model according to its membership in a landform class (Hewitt 1995; Irvin *et al.* 1995). These classes are determined by the natural clustering of the data in attribute space. Attributes central to this classification include, land element, slope, aspect, A_horizon texture, B_horizon texture, soil depth, A_thickness, dung, % gravel volume, % gravel weight, % carbon weight and % carbon volume (Hewitt 1995). Because these factors are also important to soil forming processes, soil classes should nest within landforms (Irvin *et al.* 1995). *ZONATION* adapts this philosophy by using the attributes and classes of Hewitt's criteria as fields and goals which are used to define case instances and, to store the attributes which are used to predict soil classification types of new zones.

The sequence for running *ZONATION* is as follows;

1. The user provides a case for comparison.
2. The program performs an index search and finds a subset of cases that match all the index constraints. The index constraints are taken from the field values provided by the user. The program searches the case base for the subset of cases that match all the index constraints exactly. ZONATION uses land_element as an index, therefore grouping all cases with the same land_element before making a selection.
3. If no cases match all the index constraints (for instances when there are only a few cases in the case base), the system prompts the user to search for different index values. If there are no cases which match all the index constraints, the user is informed and is prompted to enter new values for the index constraints. These may be made more general by specifying abstraction values or by specifying fewer constraints.

4. A case is selected from the subset. After the index search is completed the case matcher is invoked to scan the subset of cases to find the one with the highest weight value. This is selected and the repair rules are then applied.
5. Repairs are carried out on the selected case. On occasions additional information is requested after a case has been selected. Sometimes a repair rule can cause the current case to be abandoned and the selection process to begin again.

Figure 4. The Sequence For Running ZONATION



If the user is dissatisfied with the previous matching case(s) further cases may be examined. This is continued until they are satisfied with a matched case or until the user exhausts all possibilities.

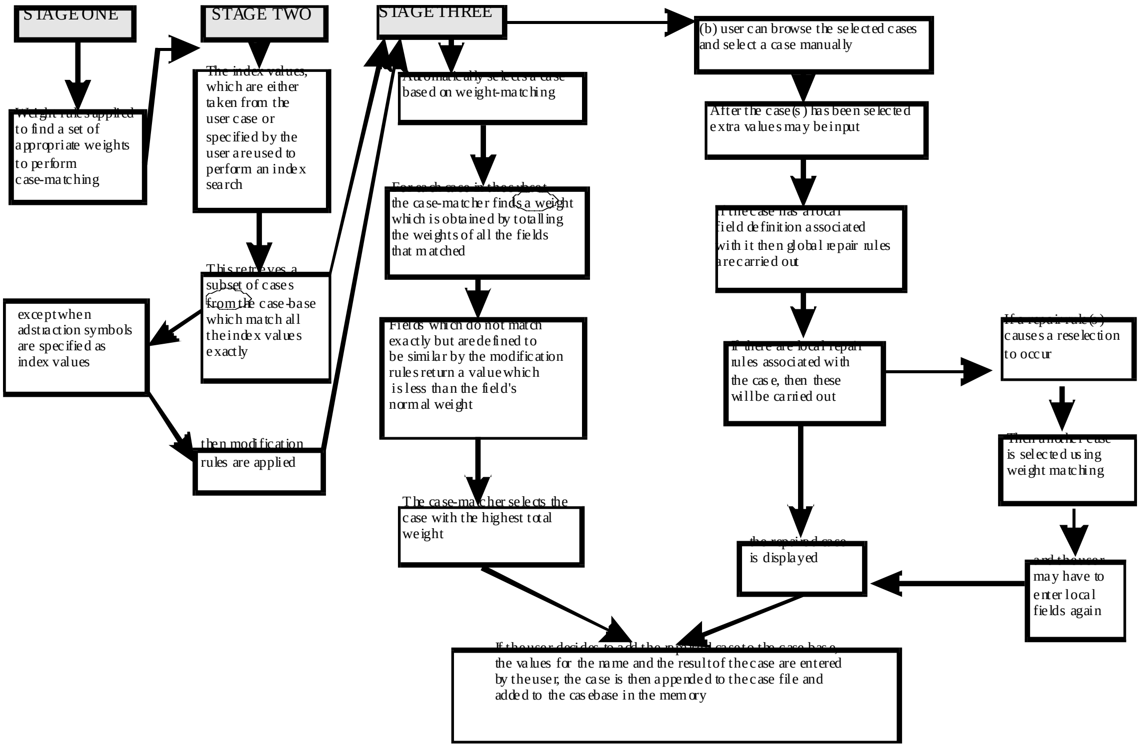
For the ZONATION case file blocks of code were written to define; *introduction*, *case definition*, *index definition*, *modification definition*, *weight rule definition*, *repair rule* *definition, and case instance*. The *introduction* block contains introductory text which is displayed when the program has finished checking the case file. The *case definition* sets the types and the weights of the problem fields that may appear in a case. The information in the *case definition* is used for checking input cases while the weights are used to aid the case-matching process. The *index definition* sets the fields used as indexes when searching for a matching case. A case base should have at least one field used as an index. The type of index field must be enumerated. The *weight rules definition* sets rules that may be applied to change the weights used for matching cases. The *modification definition* sets the modification rules and provides a means of specifying that certain symbols or numbers are similar. This is undertaken first for matching purposes and provides a means of specifying symbols as abstractions of others and second for making the search more general or for defining generalised cases. The *repair rule* definition contains the repair rules. These are used to modify the solution retrieved from the case-base making it more suitable for the current situation. Both the *modification definition* and the *repair rule definition* may be omitted. To be a complete CBR system, however, it should contain both. The last set of blocks are the case instances. These make up the case base. The case file must contain at least one case instance and will initially need to be seeded with many cases before it is operable.

The following three stages are indicative of the case matching process (figure 5);

1. Weight rules may be applied to find a set of appropriate weights for performing case-matching.

2. The index values, which are either taken from the user case or specified separately by the user, are used to perform an index search. This retrieves a subset of cases from the case base which match all the index values exactly (except when abstraction symbols are specified as index values, in the *modification rules*).
3. Once this list of cases has been retrieved the user can allow the program to automatically select a case. This is based on weight-matching. For each case in the subset the case-matcher finds a weight which is obtained by totalling the weights of all matched fields. Fields which do not match exactly, but are defined to be similar by the modification rules, return a value which is less than the field's normal weight. The case-matcher selects the case with the highest total weight. The user can browse through the selected cases and select a case manually.

Figure 5. The Stages Involved in the Case Matching Process for ZONATION.



The method of case-matching in ZONATION consists of two phases. First, the enumeration-type fields cited in the index block are used to select a sub-set of cases from the case-base. Second, a form of nearest-neighbour (other types include interquartile distance, discrimination networks and parallel retrieval (Leake 1995)) matching is used to select the best case from the subset.

The weights are not attached to the cases themselves. ZONATION parses through each case in the subset evaluating their weight. A record of the best matching cases are recorded. The importance of each field is defined in the case definition section. ZONATION uses internal rules (not to be confused with the weight rules block) to evaluate what proportion of the weight is returned for each field. If for example, the values match exactly then the full weight is returned. In comparison, if two enumeration symbols are similar then 0.75 of the field weight is returned. Strings have to match exactly or zero field weight is returned. During the parsing, of two lists of symbols and for example, if half of them match, then half of the field weight is returned.

After the case has been selected extra values may be input if the case has a local field definition associated with it and then the global repair rules (in the *repair rule definition*) are enforced. Furthermore, if there are *local repair rules* associated with the case then these will be enforced. If a repair rule causes a reselection to occur, another case is selected using weight matching and local fields may again need to be entered. The repaired case is displayed and the user is given the option of adding the repaired case to the case base. If the user adds the repaired case to the case base, the values for the name and result of the case are entered by the user the case is then appended to the case file and added to the case base.

The following table is an example of a case definition for the ZONATION case file;

Table 1. A case definition for ZONATION

```
field land_element type is (foot_sunny, shoulder_sunny, foot_shady, shoulder_shady, rolling_rise,
rolling_hollow, bluff_sunny, bluff_shady) weight is 20;
field slope type is number weight is 15;  ~degrees
field aspect type is number weight is 15; ~degree_magnetic
field soil_depth type is  number weight is 12;   ~cm
field B_tex type is (loamy_sand, coarse_sandy_loam, sandy_loam, sandy_clay_loam, loamy_silt,
silt_loam, loamy_clay, missing_data) weight is 10;
field A_tex type is (loamy_sand, coarse_sandy_loam, sandy_loam, sandy_clay_loam, loamy_silt,
silt_loam, loamy_clay) weight is 0;
field A_horizon_depth type is  number weight is 0;   ~cm
field bulk_density type is number weight is 0;   ~g/cm
field dung_freq type is number weight is 0;   ~1to10
field volume_of_gravel type is  number weight is 0;
field weight_of_gravel type is  number weight is 0;
field weight_of_carbon type is  number weight is 0;
field volume_of_carbon type is  number weight is 0;
end;
```

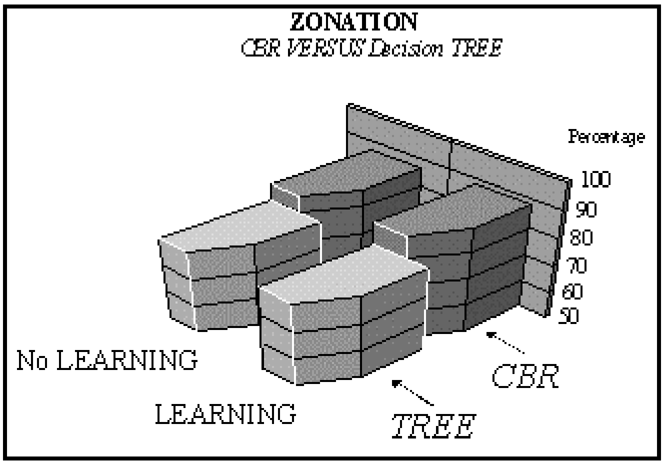
Spatial properties are defined as fields in the case file above. The case definition was used as a mechanism to process the spatial data input and the case instance was searched to fulfil the criteria of the case definition. Once a similar case instance was found it was possible to locate the similar case instance based on the fields of the case instance. Once these similar cases are located they can then be mapped and displayed. This shows the benefit a CBR provides, by increasing the GIS's analytical functionality and by adding an ability to learn. The traditional approach employed by soil scientists for soil classification uses an identification tree structure (Hewitt 1995). This example demonstrates the approach of using the GIS-CBR hybrid for soil classification.

A GIS-CBR hybrid has the ability to use explicit experiences to aid soil classification for new areas with similar spatial zones and attributes. These experiences (cases) enable the CBR to provide similar solutions (classification) to similar cases. Importantly, CBR differs from algorithms as no one solution is offered (as with an algorithm) and the user can choose a similar case from the solution set. Algorithms also need all criteria to be fulfilled while CBR allows some fields to be left blank without jeopardising the result. Results obtained through CBR improve after each new case is added to the case base. In comparison, an algorithm will predict the same answers and the associated error level each time it is used.

Explanation facilities are also easy to implement for a CBR system. The system will be used by those who feel comfortable with reasoning by analogy and who trust justifications that use data observations of past soil series incidents to support proposals, instead of chain rules that are triggered by *abstracted* threshold values. This explanation gives a system more chance of acceptance as ultimate responsibility for the decisions remains with the users. CBR systems can continuously incorporate new data in the form of cases and, in this way, adapt to long-term trends including soil degradation, loss and regeneration.

The success of the application was determined by comparing Hewitt’s model with *ZONATION* against 200 observed plots. A summary of the results are presented in the following graph. Hewitt's model while predicting 200 plots produced a 20% error. *ZONATION* was run twice. During the first prediction correct adapted cases were not saved as new cases and an error of 13.5% was produced. During the second prediction correct adapted cases were saved as new cases and the error level was reduced to 12%. The errors were then scrutinised for patterns and it was found that soils with a clay content, and with a low B_horizon value, were difficult to classify using either Hewitt's or the ZONATION model. Evaluating these errors will allow scientists to further quantify automatic soil classification problems. These patterns aid in strengthening the decision tree classification which Hewitt used and in facilitating extra rules to be added to ZONATION. It was found, for example, that soils which had sandy loam clay produced more errors when trying to predict its soil type using the decision tree. It is, therefore important to create a new branch in the tree if the soil type is sand clay loam. Alternatively, within the case base the addition of a trigger would allow for deeper case matching to try and eliminate the cause of these errors.

Table 1. Performance Graph



Whilst noting that ZONATION had a 12% error rate and evaluating the comparison between the observed values and zonation the following points should be considered;

- The performance is indicative to the software used.
- The diverseness and harshness of the central Otago environment. Landslides are frequent, especially during torrential rain, because of poor soil and the gradient of the terrain. During land slides both the source area and the region where the soil was deposited are temporally changed. Thus the Irvin (1995) and Hewitt (1995) proposal that certain soils must nest in certain landforms does not hold. Over time it is likely that the soil should again nest in its land form due to physical geomorphic process.
- A degree of uncertainty is associated with the initial control observations due user and instrument errors.
- Currently, this model is being applied to tourist spatial movements and the case history of aeroplane accidents (Higham *et al.* 1996).

CONCLUSION

Environmental problems are inherently complex. This research has proposed a novel method (GIS-CBR hybrid) to aid in modelling and solving of such problems. This CBR-GIS hybrid benefits from the functionality of both systems. Selecting spatial cases using GIS functionality (proximity, connectivity, adjacency) are such examples. Finding similar spatial cases with certain fields and in a certain proximity to or adjacency, connectivity to a spatial phenomena using CBR functionality. A case-based approach is beneficial as CBR systems have the ability to continually learn and evolve through the capturing and retainment of past experiences. This paper also illustrates the potential of AI in the spatial realm in recognising sets of patterns, predicting, providing decision support and simulating spatial phenomena. This includes recognising situations and structuring data to give spatial solutions to spatial problems. It is suggested that GIS would benefit from a CBR link. This is the first stage of *ZONATION* and an attempt to display the concept of CBR for soil modelling. It also attempts to indicate the possible extension of this concept to other environmental concerns such as geology, natural hazards and vegetation cover. The next stage of *ZONATION* should be more spatially oriented and incorporate graphical user input.

REFERENCES

Aamodt, A. and E. Plaza, 1994. Case-based Reasoning: Foundational Issues, Methodological Variations and System Approaches. *Artificial Intelligence Communications,* Vol.7 No.1.

Aha, D. W. 1994. Case-Based Reasoning, AIRIES 94, Workshop. Biloxi, MS.

Althoff, K. Wess, S. and M. Manago, 1994. Reasoning With Cases-Theory and Practice ECAI.

Benwell, G. L. 1996. A Land Speed Record? Some Management Issues to Address, The Tewkesbury Lecture, and Invited Keynote Address, *Proceedings of the Conference on Managing Geographic Information For Success*, Conference Proc, Department of Geomatics, The University of Melbourne, pp70-75.

Berger, J. 1994. Roentgen: Radiation Therapy and Case-based Reasoning. *In the Proceedings of the Conference on Artificial Intelligence Applications*, Proceedings. pp.171-177.

Branting, K. L. and J. D. Hastings, 1994. An Empirical Evaluation of Model-based Case Matching and Adaption. *American association for artificial intelligence, case-based reasoning workshop,* Proceedings. Seattle, Washington.

Burrough, P. A. and A. U. Frank, 1995. Concepts and Paradigms In Spatial Information: are Current Geographic Information Systems Truely Generic? Vol.9. No.2. pp.101-116.

Burstein, F. and H. Smith, 1994. Case-based Decision Making for Intelligent Decision Support. Department of Information Systems, University of Monash.

Fischer, M. M. and P. Nijkamp, 1993. Design and Use of Geographic Information Systems and Spatial Models. In: *Geographical information systems, spatial modelling, and policy evaluation*, Eds. M. M. Fischer and N. P. Springer Verlag, Berlin.

Frank, A. U. 1996. Qualitative Spatial Reasoning: Cardinal Directions as an Example *International Journal of Geographical Information Systems,* Vol.10. No.3. pp.269-290.

Hernandez, D. 1993. Reasoning with Qualitative Representations: Exploiting the Structure of Space. In Piera Carrete and Singh pages 493-502.

Hewitt, A. E. 1995. Soils of the Conroy Land System, Central Otago. Lincoln, Canterbury, New Zealand, Manaaki Whenua Press. Landcare Research Science Series.

Higham, E.C., Holt A. & G.W. Kearsley 1996. Tourist Flow Reasoning: The Spatial Similarities of Tourist Movements. In the Proceedings of the 8th Annual Colloquim of the Spatial Information Research Centre, Otago University, Dunedin, New Zealand, pp 69-78.

Holt, A. Benwell, G. L. & Kasabov, N. K. (In press) The Case-Based Reasoning (CBR) Approach, *Hand Book for Geostatistics and Spatial Data Analysis: Elements of Interpretation and Examples,* Editors Shibli, S. A. R. Michel, M. & D. Gregoire.

Holt, A. and G. L. Benwell, 1996. Case-based Reasoning and Spatial Analysis. *Journal of the Urban and Regional Information Systems Association,* Vol.8. No.1. pp 27-36.

Holt, A. 1996. Allowing the Environment to Model Itself. Environmental Perspectives, A Triannuual Newsletter published by the Environmental Policy and Management Reseach Centre. Issue 10.

Holt, A. and G. L. Benwell, 1995a. Spatial Reckoning using Case-based Reasoning. *8th Australian Artificial Intelligence conference incorporating the Workshop on AI in the environment,* Proceedings. Canberra, pp. 53-66.

Holt, A. and G. L. Benwell, 1995b. Case-based Reasoning with Spatial Data. *The 2nd New Zealand International two-stream conference on Artificial Neural Networks and Expert Systems,* Proceedings. Eds. N.K. Kasabov and G. Coghill. Dunedin, New Zealand, IEEE Computer Society Press, pp. 385-388.

Irvin, B. J. Ventura, S. J. & B.K. Slater, 1995. Landform Classification for Soil-landscape Studies. *Annual ESRI User Conference.*

Jankowski, P. 1995. Integrating Geographical Information Systems and Multiple Criteria Decision-making Methods. Vol.9. No.3. pp.251-273.

Jones, E. K. and A. Roydhouse, 1994. Spatial Representations of Meteorological Data for Intelligent Retrieval. *The Sixth Annual Colloquium of the Spatial Research Centre,* Proceedings. Eds. G.L. Benwell and N.C. Sutherland. Dunedin, New Zealand. pp.45-58.

Kasabov, N. and R. L. Trifonov, 1993. Using hybrid Connectionist Systems for Spatial Information Processing. *The Fifth Annual Colloquim of the Spatial Research Centre,* Proceedings. Eds. Benwell G. L. and N. C. Sutherland. Dunedin, New Zealand.pp85-96.

Kasabov, N. K. and D. Nikovski, 1992. Prognostic Expert Systems on a Hybrid Connectionist Environment. In: *Artificial intelligence- methodology, systems, applications*, Eds. B. Du Boulay and A. V. S. Gurev. Elsevier, Holland.

Kolodner, J. 1993. Case-Based Reasoning. San Mateo, Morgan Kaufmann Publishers.

Laurini, R. and D. Thompson, 1992. Spatial Knowledge. In: *Fundamentals of Spatial Information Systems,* Academic Press, London.

Leake D. B. 1995. Case-Base Reasoning: Issues, Methods and Technology, *A Tutorial for the First International Conference on Case-Based Reasoning*, Sesimbra, Portugal.

Lekkas, G. P. N. M. Avouris, & L.G. Viras, 1994. Case-Based Reasoning in Environmental Monitoring Applications. *Applied Artificial Intelligence An International Journal*, Vol.8. No.3. pp.359-376.

Leung, Y. and K. S. Leung, 1993. An Intelligent Expert System Shell for Knowledge-based Geographical Information Systems: 1. The Tools. *International Journal of Geographical Information Systems,* Vol.7. No.3. pp.189-199.

Malek, M. and A. Labbi, 1995. Integration of Case-based Reasoning and Neural Networks Approaches for Classification.

Openshaw, S. 1995. Human Systems Modelling as a New Grand Challenge Area in Science (Commentary). *Environment and Planning A,* Vol.27. pp.159-164.

Openshaw, S. 1993. Modelling Spatial Interaction Using a Neural Net. In: *Geographical information systems, spatial modelling, and policy evaluation,* Eds. M. M. Fischer and N. P. Springer Verlag, Berlin.

Skidmore, A. 1995. Neural Networks and GIS. *GIS User*, No.11 (May-July, 1995), pp.53-55.

Skidmore, A. K. Baang, J.S. and P. Luckananurug, 1992. Knowledge-based Methods in Remote Sensing and GIS. *The 6th Australasian Remote Sensing Conference,* Proceedings. Wellington, New Zealand, pp. 394-403.

Skidmore, A. K. Ryan, P. J. Dawes, W. Short, D. and E. O'Loughlin, 1991. Use of An Expert System to Map Forest Soils from a Geographical Information System. *International Journal of Geographical Information Systems,* Vol.5. No.4. pp.431-445.

Smith, T. R. and Yiang, J.E. 1991. Knowledge-based Approaches in GIS. In: *Geographic Information Systems*, Eds. D.J. Maguire, M. F. Goodchild and D. Rhind. Longman Scientific and Technical, Essex.

Webster, C. 1990. Rule-based Spatial Search. Vol.4. No.3. pp.241-259.

Williams, G. J. 1995. Templates for Spatial Reasoning in Responsive Geographic Information Systems. *International Journal of Geographical Information Systems,* Vol.9. No.2. pp.117-131.

GIS, Expert Systems and Interoperability

Linda Lilburne[1], George Benwell[2] & Roz Buick[3]

[1] Landcare Research
P. O. Box 69
Lincoln
New Zealand
Ph +64 (3) 325 6700
Fax +64 (3) 325 2418

[2] Department of Information Science
Otago University
P. O. Box 56
Dunedin
New Zealand

[3]Trimble Navigation
P. O. Box 8729
Christchurch
New Zealand

Email: lilburnel@landcare.cri.nz

1. Introduction

How should geographic information systems be developed? There is a strong demand from users for enhanced functionality and power. Vendors can and do respond to these demands. But where will this lead? Will the result be one all-embracing and all-conquering program or geographic information system (GIS)? A GIS could grow to incorporate all statistical functions, all visualisation techniques, all data management functions etc. It is possible to perceive a scenario in which GIS is developed to 'bloatware' proportions.

An alternative scenario is one in which a GIS is interfaced with other software systems. Embedding database bridges and other product-specific links, providing data import and export routines, and system calls are all ways of interfacing GIS with other systems. GIS vendors could opt to produce a 'linkware' GIS, interfaced to as many third party systems as possible.

Given these two alternatives to GIS development, an interesting set of questions arises. How far do vendors go with enhancing their systems compared with interfacing with third party systems? Is there a balance? Or do GIS users just keep calling for 'more', regardless of the solution set?

There is a balance. GIS is likely to be developed by being enhanced AND by being interfaced with third party software. In a way, this is a third developmental track leading to an increasingly functional GIS whose ability to interact with other systems is greatly improved. This interoperable GIS allows flexible combinations of system components while still providing a comprehensive range of spatial operations and analytical functions. Figure 1 depicts the three developmental tracks, leading to the 'bloated' GIS, the linked GIS, or the interoperable GIS in an environment in which systems can cooperate.

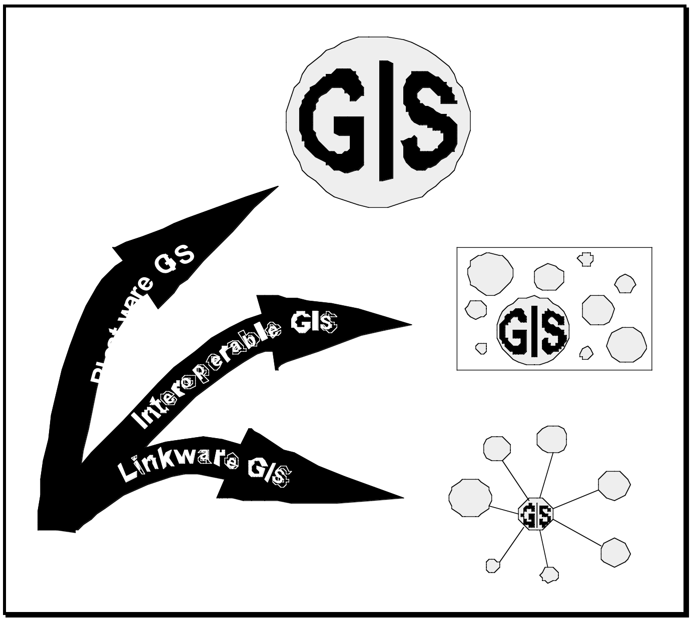


Figure 1 Three developmental tracks

Of these three developmental tracks, this paper presents an example of what can be achieved with the interoperable GIS. Expert systems are introduced along with the client/server and object-oriented paradigms. By using these paradigms, a generic, spatial, rule-based toolbox called SES (spatial expert shell) has been created. SES is described using examples and contrasted with other documented expert system – GIS linkages. But first integration is modelled in three dimensions to highlight the need for improvements in how GISs can interact with other systems.

2. The integration cube

Integration has been described by Fedra (1993), Goodchild (1992), and Nyerges (1992) from two perspectives: data and user interface. A more comprehensive model of integration can be described by adding a third perspective: interoperability. This relates to the ability of two systems, or processes, to cooperate with each other. Processes, which can exchange a variety of dynamically determined requests and information, have high interoperability. The interwoven nature of the requests allows the functions of each system to be highly integrated. More commonly however, integration is achieved by starting, executing and exiting a second process either after, or at some predefined point within, the first process. The flow of functionality between the systems is sequential. One process is followed by a second process, after which control may return to the first system. Interoperability is low.

By using the three dimensions of user interface, data and interoperability, integration can be portrayed as a cube (Figure 2). This gives a more complete picture of a coupling between two systems. The data axis progresses from no exchange of data, transfer of data, to sharing data. The user interface axis extends from a position of no link between two interfaces, to a trigger (e.g. a menu option or button) between two interfaces, to two interfaces which have a similar look and feel, through to a single interface. The interoperability axis ranges from static, sequential integrations with limited functionality to dynamic, interleaved linkages with full functionality. The degree of integration is represented by its position in the cube. The front bottom left corner represents no integration; the back top right corner of the cube represents a fully integrated system.

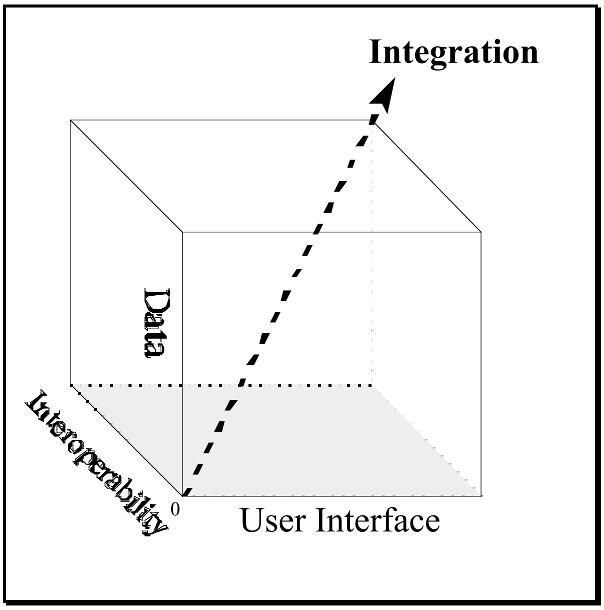


Figure 2: Integration cube

A sequentially integrated system can now be differentiated from a system in which the components of the integrated system can freely interact with each other. This model is defined in detail in Lilburne (1996). Lilburne also demonstrates how most integrations are positioned on the front plane of the cube, i.e. have low interoperability scores.

3. GIS and Expert Systems

An expert system is a "computer program designed to model the problem-solving ability of a human expert" (Durkin, 1994 p. 7). It comprises a knowledge base (KB), an inference engine, and working memory. The expert's domain knowledge is stored in the KB. The inference engine can process or reason with this knowledge to draw conclusions about a problem. The working memory contains the facts and deductions made in a session. Expert systems are used to identify, monitor, diagnose, predict, control, specify, design, configure, and plan (Jackson, 1990). An expert system can combine many types of knowledge including intuition, experience, qualitative beliefs, heuristics, empirical observations, and expert judgement.

The advantages of integrating a GIS with an expert system have been recognised by a number of authors (Burrough, 1986; Fedra, 1995; Fischer, 1994; Fisher et al., 1988; Lein, 1992; Leung, 1993; Robinson et al., 1987; Smith and Yiang, 1991; Zhu and Healey, 1992). These authors saw the expert system as having the potential to add intelligence to GIS tasks, e.g. map design, generalisation, automated name placement, feature extraction, spatial query. Fischer (1994) notes the increasing use of AI techniques to represent meta-data. Burrough (1992) envisions how an expert system could be an adviser or tutor for using a GIS. Domain knowledge represented in an expert system together with spatial data can provide a decision support environment in which users are guided by the integrated system towards a recommendation.

Table 1 contrasts the strengths and weaknesses typically observed in expert systems and GISs. This both shows strengths that are complementary and how some of the weaknesses of one system are matched by strengths in the other.

Expert system	GIS
qualitative	quantitative
imprecise data	precise data
uses symbols	uses geometric primitives eg point, line segment
integrates knowledge	integrates data
handles incomplete data and knowledge	does not easily handle incomplete data
suited to unstructured problems	suited to structured problems
no spatial capability	spatially capable
handles incomplete data and knowledge	does not easily handle incomplete data
does not cope well with lots of data	copes with large volumes of data
explanation facility	no explanation facility
can represent knowledge	is not designed to represent knowledge
can manage knowledge	can not easily manage knowledge
has inference engine	no inference or reasoning capability
opportunistic	algorithmic i.e. sequential
no mapping/graphing capability	variety of output maps/graphics
can not efficiently do arithmetic operations	can efficiently perform geometrical ops.

Table 1 Comparison of some expert system and GIS characteristics

The vision of those promoting expert system – GIS linkages was developed five or more years ago. To date, linkages have not fulfilled their vision. In part this is due to over-optimistic rhetoric about the capabilities of expert systems. It also relates to the level of interoperability between a GIS and an expert system.

For example, one application of an expert system – GIS linkage is to assist with solving or understanding environmental problems which are often very complex. GIS allows the real world to be modelled in its spatial context. Knowledge of real world objects and processes can be represented in the expert system. Sometimes the best knowledge available is in the form of heuristics. Combining spatial data with knowledge offers real opportunities in the management of our natural resources (Fedra, 1995). However, the real world is not a series of sequential processes; rather it is a complex tapestry of interactions between objects and processes. Hence the degree of interoperability between an expert system and a GIS will affect the ability of an integrated system to model the complexity of the real world. Recent technological advances offer new potential for closer interaction between systems, in particular the client/server and object-oriented technologies.

4. Technological paradigms

Client/server technology refers to the software that allows a process to receive messages from another process. These messages request services of the receiving system (the server). The service might be to perform a specified action or to return some information to the requesting system (the client). Both processes remain in memory concurrently, avoiding the loss of performance that occurs when loading a system into memory every time one of its functions is required. There is no limit to the number of requests, nor are there any restrictions on the types of requests that can be made. Both systems must conform to a common client/server protocol. There are incompatible client/server protocols, e.g. DDE, OLE (PC), RPC (UNIX), APPC (IBM).

Object-oriented (OO) technology is based on objects which have an identity, state and behaviour (Booch, 1994). ‘State’ refers to the data or values associated with the object at a particular point in time. ‘Behaviour’ is how an object acts and reacts. Key characteristics of OO technology that are important in a GIS – expert system link are abstraction, encapsulation, inheritance, and polymorphism. Abstraction is a simplified description of an object which encompasses all of its essential characteristics. Encapsulation, or information hiding, is the ability to hide implementation details from the user. Inheritance allows objects to inherit behaviour and state from parent objects. Polymorphism is the ability to redefine or override inherited behaviour from parent classes.

Abstracting the essential qualities that are useful for a given domain and grouping those objects with similar qualities is a powerful way of simplifying the computer representation of a problem. These essential qualities characterise the state and the behaviour of objects. Behaviour is defined by abstracting the essential operations that an object can perform and which can be performed on it.

A *vector* class in an expert system, defines the state and behaviour of GIS vector layers. The state of a *vector* object includes the name of the GIS layer, its description, default colour and its physical location. Behaviour is described in methods which are associated with the *vector* class. These methods are routines which encapsulate the GIS commands to draw, delete, create, modify and manipulate GIS layers. The commands are hidden from the user who does not need to know how GIS operations are implemented. Encapsulation also serves to hide the complexities of GIS data representation.

Objects inherit state and behaviour from parent classes. For example, a *roads* object, representing a GIS road network layer, inherits the state and behaviour of its parent *vector* class. Each object representing a GIS layer inherits a constructor method called *create.* This method encodes commands needed to select the required number of features from the layer, create a sub-object for each feature and transfer the GIS attributes to the sub-object. This allows the details of how to create objects that are sourced from a GIS to be associated with the appropriate object, rather than buried in a transformation routine written in a 3GL[1]. This facilitates maintenance of data exchange between the two systems.

Multiple inheritance is very useful in a GIS – expert system combination, as it allows an object to inherit behaviour from application specific classes or objects as well as GIS related classes. For example, a *road_segment* object inherits state and behaviour from the GIS *line* class, the specific *road_1* object(s) that the segment is part of and the *roads* vector layer object.

Polymorphism is useful to override inherited behaviour. For example, objects can inherit a *display* method which instructs the GIS to draw the object appropriately. Polymorphism allows a specialised *display* method to override the inherited one during processing.

5. SES design

SES (Spatial Expert Shell) integrates two commercial products: the GIS ARC/INFO (ESRI Inc., 1991), and the expert system shell, Smart Elements (Neuron Data Inc., 1994). ARC/INFO v7 includes some new commands (IAC[2]request, IACconnect, IACdisconnect, IACreturn) which create a framework for client/server communication with another process. Once a connection has been initialised, messages can be sent between the processes. Smart Elements is a combination of a hybrid frame, rule-based expert shell called Nexpert Object, and a GUI[3] developers kit called Open Interface. It has an Application Programming Interface (API) which allows C routines to access Smart Elements functions. SES is developed on a Solaris SUN Workstation platform which both ARC/INFO and Smart Elements support. Both systems use Sun's ONC-RPC client/server protocol. Smart Elements is the client and ARC/INFO is the server. A combination of C and ARC/INFO's macro language AML is used to develop the client/server interface between ARC/INFO and Smart Elements.

Essentially SES is a collection of spatial classes with predefined state and behaviour. The expert system shell is extended to include spatial classes. GIS elements (e.g. the display window, vector data, raster data) are modelled as classes under a top level class *gisObject* (Figure 3). Generalised classifications of spatial data are modelled as classes under the *gisObjectType* class. Appropriate spatial methods are associated with all of these classes. For example, polygon related state (area, perimeter) and behaviour (draw, adjacency, overlap) is defined in the slots and methods of the *polygon* class. An attribute of the state is stored in a 'slot'. Methods describe the operations or behaviour of an

---

[1] 3GL = Third generation language
[2] IAC = Inter-application communication
[3] GUI = Graphical user interface

object. *Vector* and *grid* classes have slots and methods that define vector and raster layer behaviour, e.g. how layers should be drawn, how features can be selected, and in which colour they should be drawn.

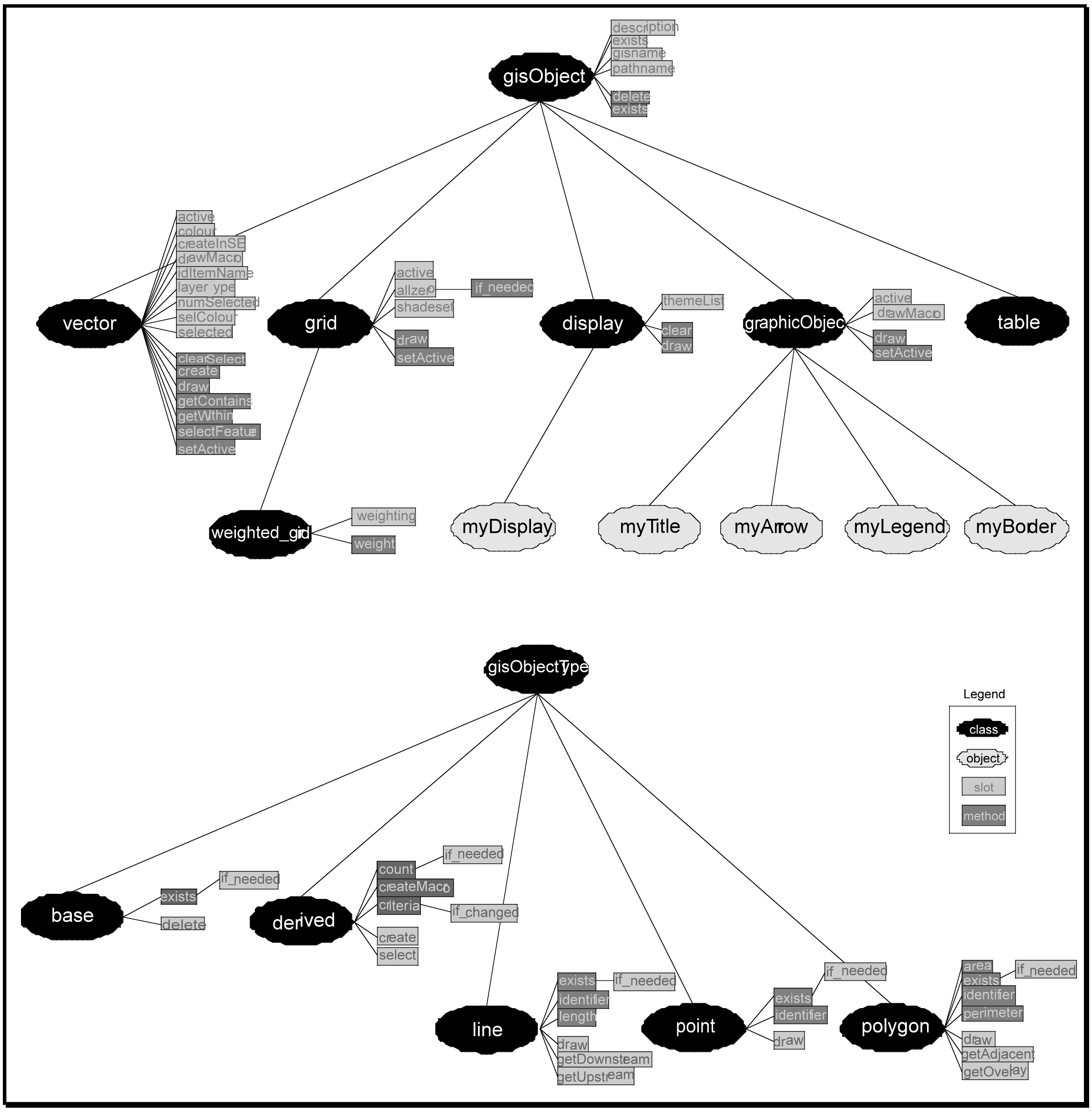


Figure 3 SES class diagram

A spatial expert system application can be developed by creating objects belonging to the spatial classes that make up SES. For example, in an agricultural application, there might be GIS layers of paddocks, tracks, streams, buildings, soil types, and a DTM. Objects would be created for each of these layers, linked to either the *vector* or the *grid* class. The objects inherit slots from these parent classes, e.g. the *paddocks* vector object inherits the *colour* slot and its default value “white” from the *vector* class. The *paddocks* object’s default *colour* value could be redefined to “yellow”. Methods are also inherited. For example, the *paddocks, tracks, streams,* and *buildings* objects all inherit the *selectFeature* method from the *vector* class. This enables an application expert system developer to transparently access a GIS function in which GIS features are selected from a map. For example, the DTM object inherits a raster specific *draw* method and default colour scheme from the parent *grid* class.

The application developer creates additional slots defining the domain state of each vector layer object. In the agricultural example, the *soils* object might have inheritable slots which represent the GIS attributes of the soil layer, e.g. soil name, soil code, pH level, soil depth. The *create* method associated with the *soils* object will dynamically create sub-objects of the *soils* object. Each sub-object, *soil_1, soil_2* etc., represents a single feature in the vector layer, e.g. a polygon. The sub-objects inherit domain slots, i.e. *soil_name, soil_depth* etc. which are populated by attribute values from the GIS. The sub-objects are linked to *gisObjectType* classes so that GIS state and behaviour can also be inherited. In Figure 4, sub-object *soil_1* inherits GIS slots and methods from the *polygon* class and inheritable domain slots and methods from the *soils* object. The *soils* object inherits GIS slots and methods from the *vector* class. These are not inherited by its sub-objects.

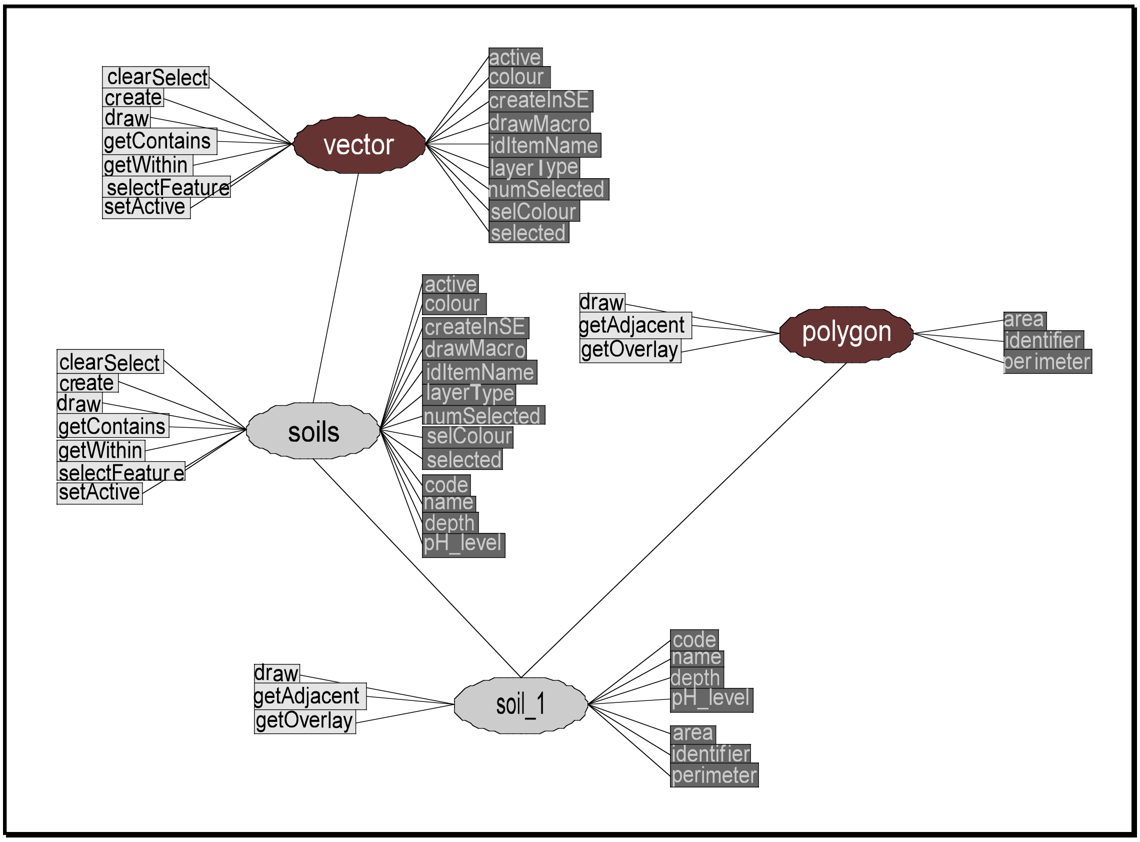


Figure 4 Soil example object diagram

The spatial methods attached to spatial classes access C procedures which perform error checking, creation and management of the information request which is sent to ARC/INFO, and manipulation of the reply. All spatial operations are executed in ARC/INFO. No locational data is manipulated in the expert system. Methods are used to access spatial relationships between objects. This enables the spatial context of an object to be referenced in heuristic knowledge. For example, a SES rule base could dynamically and transparently determine the truth of the following conditions:

- if forest stand is adjacent to road
- if the nearest firestation = "Albany"
- if bus route is longer than 5km
- if parcel is within the Christchurch District boundary
- if paddock contains sandstone
- if site is at least 200 m from the nearest water source
- if habitat is above 1000m

Each of the conditions requires the GIS to perform either an adjacency, nearest neighbour, route distance, contained within, overlay, buffer and/or a raster map algebra operation. These operations are encapsulated in generic spatial methods which are inherited by the domain objects (e.g. *parcel, forest_stand*) from spatial classes. In the first example, the *forest_stand* object sends its inherited *getAdjacent* method to the *roads* object.

In addition to the GIS providing information about spatial data and relationships, the expert system can dynamically access the full presentational and analytical functionality of a GIS. This is possible through provision of a global *gisExecute* method in Smart Elements. This method takes a string argument. First the method determines and substitutes GIS names and locations for any spatial objects referred to in the string. The string, now a valid GIS command that runs a macro, is passed to the ARC/INFO process which then executes it. The expert system can request the GIS to display the results of an inference session on a map, or complex spatial analyses such as pattern analysis, multivariate analysis, location/allocation, hydrological or viewshed analysis can be requested. The results might then be interpreted by the expert system, upon completion of the request.

Both raster objects and vector objects can be manipulated in a rule base. The Smart Elements network diagram in Figure 5 demonstrates how sample heuristics defining suitability of land for forestry can utilise combinations of raster and vector data. Forest suitability depends on the land's aspect, elevation, soil type, landuse capability code, and distance from roads. The *lri*[4] object refers to a soil polygon layer and the *track* object refers to a line layer. *Aspect, elevation, nw, high, nw_high, sunny, pukaki, luc*[5]*_6e_6c, luc_lessthaneq_5, mid, road_at_least_200* and *not_nw_high* objects all represent raster maps. The *road_at_least_200* object represents land that is at least 200m from a track. It is generated by executing its inherited *create* method with an argument string "> 200 track". This string is stored in the *criteria* slot. Similarly the raster *pukaki* object is generated by executing the *create* method with an appropriate argument which selects, then rasterises all the polygons in the lri GIS layer with a soil type of Pukaki, i.e. "lri soil cn Pukaki". The raster objects are weighted according to their relative importance. The raster objects and weights are passed to the GIS by the

---

[4] lri = land resource inventory which is a national GIS layer of soil and vegetation data
[5] luc = landuse capability code e.g. 6e

weight method which combines the GIS layers appropriately. Finally the *suit_forest* object is displayed by the GIS by sending an inherited *draw* method to itself.

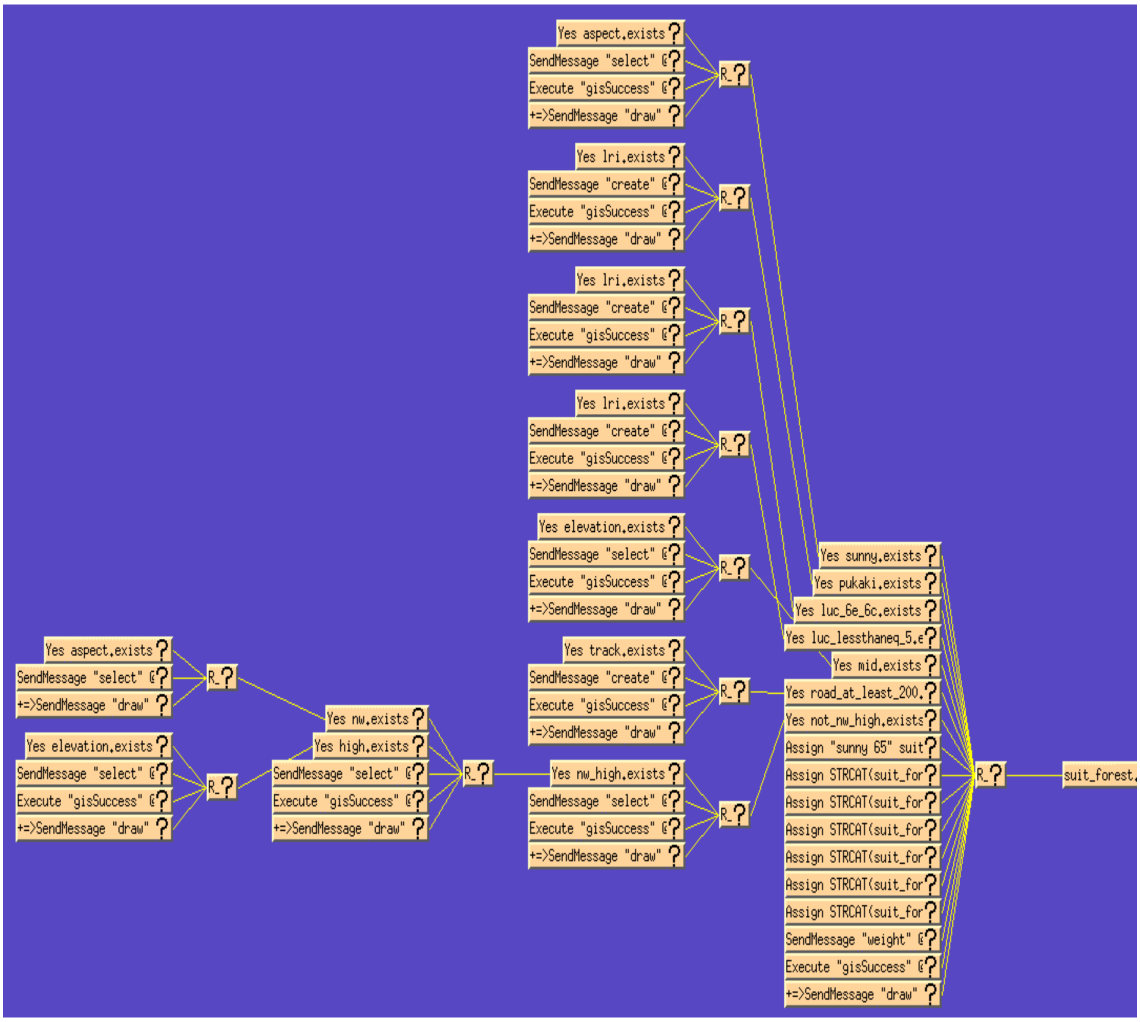


Figure 5 Network diagram modelling rule-based suitability of land for forestry

6. Discussion

The key advantage of an approach with high interoperability is flexibility. In SES, the GIS linkages required by a knowledge-based application are determined at runtime. They are not hardwired into the application. Spatial information is only accessed when it is required by the inference engine. The alternative approach is to calculate all the spatial relationships that might be required, and preload them as facts into the knowledge base, e.g. (Bleecker et al., 1990; Loh and Rykiel, 1992). This approach is only suitable for a limited range of spatial information, e.g. an adjacency matrix, stream connectivity.

The dynamic and generic nature of the linkage allows easy update of the domain knowledge base. The interwoven nature of the communication between the two processes allows the spatial and knowledge-based functions of the two systems to be effectively integrated. Figure 6 shows where SES is positioned in the integration cube.

There is very high interoperability between ARC/INFO and Smart Elements. Data is shared and there are two interfaces which can be designed to have a similar look and feel.

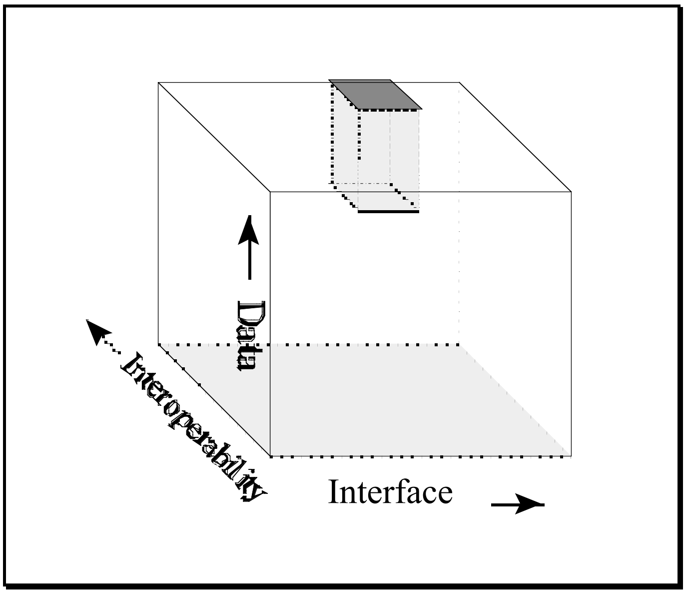


Figure 6 SES in the integration cube

This flexible linkage was achieved with minimal development effort as SES brings together the considerable power of the expert system shell Smart Elements and the GIS ARC/INFO. This was achieved by using a client/server approach.

The strengths of each system are maintained. SES maintains the computational power of a GIS by performing all geometric and raster map operations in the GIS. Raster maps are mapped to objects in Smart Elements which can be managed by the rule base. The expert system does not manipulate individual cell values. Knowledge is represented in Smart Elements using a symbolic representation (objects). This is more intuitive to work with than a quantitative representation (Sharma et al., 1994). Knowledge can be easily modified. A knowledge base developer can modify knowledge on the fly without needing to alter code or the linkage mechanism. Explanation capabilities provided in Smart Elements are available to the integrated system, as is the ability to handle incomplete data and opportunistic reasoning. GIS database management facilities can be used by the integrated system to access and manage data. The full set of GIS presentational functions are accessible in SES.

A limitation of SES is that while the intricacies of GIS operations are encapsulated in spatial classes, the use of spatial objects in rules is not completely transparent. For example, in a conditional statement:

If paddock = high

where *high* refers to the elevation, the syntax of the rule varies according to whether *high* is an aspatial attribute (i.e. “high”, “mod”or “low”), a raster map representation of cells with values over 1000m in an area of interest, or a fuzzy raster representation. The knowledge base developer must be aware of these differences and structure the rules accordingly.

The design of SES is portable to other combinations of layer-based GISs and hybrid object/rule-based expert system shells that support client/server functionality. For example, Smart Elements could use the DDE protocol to make requests of an ARCVIEW (ESRI Inc., 1994) server by passing appropriate AVENUEJ commands. For SES to be portable between GISs, spatial services must be standardised and made accessible to client processes by vendors. Requests can then be made in a format that is independent of which GIS product is the server. There is a move in this direction by the Open GIS Consortium (1996) which is developing a specification for distributed geoprocessing.

SES is based on a client/server relationship in which Smart Elements is the client and ARC/INFO the server. A reverse architecture is possible in which a client ARC/INFO process requests a Smart Elements server to run an inference session. For example, a rule base of diagnostic heuristics could be accessed from ARC/INFO. The user can be prompted for some information but all other information required by the rule base must be passed to Smart Elements before the rule(s) is processed. This is because a Smart Elements – ARC/INFO linkage is synchronous. In a synchronous linkage, the requesting process must wait for the server process to complete the request before it can continue processing. If in the diagnostic example, a GIS data value or relationship is required, this can not be dynamically accessed. The GIS is too busy waiting for its original request to finish, to service any requests made to it. The need to pass all potentially relevant facts to a rule base inhibits the flexibility of the system and is not very efficient. Spatial facts, especially distance related facts can quickly become quite extensive. An asynchronous link, in which processes do not have to wait for each other, would allow concurrent bi-directional requests.

Knowledge accessed by a client needs to be modularised into discrete reasoning segments. This is usually possible with meta-data and classification knowledge. For example, an expert system server can inform a GIS client of the validity of a value, or an expert system can be requested to fire a rule base to classify a series of data values supplied by a GIS client. The expert system-as-server/GIS-as-client architecture is appropriate in a GIS controlled application where one or more modular rule bases are accessed to perform a specified classification, diagnosis, data validation, and/or recommendation. Meta-knowledge is not appropriate in this architecture.

However, an expert system application is often a system in which many strands of knowledge are interwoven together. Control knowledge is integrated with domain knowledge, meta-data, and an intelligent interface. For example, a DSS[6] expert system application might combine an intelligent interface including appropriate question windows, recommendations, and explanations, with both knowledge about the process of determining a solution and knowledge about the domain objects. The expert system-as-client/GIS-as-server architecture (i.e. SES) is more suitable for a system in which

---

[6] Decision support system

multiple types of knowledge are integrated. GIS tasks can be modularised and thus made accessible to the expert system client which controls the integrated system.

Other integration approaches are defined by Lilburne (1996). One approach taken by some authors is to develop an in-house GIS and/or expert system which is very demanding of resources, e.g. Davis (1991), Lam (1993). Loose and merged/embedded approaches where the systems are sequentially executed have low interoperability and are usually inflexible, but these approaches require minimal resources to implement a link. An enhanced approach where one of the systems is extended to incorporate functions normally performed by the other system results in a subset of the total functionality available to SES. For example, basic GIS display functions were incorporated into an expert system based on PROLOG (Crossland, 1990). A tight approach usually requires access to the source code of the systems being integrated.

7. Conclusion

SES is a powerful, flexible, generic toolbox which can be used to represent knowledge from any domain in a spatial context. Use of the object-oriented and client/server paradigms have enabled a highly interoperable linkage, accessing the full range of functionality, to be established between two powerful systems. Minimal resources were required to achieve this.

SES supports our belief that moving towards a truly interoperable GIS is essential in today's interlinked world of distributed systems. An interoperable GIS allows effective use of external techniques and systems. In particular, an interoperable GIS expands the potential of combining knowledge with GIS. So, returning to the question of the balance between developing functionality versus interfaces, there is a need for vendors to follow the third developmental track and further improve the interoperability of GIS.

8. Acknowledgments

This research was funded by the Foundation for Research, Science and Technology, New Zealand. Rhys Gibson's cheerful assistance with Smart Elements is much appreciated, as are the comments by the reviewers: Paul Luckman, Megan Ogle-Mannering and Grant Hunter.

9. References

Open GIS Consortium, Inc. 1996. Open GIS. WWW ref: http://www.opengis.org.

Bleecker, M., Hutson, J. L., and Waltman, S. W. 1990. Mapping groundwater contamination potential using integrated simulation modeling and GIS. *Proceedings of Application of geographic information systems, simulation models, and knowledge-based systems for landuse management*, Blacksburg, VA.

Booch, G. 1994. Object-oriented analysis and design, The Benjamin/Cummings Publishing Company Inc., Redwood City, California.

Burrough, P. A. 1986. Principles of geographical information systems for land resources assessment, Oxford University Press.

Burrough, P. A. 1992. Development of intelligent geographical information system. *International journal of geographical information systems*, 6(1).

Crossland, M. D. 1990. HyrdoLOGIC - a prototype GIS expert system for examining an AI application in a GIS environment. *Proceedings of GIS/LIS '90*.

Davis, J. R., Cuddy, S. M., Laut, P., Goodspeed, M. J., and Whigham, P. A. 1991. Testing of soil moisture prediction model for army land managers. *Journal of irrigation and drainage engineering*, 117(4), p. 476-489.

Durkin, J. 1994. Expert systems: design and development, Macmillan Publishing.

Fedra, D. K. 1993. GIS and environmental modeling. *In* Environmental modeling with GIS, M. Goodchild, B. Parks, and L. Steyaert, eds., Oxford University Press.

Fedra, K. 1995. Decision support for natural resources management: models, GIS and expert systems. *AI Applications*, 9(3), p. 3-19.

Fischer, M. M. 1994. From conventional to knowledge-based geographic information systems. *Computers, environment and urban systems*, 18(4), p. 233-42.

Fisher, P. F., Mackaness, W. A., Peacegood, G., and Wilkinson, G. G. 1988. Artificial intelligence and expert systems in geodata processing. *Progress in physical geography*, 12, p. 371-388.

Goodchild, M., Haining, R., and Wise, S. 1992. Integrating GIS and spatial data analysis: problems and possibilities. *International journal of geographical information systems*, 6(5), p. 407-423.

Jackson, P. 1990. Introduction to expert systems, Addison-Wesley Publishing Co.

Lam, D. C. L., and Swayne, D. A. 1993. An expert system approach of integrating hydrological database, models and GIS: application of the RAISON System. *Proceedings of HydroGIS 93: Application of geographical information systems in hydrology and water resources*, Vienna, IAHS Publication 211, p. 23.

Lein, J. K. 1992. Modeling environmental impact using an expert-geographic information system. *Proceedings of GIS/LIS '92*, San Jose, CA, USA, p. 436-44.

Leung, Y. 1993. Towards the development of an intelligent spatial decision support system. *In* Geographic information systems, spatial modelling and policy evaluation, M. Fischer and P. Nijkamp, eds., Springer-Verlag.

Lilburne, L. 1996. The integration challenge. *Proceedings of SIRC '96*, Dunedin.

Loh, D. K., and Rykiel, E. J. 1992. Integrated resource management systems: coupling expert systems with database management and GIS. *Environmental management*, 16(2), p. 167-177.

Nyerges, T. L. 1992. Coupling GIS and spatial analytical models. *Proceedings of 5th international symposium on spatial data handling*, Charleston, South Carolina, p. 534-543.

Robinson, V. B., Frank, A. U., and Karimi, H. A. 1987. Expert systems for GIS in resource management. *AI Applications*, 1(1).

Sharma, J., Flewelling, D. M., and Egenhofer, M. J. 1994. A qualitative spatial reasoner. *Proceedings of 6th international symposium on spatial data handling*, p. 665-681.

Smith, T. R., and Yiang, J. 1991. Knowledge-based approaches in GIS. *In* Geographical information systems: principles and applications, D. J. Maguire, M. F. Goodchild, and D. W. Rhind, eds., Longman Group UK Ltd., Marlow, UK.

Zhu, X., and Healey, R. 1992. Towards intelligent spatial decision support: integrating geographical information systems and expert systems. *Proceedings of GIS/LIS '92*.

Environmental decisions with spatial process modelling

Samuel Mann
Spatial Information Research Centre
Department of Information Science
University of Otago, PO Box 56, Dunedin,
New Zealand
Phone: +64 3 479 8301
samann@commerce.otago.ac.nz

1 Abstract

This paper first describes the difficulties inherent in supporting a class of environmental problems, those involved in Regional Environmental Decision Making. A set of conceptual criteria are presented along with discussion on how the criteria might be approached. It is shown that a major obstacle is the need for a system that integrates components of Geographic Information Systems with process modelling functions. A new approach, Spatial Process Modelling is proposed. More detailed design criteria for this system are developed which are then used to develop a prototype system. The system is described and benefits and limitations discussed.

2 Introduction

This paper discusses the background, methods and initial testing of a Spatial Process Modelling System. While the system is intended for general use, a case study of a class of environmental problems is described for which current support is poor. These are the issues involved in regional environmental decision making (REDM).

2.1 Regional Environmental Decision Making

Two major forms of issue can be identified in environmental management. The first is that over the siting of particular noxious land uses, the “yes, but not in my backyard” problem. Such problems are suited to traditional decision making and to the array of tools currently available. Opposing groups present their graphs and maps in order to convince decision makers. These conflicts are also well served by bargaining techniques (computerised or not) where a residents group may be able to negotiate the establishment of a park on a corner of a large industrial development. Tools such as geographic information systems (GIS) may be of use here.

Regional Environmental Decision Making (REDM) forms a second group of issues. It is not well served, either by traditional decision making or by the tools used in these processes. These issues are those which are considered to be of regional significance and include decisions about grazing practices, the introduction of alien organisms, the impact of tourism and so on. Mann (1995) has shown that current REDM does not satisfy measures of good decision making (eg. Cleaves 1995).

2.2 Computer Support for Environmental Problems

The development of strategies for improving REDM is poorly supported by current decision support systems. Marr and Benwell (1996) showed that only 20% of local and regional government agencies were performing land use analysis with GIS despite 70% using GIS in the organisation. Data from a survey conducted by the author of NZ resource management practitioners (NZARM) shows that 42% were unsatisfied with current computer support for their work. The development of computer aided support is hampered by the inability to model and understand the complexities of specific environments (Mann and Benwell 1995). These include;

- uncertainty in goals (eg: ‘sustainability’, ‘biodiversity’)
- uncertainty in knowledge, often with conflicting evidence
- a dynamic environment
- importance of varying scale in time and space
- complex processes including feedback loops
- interaction of natural and human environments and the incorporation of policy and ‘precedent’ in decisions.

3 Tools to meet conceptual criteria

Mann (1996) took these perceived difficulties and developed a set of criteria required for development of useful tools for REDM. These criteria are shown in Table 1.

GIS applications are widely used in resource management situations and can be shown to improve performance. Crossland *et al.* (1995) showed that for well structured problems, GIS “makes positive contributions to decision maker performance, as evidenced by lower solution times and greater accuracy” (pg231). It is not clear though how the simple tasks measured by Crossland would generalise to the more complex area of REDM.

Table 1: Conceptual Criteria

1. a toolbox approach to allow flexibility while retaining powerful processing
2. spatially based analysis and display (including functions usually associated with geographic information systems, GIS)
3. process modelling functions
4. emphasis on facilitating human interaction and thinking for both workshop situation and single user
5. ease of use
6. facilitating requirements for organisational decision making.

The author's NZARM survey shows that of a wide range of methods used to represent the environment as part of a decision making process, personal conceptual models and spoken and written conceptual models were the most important. More formal methods, equations, textbook diagrams and programming, were least important. When asked to name the information sources used in their situations, again, a wide range of sources were identified but paper and digital maps were identified as the most frequently used. Decision support for REDM requires a mix of conceptual modelling and spatial analysis. Burrough and Frank (1995) however argued that "there is a large gap between the richness of the ways in which people can perceive and model spatial and temporal phenomena and the conceptual foundations of most commercial geographical information systems" (pg105). Unless this gap can be closed GIS may remain unsuitable for REDM support.

If serious attempts are to be made to manage the environment then appropriate information must be available as to the consequences of any actions. Because most environmental action is irreversible, modelling provides a powerful way of non-destructively testing actions and outcomes. There are many examples of simulation models being used in environmental decision making and both models and model*ling* can be seen to have several benefits.

A first benefit is in prediction. Simulation predictions (Kirchner 1994 prefers ‘system characterisations’) can assist decision makers in testing the potential outcomes of management alternatives. Often, as Pandey and Hardaker (1995 pg443) argued, “modelling results may bring the reality home to decision makers”. In this vein, Ball (1994) considers a good model as “one that is capable of reproducing the observed changes while producing insight into the dynamics of the system”.

A second benefit of modelling is in the generation of ‘new’ knowledge. Leimback (1994) describes a system of simulation based knowledge acquisition where, if in model development a piece of knowledge is missing, smaller model runs are initiated to populate the larger models.

A third, possibly more important benefit of modelling is the modelling process itself. Holling (1978) presented a process called Adaptive Environmental Assessment and Management (AEAM) that “uses the construction of dynamic models as an intellectual device to help people clarify issues, communicate effectively about shared concerns and explore objectively the construction of alternative policy options” (Walters 1986, pg43). Grayson *et al.* (1994) described workshops where the aim is to develop a simulation model which can be used to evaluate the effects of various management options. The model is the tangible outcome but “the modelling workshops are shown to be of primary importance” (pg245), indeed there is no method of saving results from the model. Key benefits include the atmosphere of ownership of the problem and solution, the use and critical appraisal of available information and that “capabilities and limitations of model are well understood by all...a different approach compared to many computer simulations of natural systems wherein the assumptions and limitations are often concealed rather than revealed” (pg 251).

4 Development of design criteria

4.1 Integration

There have been many calls for the integration of modelling and spatial information systems. So what are the obstacles to this development?

Data integration is identified by a number of authors as a recalcitrant factor. For example Kirchner (1994) argued that the shift from static models (including maps) to models where the structure is not predefined has implications for data requirements. Organisation and bookkeeping of inputs and outputs must take on increased importance. Coleman *et al.* (1994) suggested that “transfer of data between ecological models and GIS can be time consuming, and requires unique solutions for each model-GIS interaction” (pg398) but this is not a very practical solution. A further problem is the complexity and size of processing required by models containing spatial and temporal aspects. Coleman’s solution was to use distributed processing over several UNIX machines but this severely limits the portability of systems. At present this extended processing time may be a problem that we have to live with.

Despotakis *et al.* (1993) argued that GIS is primarily data driven while modelling is essentially process driven. This results in conflicting paradigms, GIS has space as the independent variable (ie. is frozen in time) while modelling is the converse (frozen in space). They concluded that that there is a “missing node between the field of GIS modelling and non-spatial modelling which would be necessary to integrate the benefits from both fields in a dynamic sense” (pg236). This view is supported by Nyerges (1992) who identified emphasis on content and structure for GIS and content and process for modelling.

A further problem is what Steyaert and Goodchild (1994) described as the cumbersome interfaces of GIS. This however is more of an issue than just the ‘front end’ of GIS and can be traced back to the previous discussion about emphasis. In describing the gap between what can be represented in GIS and the perception and modelling abilities of people, Burrough and Frank (1995) argued that the reductionist approach employed by GIS is sensible when dealing with simple, easy to combine abstract objects but is not suited to the natural world. A major conflict is the way in which dynamic processes are represented. An environmental scientist might produce a map of evaporation on a particular day but would find it harder to use a GIS to represent the processes involved in evaporation. This would be needed to predict evaporation or to model the effects on rainfall of a revegetation programme.

There is then a mismatch in interface, in data organisation and in general approach. Most of these problems come down to a difference in what is represented in terms of perceptions; pattern (data) or process (dynamics). This should not be an overwhelming obstacle, much of Geography could be described by the interaction of these themes (see Chapman 1979). The task then is to develop a GIS/model hybrid that best mixes the advantages of both pattern and process while overcoming the inherent conflicts. Such a system is referred to as Spatial Process Modelling (Figure 1). This figure can be seen as a metaphor for the problem at hand. Environmental problems occur in the real world. Both process modelling and spatial model are projections away from the real world but as previous discussion suggests, the two projections are in different directions. It is proposed that a modelling system which combines the two approaches will better approximate the 'real world' and allow improved decision making.

4.2 Design Criteria

This section examines approaches to modelling with the intention of developing design criteria for Spatial Process Modelling. Rather than characterise integration according to coupling intensity (*sensu* Goodchild and Wise 1992) or method of integration (*sensu* Lilburne 1996), the focus here is on the degree of flexibility in modelling.

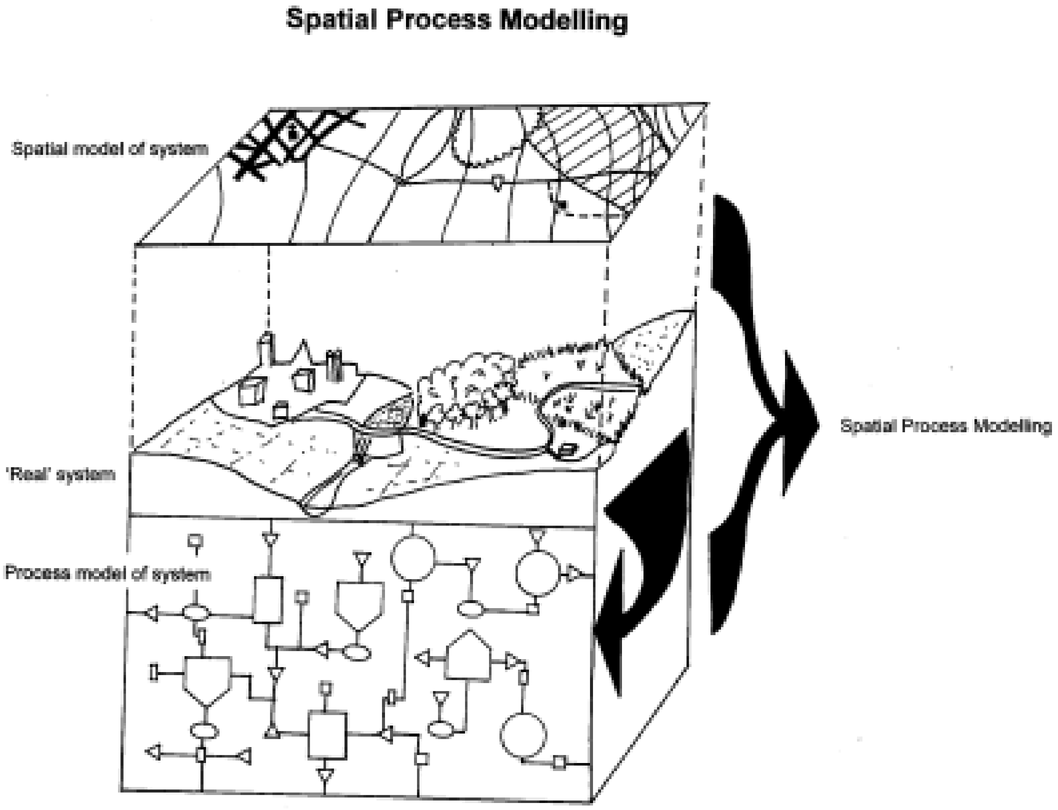


Figure 1: Spatial process modelling integrates GIS and process modelling

There are many successful models in environmental management. Most however are 'preformed', allowing the user to examine a limited number of options. These systems are really Decision support systems in the traditional sense of providing answers to predictable questions (eg Geraghty 1993). The rigid structure however eases the problems of data management and allows sophisticated display (Bishop 1995).

At the other end of the scale are approaches that rely on an ability to write code within a formal programming language. The AEAM workshops described by Grayson *et al.*

(1994) are examples of this. While participants are encouraged to feel an ownership of the model, they are in fact, separated from it, being forced to work through a programmer. Further, it is difficult for participants to get an overall feel of the model as it is implemented in code (at present QuickBasic, Grayson *pers comm* 22/1/96).

The ECO-LOGIC program of Robertson *et al.* (1991) aimed to provide assistance in these problems of comprehending model structure and writing code. Their approach was for the program to ask questions to build the model from key points, for example: ‘what do rabbits eat?’ and ‘how is growth represented?’. Templates are then used to write code (Prolog) that can be compiled and run as a simulation. The disadvantages of this approach are that it is domain specific and that information used to describe models is often too vague for code generation. The models also have to fit into a relatively predictable structure, a user is unable to deviate from the predefined structure.

Lowes and Walker (1995) described a high level language that allows a user to specify a model structure using pseudo-English. They found that this “high level, domain specific task or macro language” (pg2) was sufficient for representing the decision makers model of problems and enabled generation of a tool (code). However the task language was “too difficult” so they now favour diagrammatic approaches.

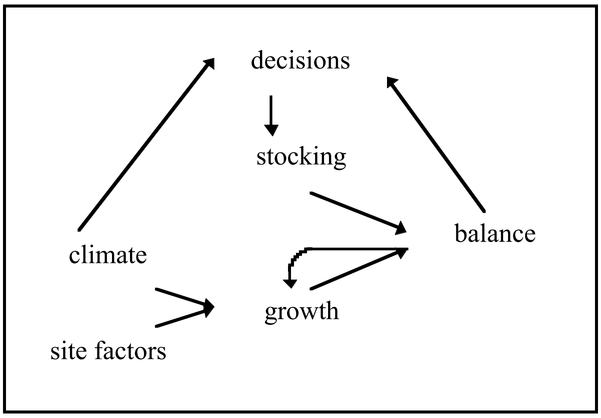


Figure 2: A simple model of a farming system

Diagrammatic approaches have an advantage in that they are closely aligned with conceptual diagrams. Such diagrams in varying forms can be found within the pages of textbooks on almost anything, from simple flow charts of water cycles to complex representations of agroecosystems. Figure 2 shows a model of a farming system. While very simple, this model could be used to facilitate discussion among stakeholders. Farmers for example might wish to add a link between the land (site factors) and farming decisions.

With computerised diagrammatic techniques, diagrams such as Figure 2 can become much more powerful. Dynamic modelling packages such Stella (HPS 1990) facilitate the development of system models through a toolbox library system of model components. Although the model appears sketched, the package allows the users to run simulations to examine the effects of varying parameters and/or model structure. The user, then, is separated from the business of generating code, but still has access to a powerful and flexible modelling system. The problem though, is that these modelling systems are presently aspatial.

A list of design criteria is given in Table 2. It is argued that development based on these criteria while still meeting the broader objectives of Table 1 will make a useful support system for REDM.

Table 2: Design Criteria

- **Visual**: The system must be easy to use with a Graphical user interface.
- **Interactive**: The user must be able to develop scenarios interactively.
- **System dynamic modelling**. There are three components to this criteria.

1. the model must be able to support a systems approach
2. the system and the systems it represents are dynamic which requires flexibility in design and use
3. the use of modelling. The system must support the development of models rather than fixed models.

- **Spatial:** The system is concerned with the management of the environment. The support of environmental decisions requires an ability for spatial manipulation and presentation.
- **Model database**: A database for model structure (as distinct from the environmental data).
- **Integrated**: Models are not inherently spatial, in fact most are aspatial and may be represented in modelling packages such as Stella and Extend. The environment is spatial and is represented within geographic information systems. The linking of these two approaches should be integrated by the system such that the model components include the spatial objects. This integration is a fundamental criterion.
- **Generic**: The system should be generic. That is, the system should operate as a toolbox independent of the domain.
- **Portable and PC based**

5 Spatial Process Modelling System

5.1 Overview

This section describes a prototype implementation of a Spatial Process Modelling System. The SPMS essentially inserts map objects directly into a Stella-like visual modelling toolbox. This step has resulted in a powerful yet flexible tool but has also raised many methodological issues. First an overview of the system is presented from the users perspective. This is followed by a description of the structure of the system. Benefits and limitations of the current implementation are also highlighted.

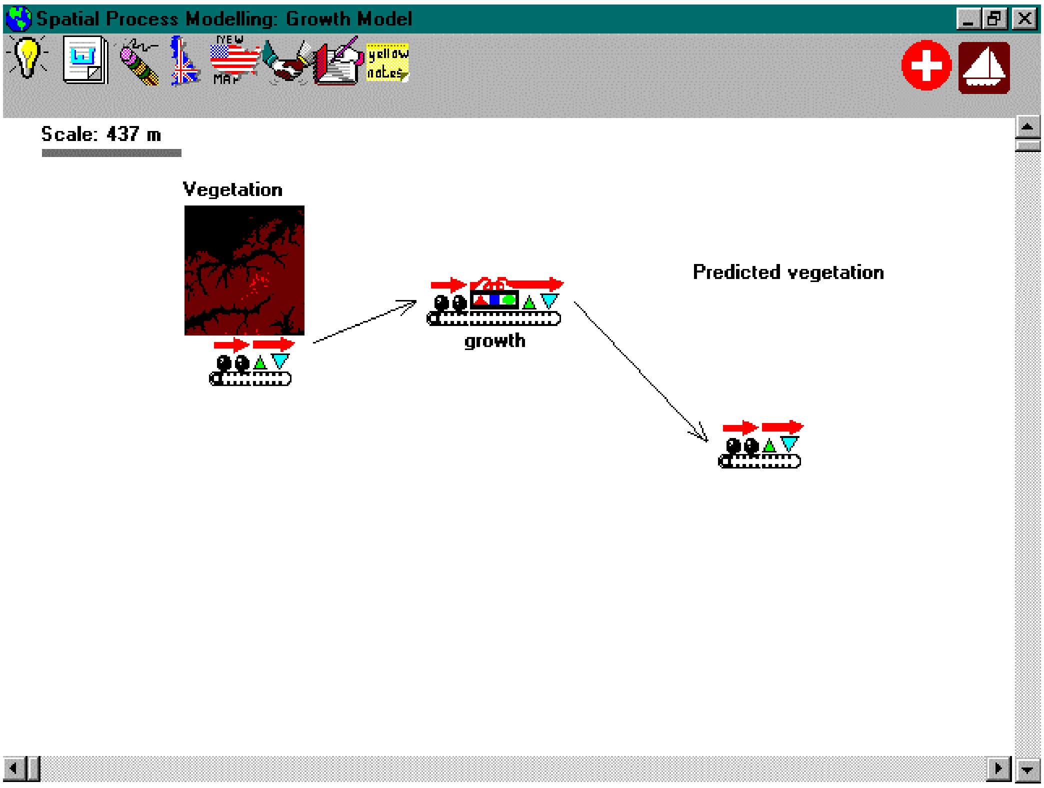


Figure 3: Growth model under development

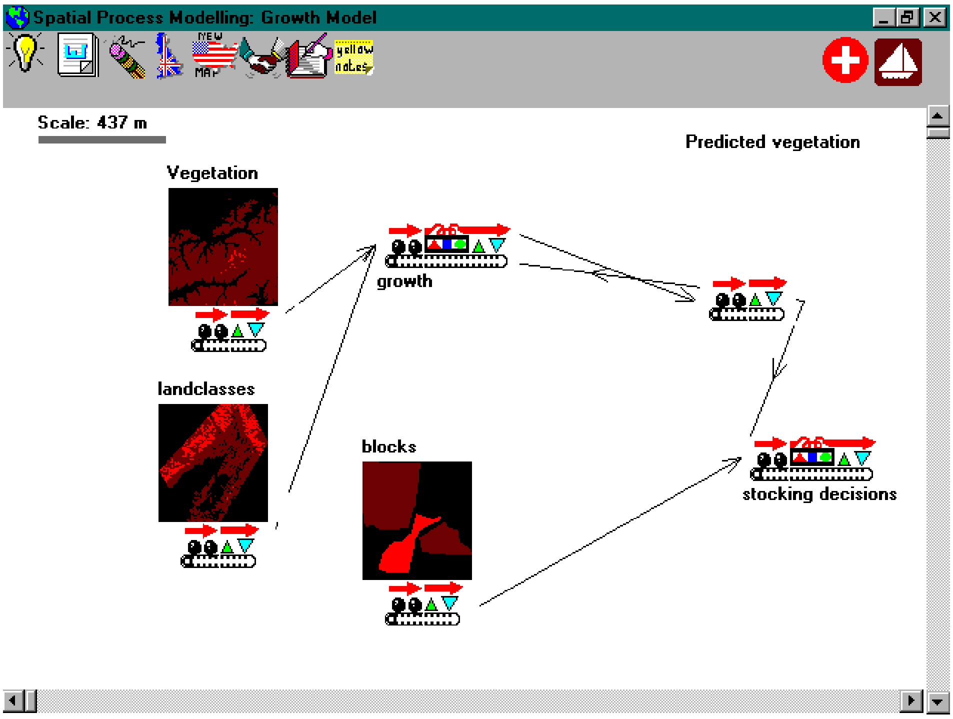
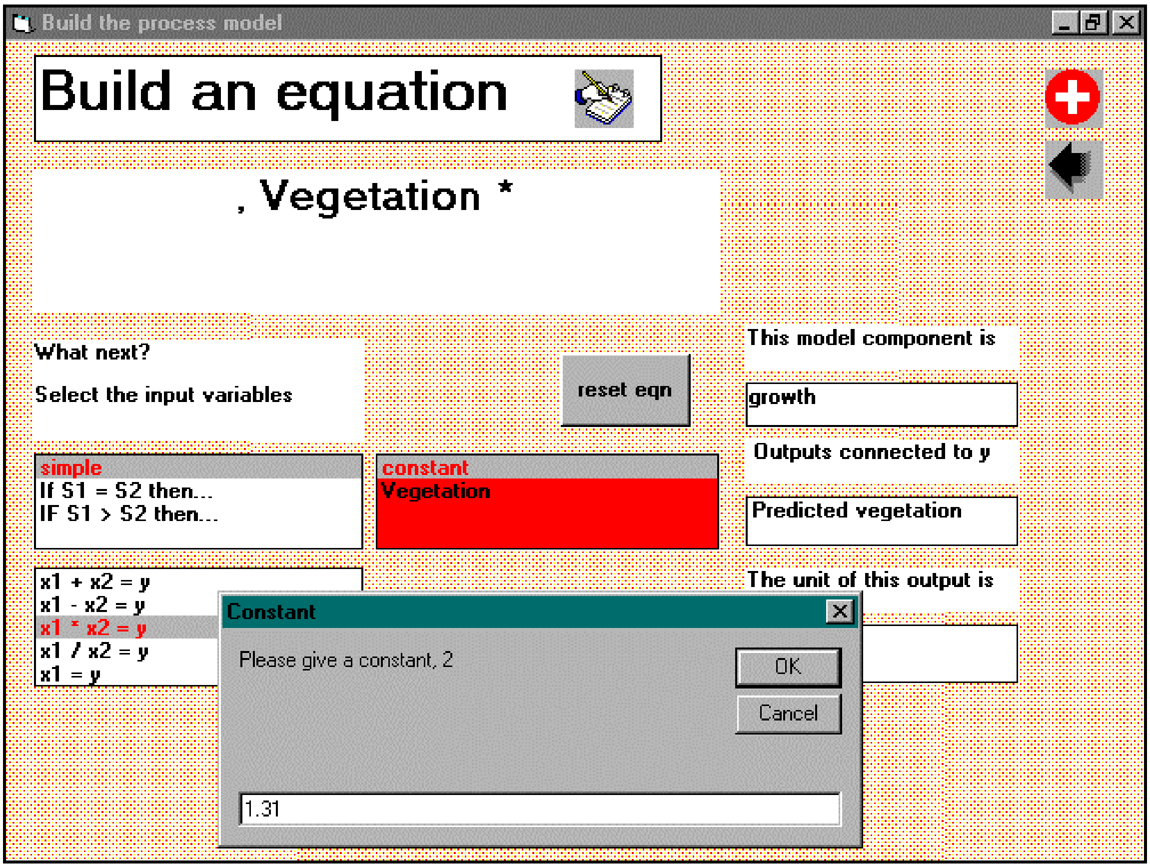


Figure 4: More complex growth model

In the SPMS users build process diagrams interactively (see Figure 3 and Figure 4). At the core of the program, the GIS analysis[1] and modelling functions are combined. Within a graphical environment, process models (emulating objects) are linked together visually. Also available to form part of the model are spatial objects, with inputs and outputs. Objects may be joined to form complex structures allowing feedback mechanisms.

---

[1] GIS input and preparation are done elsewhere (eg Idrisi) It would also be desirable to use methods such as the federated schema suggested as part of Lake's (1996) electronic round table for environmental management.



Build an equation

, Vegetation *

What next?

Select the input variables

If S1 = S2 then...
IF S1 > S2 then...

Vegetation

x1 + x2 = y
x1 - x2 = y
x1 * x2 = y
x1 / x2 = y
x1 = y

This model component is

growth

Outputs connected to y

Predicted vegetation

The unit of this output is

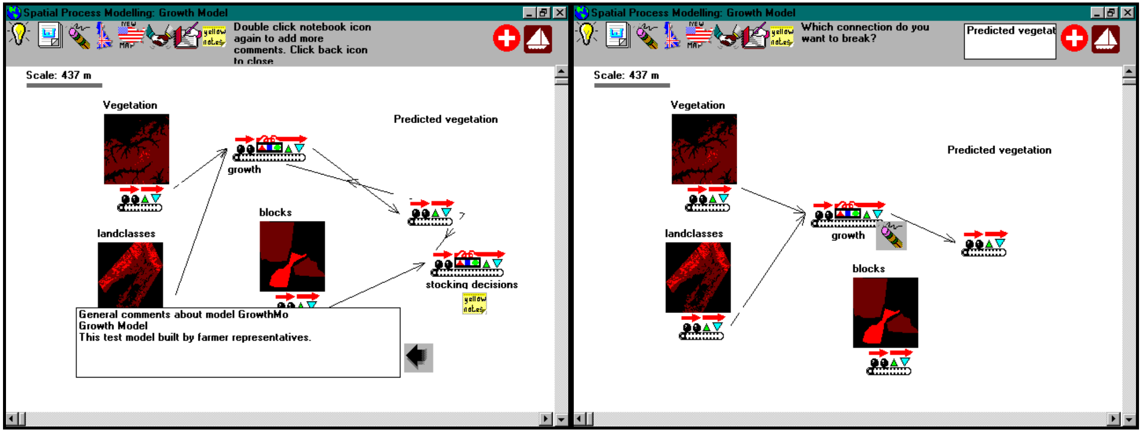
Constant

Please give a constant, 2

1.31

Figure 5: Equation builder, presents linked components and leads user through equation definition.

Lines are drawn linking the objects and then the equation for each process model is defined according to the inputs and outputs. The equation builder (Figure 5) presents the links connected to the process component and leads the user through defining the equation. In the current implementation this equation is relatively simple, for example add 5 to input 1 when input 2 = 10. It would however be beneficial to include more complex geographical operators (eg Albrecht's 1996 Universal GIS Operators) and more complex conditional statements. For example in Figure 4 different decisions may be made



Double click notebook icon again to add more comments. Click back icon to close

General comments about model GrowthMo
Growth Model
This test model built by farmer representatives.

Which connection do you want to break?

Predicted vegetat

Figure 6: The system is flexible and allows deletion of links and whole components. Annotations may refer to the whole model or to specific components.

for each farm block but the SPMS cannot currently deal with this. A method of including tabular data is needed to achieve this goal.

The equation may also be annotated with any assumptions made, for example, 'this is how tussock grass responds to burning, I know that is different in very dry years' (Figure 6). This has the duel advantage of clearly laying out assumptions, and in directing areas where further research is needed. Including technical details in the model but hiding them from view may also facilitate the integration of research findings from a number of disciplines.

5.2 Technical Description

The SPMS has four major components, the interface control, model structure tables, model launcher and report generator (Figure 7).

The system was developed in Visual Basic (4 Professional) as this encompassed the values of rapid software development, is PC based and is compatible with a large range of existing software (notably MS-Access and MS-Word). Access is used for the model structure data while OLE links to MS-Word allows generation of reports from modelling sessions.

The main form (at the top right of Figure 7) has an expandable drawing space and a toolbox of icons. At the start of a session a dialogue box asks for the model name, which is used to generate a new Access table for that model. All model components are then represented by a row in that table. Fields are shown in Table 3. The object type refers to whether the component is a process or map object. A further two object types, tabular and chart are intended in future development. The interface controller uses the screen coordinates to position each component and associated links. The apparent redundancy

Table 3: Model structure fields

Session name
Model
Index number
Object name
Object type
Unit
Map file name
Screen positional data
Links
Inputs (from model number)
Outputs (to model number)
Line references
Equation information
    (equation type, constants, referential index values)

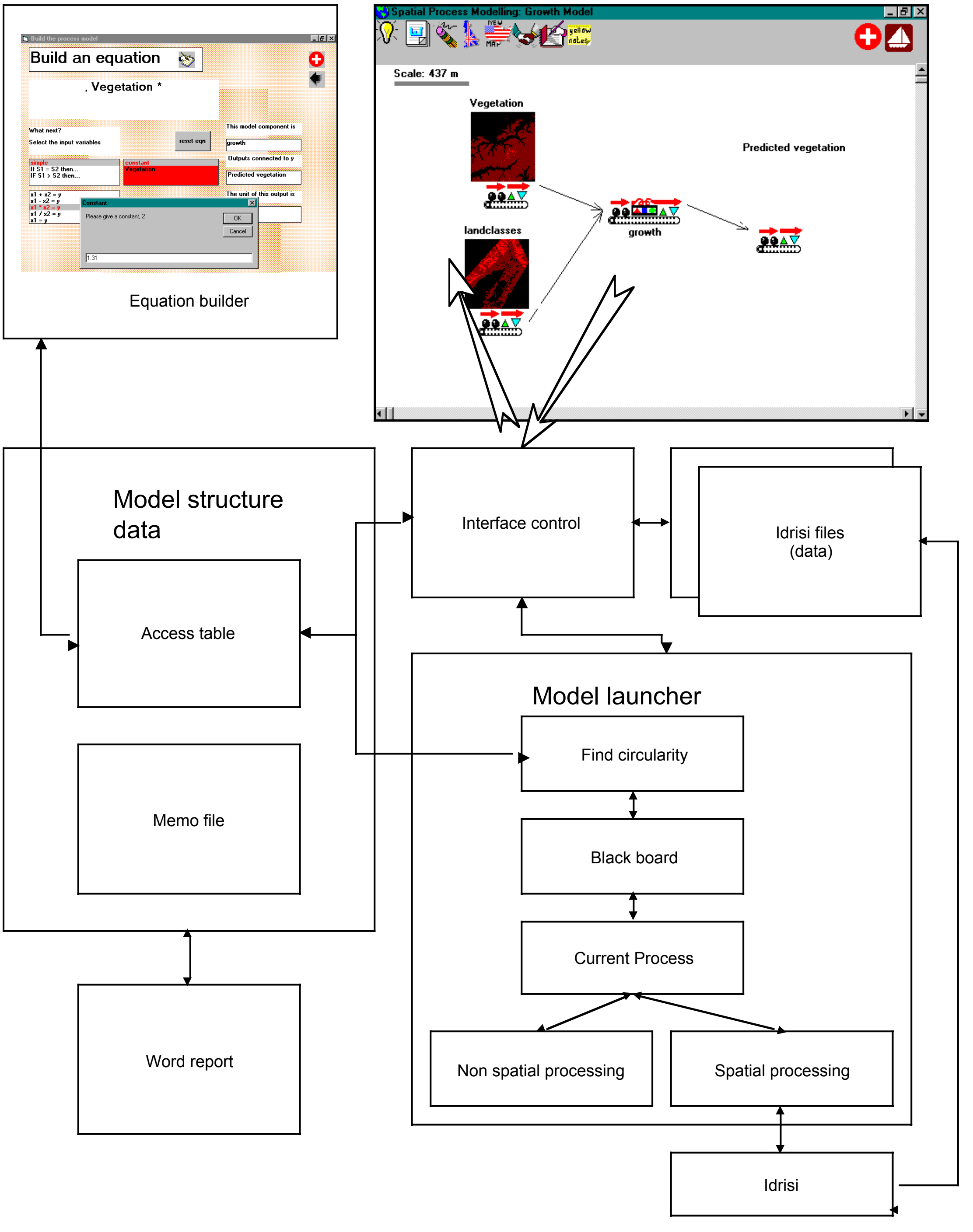


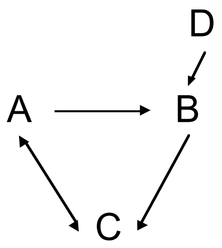
Figure 7: SPMS Structure

in the input and output links is to overcome a limitation in Visual Basic in that lines do not have properties enabling identification and selection. As Figure 6 shows, links can be selected and deleted, as can whole model components.

The model can be launched at any stage (but will not give results unless links are complete). This will 'compile' the model and run it. This is performed by a series of functions and arrays that are used to interpret the model structure table. Figure 8 shows the structure of the arrays needed to launch a simple test model structure. For the moment, assume that all components are aspatial and that equations have two components without conditional operators. This then is a relatively simple feedback loop with equations assigned to each process. The first function moves through the model structure ('Modelboard array') following the links. Component 'A' requires a constant; '20' and component 'C'. This means 'C' must be either initialised or computed before 'A'. 'C' is found to require 'A' which means the model contains circular references, 'C' must be initialised and the user is so prompted. When all references are resolved, the model is reordered onto the 'Blackboard array' so it can proceed. The 'Current process function' then moves through each component in turn performing calculations and feeding the values back to the 'Blackboard'. In this case all processing is performed within the SPMS but if one or more components are spatial then Idrisi (Eastman 1992) is used. This is made more complex as spatial components feed through the model. In the example case, if 'D' was a map object, three methods of performing calculations are required; both aspatial, one spatial one aspatial, and both spatial (see Table 4). Calculations performed in Idrisi are driven for each 'Current process' in turn by command line exported from SPMS via .bat and .pif files. This is transparent to users. Also note from Table 4 that by the end of two iterations, five maps the size of the original represented by ‘D’ have been created. This can quickly lead to a data storage problems so in the current implementation, only the first and last values are stored. When processing is complete control is returned to the interface control and any changed maps are updated.

Testing with various structures has shown that the prototype SPMS can manage structures such as that represented in Figure 8, and even more complex combinations of test configurations with moderate success. Unfortunately when applied to more realistic cases (such as Figure 4) there is less success. Firstly as already established, the equation builder does not accept complex or tabular concepts. Second, the growth model has links from both ‘predicted vegetation’ and ‘vegetation’. Intuitively the model should use the ‘vegetation’ on the first cycle and ‘predicted’ thereafter, but is currently unable to do this. What is needed is a method of including the temporal dimension into the model structure and expressing this in both the equation builder and the launch processor.

Figure 8: Array structure to launch test model structure



A= C + 20
B = A/D
C = A-B
D = 20
i = 2

Model Board				
Ref	x1	x1type	x2	x2type
A	C	ref	21.	Cons
C	A	ref	B	ref
B	A	ref	D	ref
D	20	cons	-	-

Blackboard (t=0)			
	Current Value	Current type	Current time
A		nonspatial	0
D		nonspatial	0
B	40	cons	0
C	150	Initialised	-1

Current component	Time	x1	x2	operator	returns
C	0	150	na	=	150
A	1	C	20	+	170
D	1	40	na	=	40
B	1	A	D	/	5.67
C	1	A	B	-	164.33
A	2	C	20	+	184.33
D	2	40	na	=	40
B	2	A	D	/	4.61
C	2	A	B	-	179.72

Table 4: Effect of Component 'D' being spatial

Current component	t	x1	x2	op	returns	Performed by
C	0	150	na	=	150	SPMS
A	1	C	20	+	170	SPMS
D	1	spatial	na	=	map_D	na
B	1	A	D(spatial)	/	map_B1	Idrisi Scalar
C	1	A	B(spatial)	-	map_C1	Idrisi Scalar
A	2	C(spatial)	20	+	map_A2	Idrisi Scalar
D	2	spatial	na	=	map_D	na
B	2	A(spatial)	D(spatial)	/	map_B2	Idrisi Overlay
C	2	A(spatial)	B(spatial)	-	map_C2	Idrisi Overlay

6 Discussion

Difficulties encountered when dealing with the environment result in inadequate regional decision making. This situation is not helped by the current generation of computerised aids. Conceptual criteria have lead to a desire to combine GIS and process based modelling in a way that is useful for decision making. This has lead in turn to design criteria for a Spatial Process Modelling System. A prototype development of such a system has shown that while most of the criteria are met there remains computational difficulties in developing a functional system. The three most crucial of these are needs for an explicit representation of time, an improved equation builder with better conditional statements, and the incorporation of tabular data. Once these features are operational, more long term goals include the development of a hierarchical representation, modular links to other modelling tools, and data management. Work on the representation of error with assessments of sensitivity to variation in model components would also be beneficial.

When these problems are overcome the SPMS with feedback loops and the ability to modify model structure within a spatial paradigm may become a useful tool in Regional Environmental Management. The system relies on the successful integration of spatial processing and display with the characteristics of system dynamic modelling.

7 References

ALBRECHT, J. H. 1996. Universal GIS operators for environmental modelling. Third International Conference Integrating GIS and Environmental Modelling, Sante Fe, NCGIA.

ALBRECHT, J. H., JUNG, S., MANN, S. 1996. VGIS-- a GIS shell for the conceptual design of environmental models. GISRUK'96, Canterbury UK.

BALL, G. L. 1994. “Ecosystem modelling with GIS.” *Environmental Management* 18: 345-349.

BISHOP, I. 1995. Creating an interactive immersive decision environment: linking modelling to visualisation. Proceedings of the 23rd Annual International Conference of the Australasian Urban and Regional Information Systems Association Incorporated, Melbourne, AURISA. 474-482

BURROUGH, P. A. AND FRANK A. U. 1995. “Concepts and paradigms in spatial information: are current geographical information systems truly generic.” *International Journal of Geographic Information Systems* 9 2): 101-116.

CHAPMAN, K. 1979 *People, pattern and process: an introduction to human geography.* London, Edward Arnold.334

CLEAVES, D. A. 1995. “Assessing and communicating uncertainty in DSS.” *AI Applications* 9 (3): 87-102.

COLEMAN, M. B., BEARLY, T. L. BURKE, I.C. AND LAUENROTH, W.K. 1994. Linking ecological simulation models to geographic information systems: an automated solution. *Environmental Information Management and Analysis: Ecosystem to Global Scales* Eds. W. K. Michener, J. W. Brunt and S. G. Stafford. London, Taylor and Francis. 397-412.

CROSSLAND, M. D., WYNNE, B. E. AND PERKINS, W.C. 1995. “Spatial decision support systems: an overview of technology and a test of efficacy.” *Decision Support Systems* 14(3): 219-237.

DESPOTAKIS, V. K., GIAOUTZI, M. AND NIJKAMP, P., 1993. Dynamic GIS models for regional sustainable development. *Geographic Information Systems: spatial modelling and policy evaluation* Eds. M. M. Fischer and P. Nijkamp. Berlin, Springer-Verlag. 235-261.

EASTMAN, J. R. 1992. “IDRISI: A grid based geographic analysis system.” *Clark University Graduate School of Geography.*

GERAGHTY, P. J. 1993**.** “Environmental assessment and the application of expert systems: an overview.” *Journal of Environmental Management* 39: 27-38.

GRAYSON, R. B., DOOLAN, J. M. AND BLAKE T. 1994. “Applications of AEAM (Adaptive Environmental and Management) to Water Quality in the Latrobe River Catchment.” *Journal of Environmental Management* 41: 245-258.

GOODCHILD, M. AND R. H. S. WISE 1992. “Integrating GIS and Spatial Data Analysis: Problems and Possibilities.” *International Journal of Geographical Information Systems* 6(5): 407-423.

GROSSMAN, W. D. ANDEBERHARDT S. 1994. Geographic information systems and dynamic modelling: potentials of a new approach. *Innovations in GIS* Ed. M. F. Worboys. London, Taylor and Francis. 167-180.

HIGH PERFORMANCE SYSTEMS 1990. Stella II. Hanover, New Hampshire.

HOLLING, C.S, (ED). 1978. Adaptive environmental assessment and management, International Institute for Applied Systems Analysis. pg377

KIRCHNER, T. B. 1994. Data management and simulation modelling. *Environmental Information Management and Analysis: Ecosystem to Global Scales* Eds. W. K. Michener, J. W. Brunt and S. G. Stafford. London, Taylor and Francis. 357-375.

LAKE, R. 1996. Electronic round table: a contribution to co-operative land management. GIS Applications in Natural Resources. M. Hect, H. Dennison-Parker and A. Shortreid, GIS World**:** 66-74.

LEIMBACK, M. (1994). "Expert system-model coupling within the framework of an ecological advisory system." Ecological Modelling **75/76**: 589-600.

LILBURNE, L. 1996. The Integration Challenge. SIRC'96: Proceedings of the 8th Annual Colloquium of the Spatial Information Research Centre, University of Otago, University of Otago.

LOWES, D. AND WALKER D. 1995. Environmental Management: Opportunities and Challenges in the Application of AI. Proceedings of the Workshop on AI and the Environment: Eighth Australasian Joint Conference on Artificial Intelligence, Australian Defence Force Academy, Canberra, University College, University of New South Wales. 1-11

MANN, S. 1995. “A Case Study in Decision Making.” *Environmental Perspectives* 8(August):1-4

MANN, S. 1996. Spatial Process Modelling for Regional Environmental Decision Making. SIRC'96: Proceedings of the 8th Annual Colloquium of the Spatial Information Research Centre, University of Otago, University of Otago.

MANN, S. AND BENWELL G. L. 1995. GIS in environmental management. AURISA/ 7th Colloquium of the Spatial Information Research Centre, Palmerston North, NZ, 295-310

MARR, A. J. AND BENWELL G. L. 1996. “The evolution of the use of GIS in New Zealand Local Government.” New Zealand Surveyor(286): 30-35.

NIELSON, N. R. (1991). Application of artificial intelligence techniques to simulation. Knowledge-based simulation - methodolgy and application. P. A. Fishwick and R. B. Modjeski. New York, Springer-Verlag. **4:** 1-19.

PANDEY, S. ANDHARDAKER J. B. 1995**.** “The role of modelling in the quest for sustainable farming systems.” *Agricultural Systems* 47: 439-450.

ROBERTSON, D., BUNDY, A. MUETZELFELDT, R. HAGGITH, M. AND USCHOLD, M. 1991**.** *Eco-logic: logic based approaches to ecological modelling*. Cambridge MA, MIT Press.

STEYAERT, L. T. AND GOODCHILD M. F. 1994. Integrating geographic information systems and environmental simulation models: a status review. *Environmental Information Management and Analysis: Ecosystem to Global Scales* Eds. W. K. Michener, J. W. Brunt and S. G. Stafford. London, Taylor and Francis. 333-355.

WALTERS, C.J.. 1986. Adaptive Management of Renewable Resources. New York, MacMillan. pg374

GIS Maturity and Integration

Andrew J. Marr and George L. Benwell
Spatial Information Research Centre,
Department of Information Science,
University of Otago, PO Box 56, Dunedin, New Zealand,
Tel. +64 (3) 479 8301, Fax. +64 (3) 479 8311,
ajmarr@commerce.otago.ac.nz
gbenwell@commerce.otago.ac.nz

1. Abstract

This paper discusses the concept of maturity in the use of GIS and then formulates a computational method for measuring an organisations maturity level from the construction of a surrogate indicator. Generation of this model is made under the proposition that maturity is linked to the level that GIS has been integrated and utilised on an organisation wide basis in day to day activities. The research focuses on New Zealand local government and incorporates parallel studies of conventional information technology (IT) with recently collected data to provide support for the concepts and techniques used. It is postulated that due to similarities of function found in other local authorities, that the model has the potential, with further research for wide application.

1. Introduction

In New Zealand as in other parts of the world, Geographic Information Systems (GIS) have become a firmly established production tool in many areas of regional and local government. New Zealand local authorities have traditionally possessed large amounts of spatial and non-spatial data (Fraser and Todd, 1994; and Anderson and Benwell,

1992) and GIS is increasingly being used as a tool for data management and analysis (Marr and Benwell, 1996a). For most organisations GIS offers a variety of benefits both tangible and intangible over more traditional and less automated methods (Mackaness, 1989; Dickinson and Calkins, 1988).

The majority of GIS implementations in New Zealand have occurred in the last three years, but some date back prior to 1988 (Marr, 1996). This situation has occurred over a period of major technological improvement, enhancing the computational functions and abilities of GIS. In addition, there has been a rapid increase in the amount, availability, and quality of digital data (Marr and Benwell, 1996a). These factors over time have led to contrasts in the way GIS has been implemented among different organisations. The proposition is made that trends observed in the local authorities represent an identifiable maturity process from which a principle component is the organisation wide integration of data resources.

1. Maturity and Integration

Some organisations consider the development of a totally integrated information system as the logical process of increasing organisation wide overall efficiency (Marr and Benwell, 1996b). The result of organisation wide data integration may be that GIS becomes less of a standalone system, and more of an incorporated tool responsible for spatial problem solving (Mayr, 1995; Zwart, 1992).

Although not regarded as authoritative, "GIS maturity is defined as the degree to which systems are actually used, which in turn relates to the number of users" (Mayr, 1995, p30).

Figure 1 lists some of the benefits that may be achieved if organisation wide integration is attained. It is suggested that where an organisation is found to have widely integrated its data resources, the relative level of maturity will be high in the use of those resources. As levels of integration increase, the number of users for which use of the new system would be advantageous also potentially increase resulting in a higher maturity level inline with the definition by Mayr.

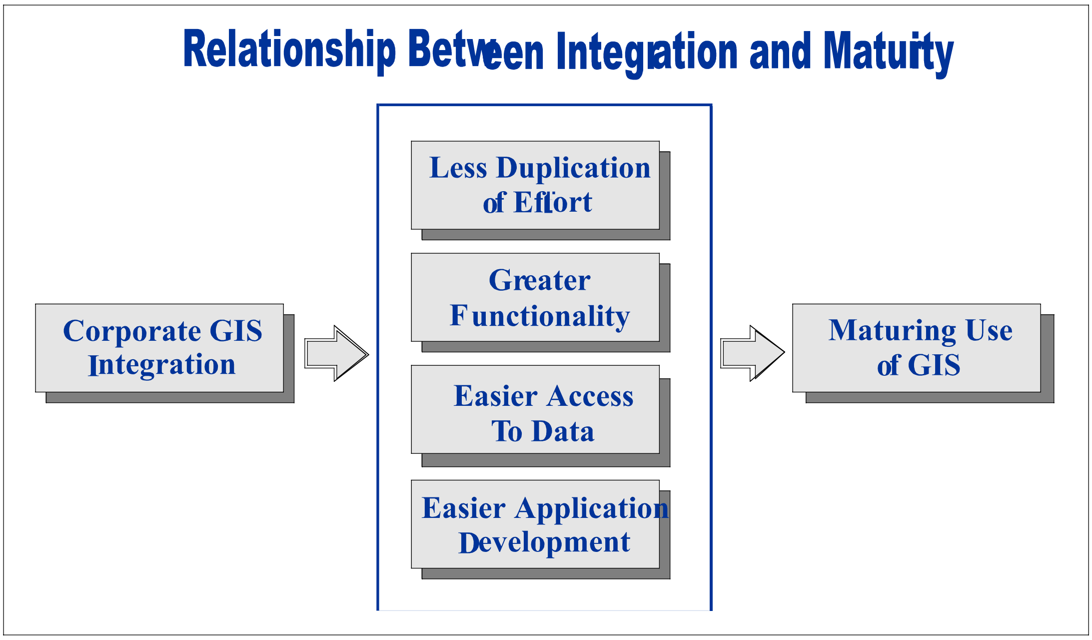


Figure 1 - Relationship Between Integration and Maturity
from Lopez and John (1993)

The effects of corporate integration were analysed based on both New Zealand and Australian local government surveys carried out between 1992 and 1995. The 1995 survey was carried out and reported on by Marr (1996). The survey was conducted on all New Zealand local government organisations, and resulted in a response rate of

74.4%. Of this response, 70% of the organisations indicated they had GIS. While most organisations perceived corporate integration as desirable, most were far from such a reality. Related to the integration of data resources is the inherent need for GIS analysis to be generally available within an organisation. Typically in the past, GIS implementation has focused on a particular application and therefore a related department of the organisation. As GIS management moves to a corporate role, there is often a need to re-engineer internal processes to cope with the changes which commonly see GIS control move to the information services department or equivalent (Marr and Benwell, 1996a).

To illustrate this trend, Figure 2 shows the primary activity of the person most responsible for managing GIS in the organisation based on the returned responses.

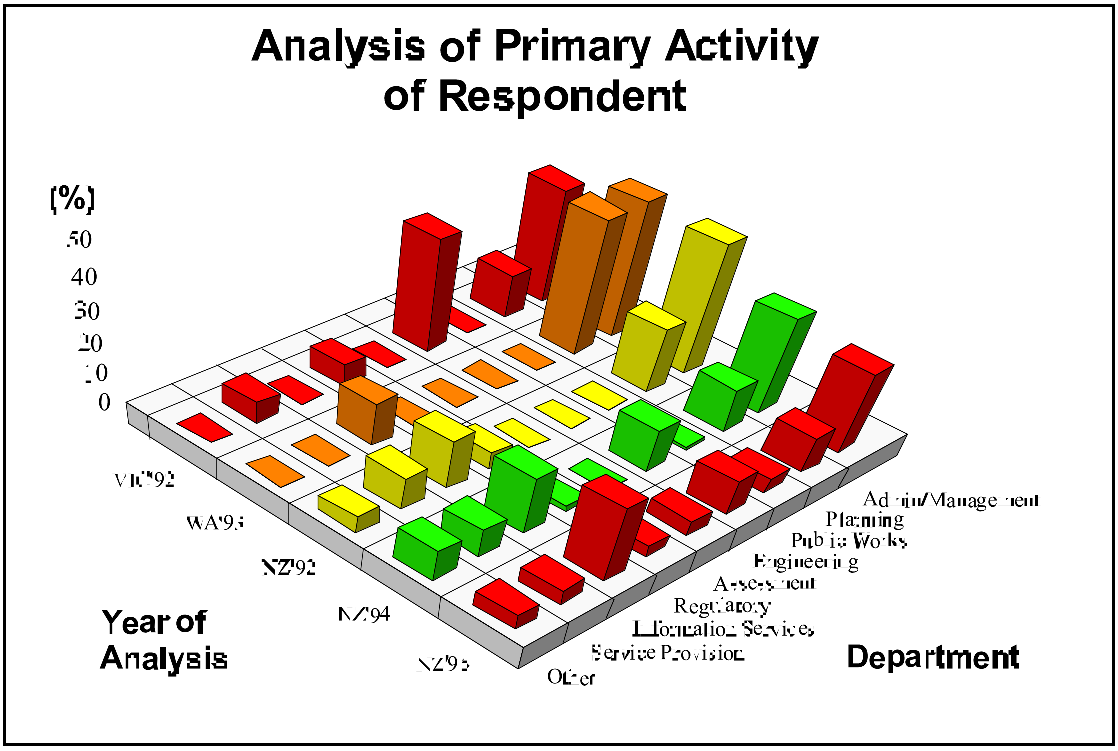


Figure 2 - Analysis of Primary Activity of Respondent
from Marr (1996)

The trends suggest that the role of the information services department in managing the corporate resources is increasing while the role of more traditional departments such as administration/management and planning is decreasing. These trends are further corroborated by the analysis of Figure 3.

This graph shows the department responsible for GIS within the organisation through the successive studies.

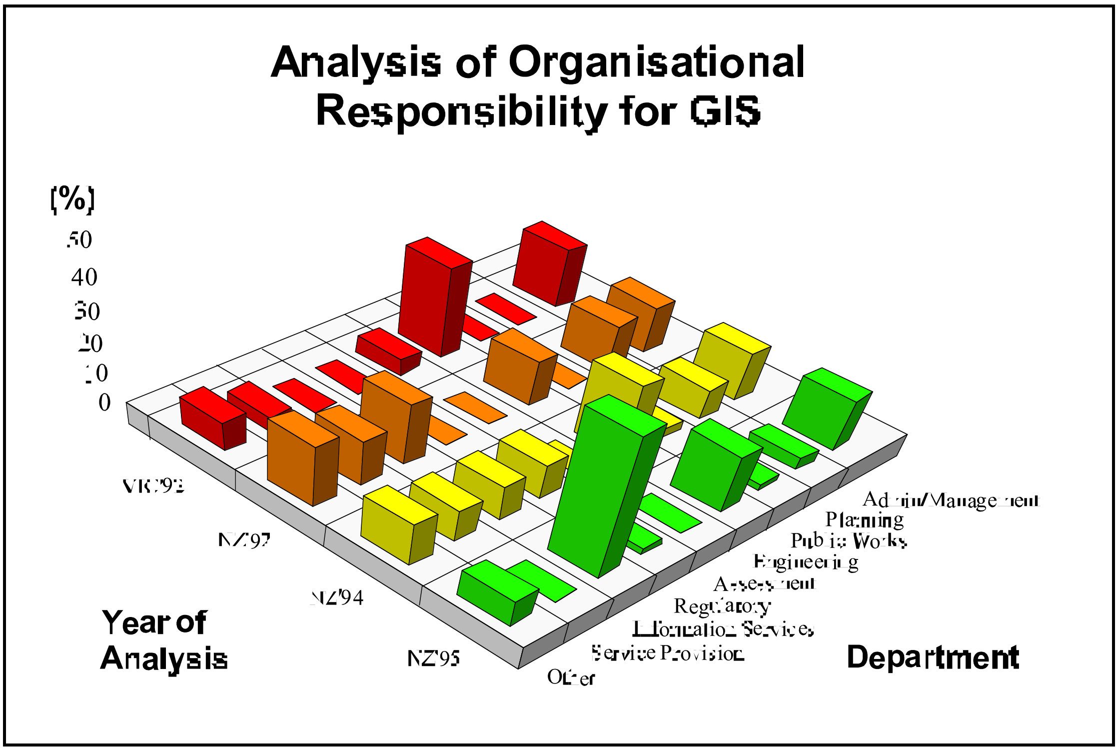


Figure 3 - Analysis of Organisational Responsibility for GIS
from Marr (1996)

There appears to be a marked increase in the percentage of information services departments assuming responsibility for GIS in the organisation.

2. The Nolan Model

There exists a very close relationship between the technologies of GIS and other forms of IT. Indeed, organisations with a policy of corporate integration find that it eventually becomes difficult to differentiate GIS from other IT. Total integration would see this differentiation disappear altogether with the development of generic tools for spatial manipulation and analysis.

Among most local government organisations in New Zealand and probably elsewhere, GIS implementations may be regarded as immature and not fully developed. However extensive research has been carried out into conventional IT where greater development has had time to occur. Some of the more fundamental work has evolved in the form of the 'Nolan Model' shown in Figure 4 which suggests the components of an IT maturing process (Nolan, 1979; Nolan, 1977; Gibson and Nolan, 1974; Nolan, 1973). The development of the model and associated management guidelines are widely considered to be the most comprehensive research in this area (Sprague and McNurlin, 1993; Jackson, 1986; Benbasat *et al.*, 1984).

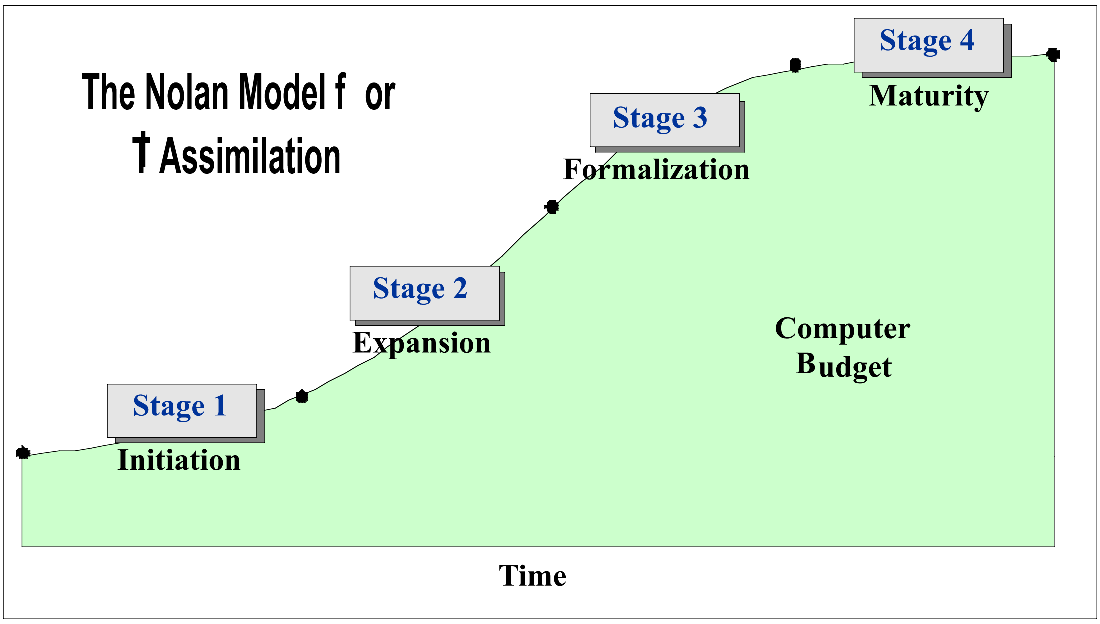


Figure 4 - The Nolan Model for IT Assimilation
From Gibson and Nolan (1974)

Based on the observation of a number of organisations, Nolan hypothesised that the IT implementation could be identified to pass through a series of stages beginning with initiation and ending with maturity. These stages were identified based on the analysis of computer budget expenditure for each organisation and plotted as a curve (Figure 4). It was suggested by Gibson and Nolan (1974) that the curve, represented the growth of applications, the growth of personnel specialisation, and growth of formal management techniques. From the analysis of the organisations it was postulated that the further through the stages an organisation was, the greater the efficiency achieved in its use of IT (Nolan, 1979; Nolan, 1977; and Gibson and Nolan, 1974). This is attained through the integration, standardisation, and dispersal of IT throughout the organisation.

The assertion that the computer budget may used as an surrogate indicator for a number of aspects relating to IT management has been disputed on several occasions (King and Kraemer, 1984; Drury, 1983; and Lucas and Sutton, 1977).

Whilst acknowledging its limitations, the appeal of the model is its simplicity and logical progression. Most IT managers can relate their organisation to a position on the development model and therefore may take advantage of the management policies suggested by Nolan for each stage. These policies attempt to facilitate innovation and development while restricting cost overruns. While Nolan (1973) acknowledges that the use of this indicator is less than ideal, the point is made that few alternatives exist for the convenient verification of the assertions made.

3. Maturity and Integration with GIS

Organisations identified as having a higher level of maturity, appear to be integrating their GIS with conventional IT (Marr and Benwell, 1996b). It is therefore suggested that the Nolan model provides a suitable basis for the development of a parallel GIS model. New Zealand local government organisations were asked to determine their relative position on a development line similar to that shown in Figure 5. The line is not intended to represent a linear relationship, but more a general progression of development. Respondents in the survey were asked to indicate via the addition of a vertical line, the position they felt their organisation had reached. By the inclusion of an arbitrary scale on the X axis, their response could be measured numerically.

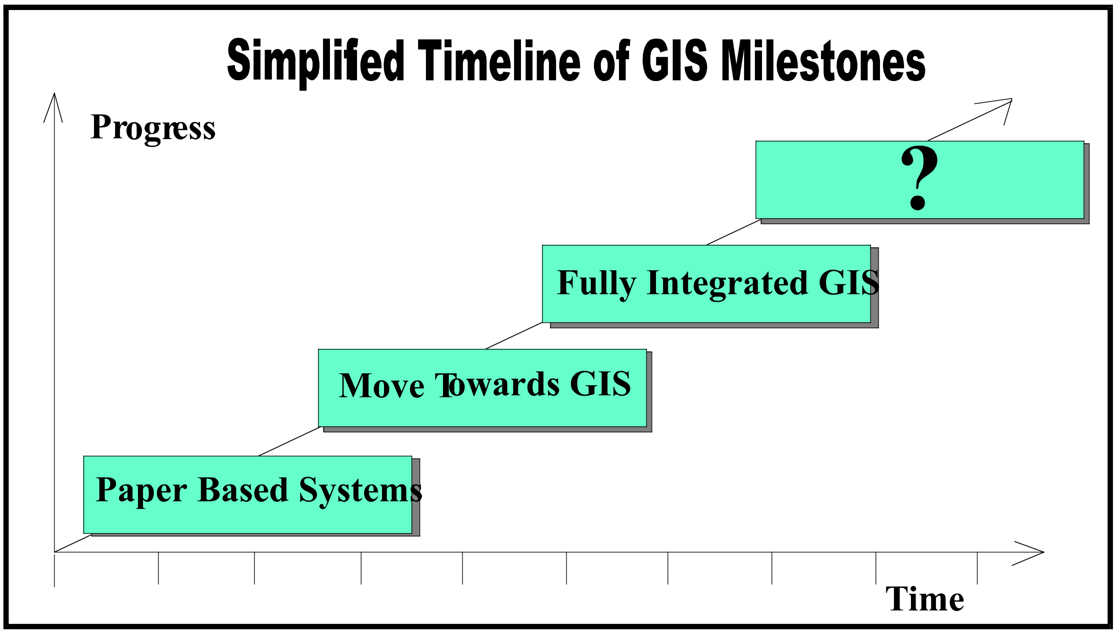


Figure 5 - Simplified Timeline of GIS Milestones
From Marr (1996)

The definition of '*Paper based systems*' is where primary operation in the organisation is undertaken using new and existing paper plans, documents, and microfiche. *'Move towards GIS*' describes the situation where initial GIS purchase, both hardware and software has been made, and electronic data capture is in progress, resulting in some limited functionality. When 100% of the data capture is complete (or very close), and the GIS is fully operational achieving original objectives, it is described as being at the '*Fully Integrated Stage'*.

The respondents were also asked to name the fourth unidentified stage. This was designed to illicit the response of 'Integration of Corporate Data Resources'. This is where all organisational data is integrated as a corporate resource. Based on the response to the question 56% of the respondents correctly identified this stage. This suggests that while for most organisations the GIS implementation process is far from complete, there is clear focus on the end objectives. The timeline is highly simplified and subjective, but it is believed that most organisations with GIS would relate to the general concept shown. The respondents were provided the opportunity to comment on the model. Comments were made on the length of time required during the data capture phase, and some raised the issue as to whether corporate integration would ever be achieved. While these points are acknowledged, none of the respondents dismissed the model outright, leading to the confidence that the model and resulting responses are appropriate.

The graph not only represents a general progression of development, but is also indicative of a rise in the maturity level in the use of GIS in the organisation. This is founded on the assertion that as integration increases, the number uses and users also increase. This is inline with maturity definition by Mayr (1995). Although a numerical indicator is created, it suffers in accuracy from the subjective and continuous scale from which it is derived.

The remaining portion of this paper is devoted to the formulation of a computational method for measuring an organisation's maturity level based on the analysis of discrete data. It is suggested that such a method would provide an improved model, that was free from bias, from which to perform cross-organisation comparison.

4. A Computational Indicator of Maturity

Six individual variables (Table 1) believed to be potentially related to GIS maturity by the authors, were identified from the survey and a correlation matrix (Table 2) produced using SPSS software (Noru_is and SPSS Inc., 1993). In addition, the variable POS was also included in the matrix representing the maturity level identified for each organisation from the development line. The assumption was made that variables achieving high correlation coefficients in relation to the POS variable, could be regarded as significant in the construction of a maturity indicator. POP and AGE were the only unconstrained variables, the others related only to the options provided. The effect of this aspect with regards to the creation of a generic indicator of maturity is to provide standardisation.

ID Code	Description of Variable
ACCEPT	The Degree of Acceptance of GIS in the Organisation
DEPT	The Department Responsible for GIS
NUMD	The Number of Departments in the Organisation Using GIS
NUMUSE	The Number of the Uses which GIS is Assisting
POP	The Population Base of the Local Government Organisation
AGE	The Age in Years of the GIS Implementation

Table 1- Description of the Selected Maturity Variables

Marr and Benwell (1996b)

	ACCEPT	AGE	DEPT	NUMD	NUMUSE	POP	POS
ACCEPT	1	0.0360	0.0684	0.3075	0.3291	0.0525	0.1586
AGE		1	0.1286	0.4865	0.4329	0.2877	0.3132
DEPT			1	0.1564	0.1528	0.1774	0.0400
NUMD				1	0.4979	0.2226	0.3613
NUMUSE					1	0.2763	0.5314
POP						1	0.0610
POS							1

Table 2- Coefficients of Possible Maturity Variables

From Marr and Benwell (1996b)

The highest three correlation matrix values in relation to POS were NUMUSE, NUMD, and AGE with 0.5314, 0.3613, and 0.3132 respectively. To further analyse the relationship between each variable and POS, Chi square tests (χ^2) were performed (Daniel, 1990). This additional statistical analysis was used to show that each variable and POS are dependant and further justify the variables selected. The values for each variable were categorised in matrix form against the POS value for each organisation. When required, real values on the boundary of categorisation, were rounded up to the nearest integer. Associated with the Chi square test is the Cramér statistic value which used to determine the strength of an identified relationship (Table 3). The results confirm the use of these variables as significant

ID Code	X^2	χ^2	Deg. of Freedom	Cramér Value
NUMUSE	18.402	16.812	6 (to 0.99)	0.50
NUMD	37.457	21.666	9 (to 0.99)	0.62
AGE	N/A	N/A	N/A	N/A

Table 3 - Statistical Analysis of Maturity Variables

From Marr and Benwell (1996b)

In the case of AGE, the average expected frequency failed to obtained the minimum of 2.0 recommended by Daniel (1990). This would have introduced uncertainty in any resulting statistics if followed through. Still, a relationship clearly exists as shown in Table 2.

Although some circularity and inter-variable dependence is likely to exist, it is considered unavoidable since the assumption has been made that all the variables combine to form an approximate maturity level of GIS use. Observation of organisations with GIS suggests that no single variable is capable of being used as an appropriate indicator of maturity.

To create the required surrogate indicator of maturity in GIS use, there is a need to incorporate the values of the three variables (NUMUSE,NUMD, and AGE) to form one new computational indicator of maturity. One method of achieving this in diagrammatic form is shown in Figure 6. Each organisation may be represented in three-dimensional space based on the value of the three variables identified above. The variables form the three axis of the model. An extended line is drawn from the origin, through an axis represented by the corresponding values for each variable on the correlation matrix (Table 2). Values on the line are then taken to represent the level of integration and thus a surrogate measure for maturity of GIS use. This technique has previously been used by Lilburne (1996) to show the inter-relationships of interface, data, and functionality in systems integration.

The position of an organisation on the integration line is determined by computing the shortest distance between the (X,Y,Z) of the organisation and the line itself. The imaginary line from each point forms a 90º angle with the line of integration. The standard algebraic equations to analyse these are shown in Figure 7, and may be incorporated in any spreadsheet package. Also shown is the formula to compute the length of a point on the line from the origin. This distance then becomes the amalgamated indicator of maturity.

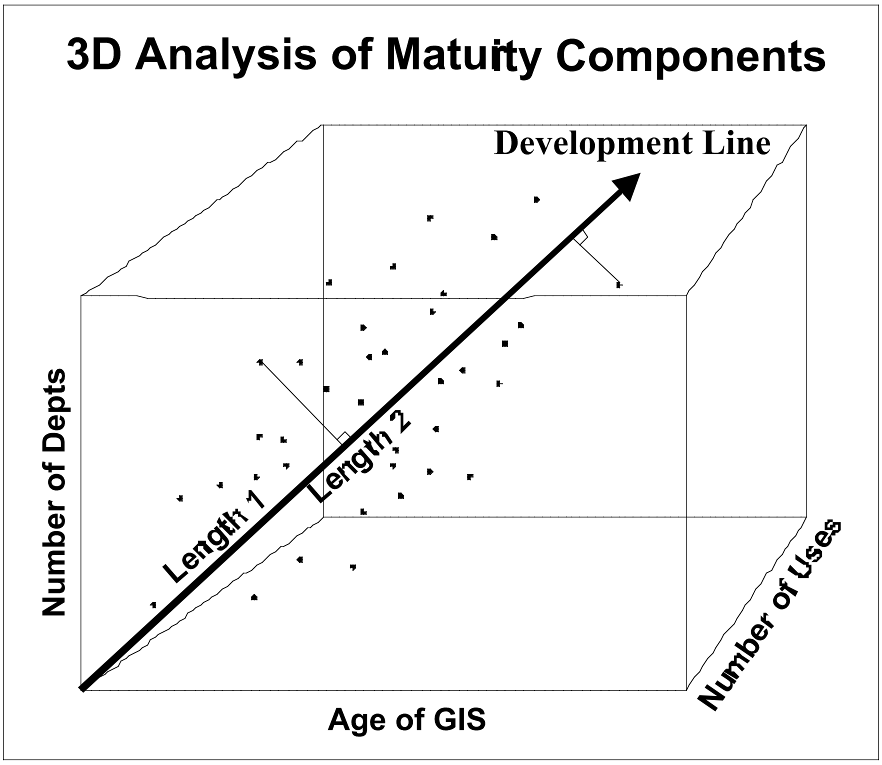


Figure 6 - 3D Analysis of Maturity Components with Two Examples

from Marr and Benwell (1996b)

To test this concept, a correlation matrix was produced to assess the relationship between the derived maturity value from the model and the position of the responding organisation as indicated in Figure 5. This resulted in a positive correlation coefficient of 0.5233. In addition, the Chi-square test was also performed. This resulted in X^2 of

24.598 which is greater than χ^2 of 13.277 with 4 degrees of freedom. The Cramér statistic produced a value of 0.577 indicating a very strong relationship. For additional analysis the values for each organisation derived from the integration model were plotted against the POS values indicated by each respondent. The resulting graph (Figure 8) shows a positive relationship between the two values for each organisation as expected.

Where

$\underline{n} = (n_1,n_2,n_3)$ = coefficients(*numuse, numd, age*)
$\underline{a} = (a_1,a_2,a_3)$ = (*NUMUSE, NUMD, AGE*)
$\underline{z} = (z_1,z_2,z_3)$ = *Standardised Position Vector*

Then

$$\underline{\hat{n}} = \frac{\underline{n}}{|\underline{n}|} = \frac{1}{\sqrt{n_1^2 + n_2^2 + n_3^2}}(n_1,n_2,n_3)$$

Position Vector = (Z_1, Z_2, Z_3) =

$$(\underline{a}.\underline{\hat{n}})\underline{\hat{n}} = \left(\frac{a_1n_1 + a_2n_2 + a_3n_3}{n_1^2 + n_2^2 + n_3^2}\right) \times (n_1,n_2,n_3)$$

$$\text{Distance } |\underline{Z}| = \sqrt{Z_1^2 + Z_2^2 + Z_3^2}$$

Figure 7 - Equations for Position Determination

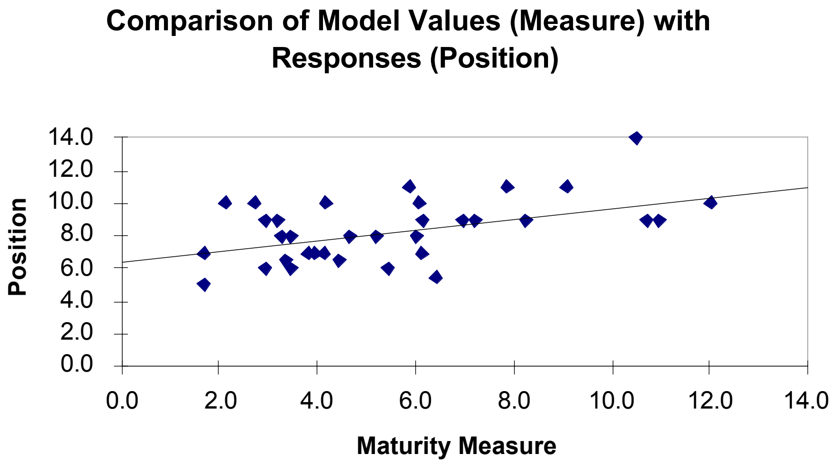


Figure 8 - Comparison of Model Values and Response Values
From Marr and Benwell (1996b)

The graph and the previous statistical analysis suggest that the number of uses of GIS, the number departments using GIS and the age of the GIS, can be used in combination to form an approximate measure of GIS maturity in New Zealand local government. While potential problems have been acknowledged such as circularity and inter-dependence, the model represents a serious attempt to map an organisation's effective GIS development. The development of the indicator draws on previous research in this field relating to general IT. It is suggested that the Nolan model forms a suitable foundation, due to the observed parallel nature of the two technologies, particularly in those perceived to be at a more mature level of operation. The model uses data identified as being related to maturity (Mayr, 1995), but also encompasses the concept that the differences between maturity of use and integration appear minimal.

5. Conclusion

The aim of this paper was first, to discusses the concept of maturity in the use of GIS and second, formulate a computational method for measuring an organisation's maturity level from the construction of a surrogate indicator. From the accumulated data and discussion it has been found that the process of a maturing GIS can be mapped empirically with some degree of confidence. With the use of statistical analysis techniques, a comparison has been made as to which organisations have the greater level of GIS maturity based on the variables NUMUSE, NUMD, and AGE. This model can be used to assess an organisation's progress in relation to other comparable authorities. It is also suggested that with more testing, the model concept may possibly apply to other organisations outside local government.

6. References

ANDERSON, M.R. AND BENWELL, G.L. 1992. A Survey of GIS Usage Amongst Local Authorities in New Zealand and Victoria, Australia, In: *The 4th Annual Colloquium of the Spatial Information Research Centre*, University of Otago, Dunedin, New Zealand, 18 - 20 May, p11 - 24.

BENBASAT, I.DEXTER, A.S.DRURY, D.H. AND GOLDSTEIN, R.C. 1984. A Critique of the Stage Hypothesis: Theory and Empirical Evidence, *Communications of the ACM,* 27(5): 476 - 485.

DANIEL, W.W. 1990. *Applied Nonparametric Statistics*, PWS-Kent Publishing Company, Boston, 633p.

DICKINSON, H.J. AND CALKINS, H.W. 1988. The Economic Evaluation of Implementing a GIS, *International Journal of Geographical Information Systems,* 2(4): 307 - 327.

DRURY, D.H. 1983. An Empirical Assessment of the Stages of DP Growth, *MIS Quarterly,* June:59 - 70.

FRASER, K.G. AND TODD, E. 1994. The Use of GIS in New Zealand Local Authorities - A Survey Comparing the Situation in July 1993 with February 1992, In: *The 4th Annual Colloquium of the Spatial Information Research Centre*, University of Otago, Dunedin, New Zealand, 17 - 19 May 1994, p147 - 159.

GIBSON, C.F. AND NOLAN, R.L. 1974. Managing the Four Stages of EDP Growth, *Harvard Business Review,* 52(1): 76 - 88.

JACKSON, I.F. 1986. *Corporate Information Management*, Prentice-Hall, New Jersey, 338p.

KING, J.L. AND KRAEMER, K.L. 1984. Evolution and Organizational Information Systems:An Assessment of Nolan's Stage Model, *Communications of the ACM,* 27(4): 466 - 475.

LILBURNE, L. 1996. The Integration Challenge, In: *The 8th Annual Colloquium of the Spatial Information Research Centre*, Spatial Information Research Centre, Dunedin.

LOPEZ, X.R. AND JOHN, S.A. 1993. Data Exchange and Integration in UK Local Government. *Mapping Awareness & GIS in Europe* 7,(5): 37 - 40.

LUCAS, H.C.J. AND SUTTON, J.A. 1977. The Stage Hypothesis and the S-Curve: Some Contradictory Evidence, *Communications of the ACM,* 20(4): 254 - 259.

MACKANESS, W. 1989. Introducing GIS into Organisations, In: *The Inaugural Colloquium of the Spatial Information Research Centre*, University of Otago, Dunedin, New Zealand, 29 Nov - 1Dec 1989, p106 - 119.

MARR, A.J. AND BENWELL,G.L. 1996(a). The Evolution of the Use of GIS in New Zealand Local Government, *Journal of the New Zealand Institute of Surveyors,* April, 286:30 - 35.

MARR, A.J. AND BENWELL,G.L. 1996(b). Maturing GIS in New Zealand Local Government, In: *The 8th Annual Colloquium of the Spatial Information Research Centre*, Spatial Information Research Centre, Dunedin.

MARR, A.J. 1996. Geographic Information Systems Maturity in New Zealand Local Government. University of Otago, Unpublished M.Sc. Thesis, 126p.

MAYR, W. 1995. Levels of GIS Maturity in Malaysia. *The Australasian Geographic Information Systems Applications Journal* ,(12): 30 - 31.

NOLAN, R.L. 1979. Managing the Crisis in Data Processing, *Harvard Business Review,* 57(2): 115 - 126.

NOLAN, R.L. 1977. *Management Accounting and Control of Data Processing*, National Association of Accountants, New York, 192p.

NOLAN, R.L. 1973. Managing the Computer Resource: A Stage Hypothesis, *Communications of the ACM,* 16(7): 399 - 405.

NORU_IS, M.J. AND SPSS INC., 1993. *SPSS for Windows - Base System User's Guide, Release 6.0*, SPSS Inc, Chicago, 828p.

SPRAGUE, R.H. AND MCNURLIN, B.C. 1993. *Information Systems Management in Practice*, Prentice-Hall International Inc., New Jersey, 554p.

ZWART, P. 1992. Mature GIS - or Taking the G out of GIS, In: *The Proceedings of the Fourth Annual Colloquium of the Spatial Information Research Centre*, The Spatial Information Research Centre, Dunedin, 18 - 20 May 1992, p1 - 9.

Spatial Data Acquisition from Motion Video

Mark Williams

Departments of Computer and Information Science

University of Otago, PO Box 56, Dunedin, New Zealand

Phone: +64 3 479 8488

mrwilliams@otago.ac.nz

ABSTRACT

Geographic information systems are an important tool for the field of geocomputing. A key component of every system is the data—spatial data has traditionally been labour-intensive to collect, and hence expensive. This paper establishes a new method of acquiring spatial data from motion video. The proposed method is based upon the principles of photogrammetry, but allows position to be calculated with feature tracking rather than point correspondence. By doing so, it avoids many constraints imposed by previous solutions. The new method is demonstrated with linear and rotational motion.

INTRODUCTION

The field of geocomputing utilises computer models of the real world. While in some cases there is no spatial correspondence between the computer model and reality, many cases require an accurate representation of the real world. For example, geographical information systems (GIS) commonly require accurate, scaled maps of land. A simple GIS may incorporate a flat two dimensional map from which the computer could, for example, measure distance or calculate areas. A more complex version may include a topographic map, which is essentially a layered set of two dimensional maps at fixed height intervals. From this information the GIS can calculate volumes of hills, or viewshed images that determine the visibility of, for instance, a large structure on a hilltop. More detailed GIS models require true three dimensional models of objects, such as tunnels, to measure volume. In all cases a GIS is only as good as the source data collection, hence the data collection technique needs to be both accurate and practical for the size of the object being modelled. In the case of traditional GIS, photogrammetry is well established as a data gathering method. This large-scale technique has also been proven successful for some small-scale tasks. For example Fraser (1988) demonstrated a semi-automated method for measuring ships' hulls and satellite dishes. On an even smaller scale, Rivett (1983) showed photogrammetry has been used for measuring human faces to evaluate post-operative swelling.

The quintessential problem of photogrammetry is to establish correspondence, which involves identifying the points in each image that represent a given point in space. This has proven to be an easy task for a trained human operator, but very difficult for a computer. Solutions often require that easily identifiable reference objects be placed within the scene. They may also require controlled lighting and manual identification of corresponding target points. There is, however, a class of problems where alteration of the scene is undesirable, making automated measurement very difficult.

A new method is proposed that aims at building a three dimensional model from easily obtained visual data. Currently the requirements of this method are normal lighting and controlled, functional camera motion.

This paper reviews methods of capturing three dimensional spatial data from two dimensional images. A new approach is proposed that is based upon the principles of photogrammetry, while circumventing the correspondence problem. The mathematical foundations of this method are described, then several worked examples are given to demonstrate this technique.

EXISTING METHODS OF SPATIAL DATA CAPTURE

The most common method for obtaining data for a GIS is photogrammetry. In the traditional case of obtaining land data, stereo pairs of aerial photographs are taken; smaller scale data capture may use terrestrial photogrammetry. Both cases process the stereo images to calculate depth of objects within the scene.

Stereo Geometry

Figure 1 shows the basic form of stereo geometry: consider this as a plan view of two cameras capturing a scene. The cameras have focal length f and are separated by the stereo baseline b. For simplification, the cameras are shown with parallel viewing directions, and subsequently parallel image planes. This is not essential, but if the cameras are not parallel a transformation is required to align the coordinate systems.

Two points A and B in the scene appear at A_l and B_l in the left image and A_r and B_r in the right image, and are at distances z_A and z_B from the cameras. Because of the different viewpoints of the cameras, the points appear in different places within each image—this difference is known as *disparity* or *parallax*. The parallax of each point is calculated from the x-coordinate of the image points:

$$p_A = A_l - A_r \qquad\qquad p_B = B_l - B_r$$

Using similar triangles, the depths of points A and B can be calculated:

$$z_A = \frac{fb}{p_A} \qquad\qquad z_B = \frac{fb}{p_B}$$

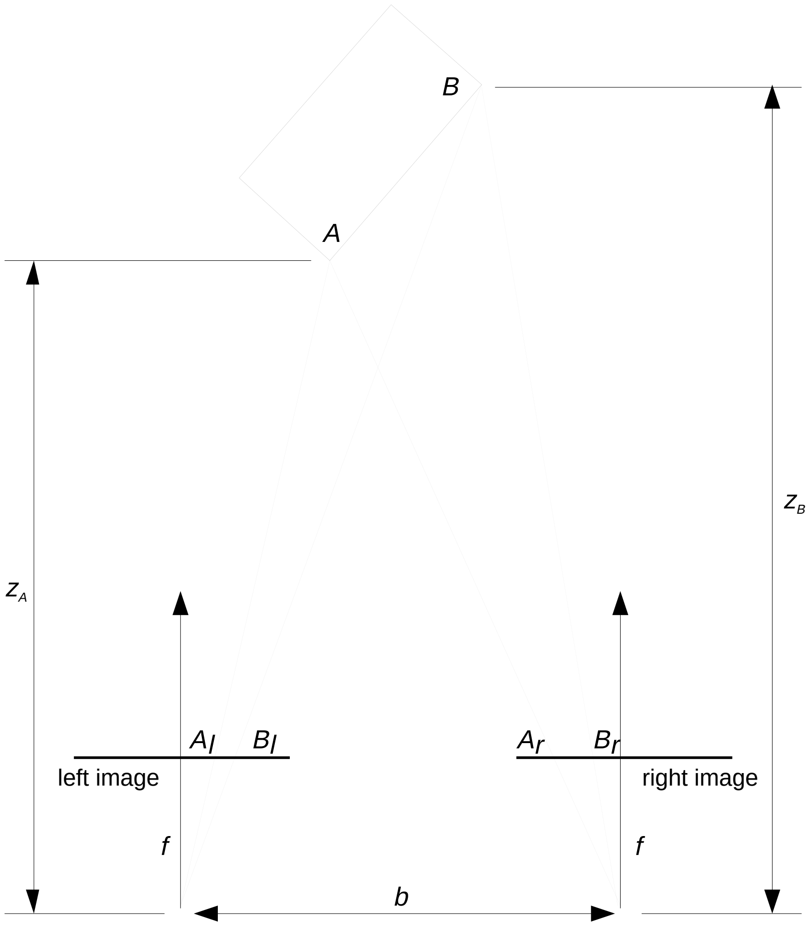


Figure 1: Geometry of stereo vision

These equations demonstrate the requirements of a depth-from-stereo algorithm: given a known baseline and focal length, align the images to reconstruct the relative orientation at the time they were captured, then measure the disparity of each matching pair of points to determine depth. While this process appears very simple, there is one fundamental problem: determining which point in the right image corresponds to a given point in the left image. Thus the key issue for calculating depth from stereo images is the *correspondence problem*. An important technique used by many researchers to assist solve the correspondence problem is the *epipolar constraint*. Figure 2 demonstrates the definition of the epipolar plane. For any point *P* in the scene, an epipolar plane is defined by the point and its representation in the left and right images. The epipolar plane intersects the images forming an epipolar line in each. When the two images are correctly aligned, these epipolar lines will be horizontal. Having located a point in one image, it is only necessary to search along the epipolar line to find the corresponding point in the second image. Hence the epipolar constraint reduces the correspondence problem to a one dimensional search of a scan line for the matching point.

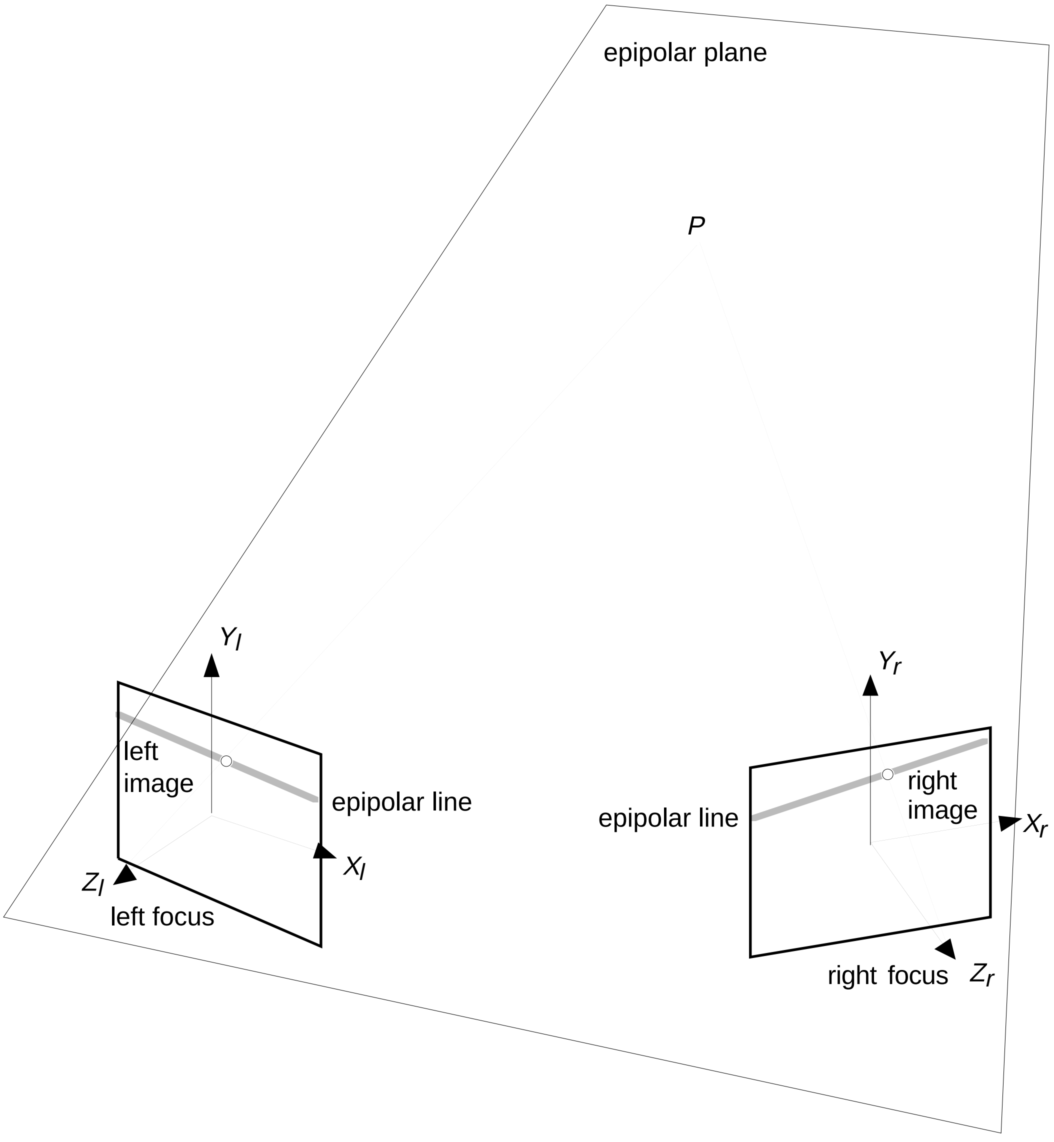


Figure 2: Epipolar Geometry for Stereoscopic Viewing

The epipolar plane is defined by a point *P* in space and the representation of that point in each image. The plane forms epipolar lines where it passes through each image. Having found a feature point in one image, it is only necessary to search for a matching point along the epipolar line in the other image.

Photogrammetry

Automatic photogrammetry systems have taken many different approaches to solving the correspondence problem. These approaches can be summarised in three categories: interactive methods which require a human operator to identify a match, target based systems that attach or project easily identifiable targets onto the object to simplify matching, and fully automatic systems that attempt to identify natural features within the scene to match.

Interactive systems, such as those demonstrated by Erez and Dorrer (1984) and Beyer (1992), are the most simple solution. An operator looks at a stereo pair of digital images and identifies matching feature points in each. The system then refines the location of each feature, for example by performing a least squares match with the intensity value of the pixels. The sub-pixel coordinates achieved by the refinement are then used to obtain an accurate measure of the parallax for that feature.

Target based systems aim to remove the need for operator assistance. This requires making the matching task as simple as possible. A common technique, as demonstrated by Fraser and Shortis (1994), Clarke *et al.* (1995) and Peipe and Schneider (1995), utilises retro-reflective targets and controlled lighting to ensure the targets are the only objects visible in the images. Using the epipolar constraint, matching targets in the stereo image pair is greatly simplified. As with the interactive case, the location of each target is refined by a least squares method using pixel intensities.

The most ambitious method of stereo matching attempts to identify the same point in each image without targets or controlled lighting. Rosenholm (1987a/b) attempted to match the intensity values of a patch or *window* of pixels in each image. Even using the epipolar constraint to reduce the search space to a single dimension, pixel based matching is extremely slow. Kölbl *et al.* (1991) addressed this issue by developing specialised hardware to perform the search in parallel. Koch (1992) reduced the search space by using a more complex evaluation criteria: instead of matching just the intensity of pixels, a gradient vector and match confidence is calculated for a pixel window, giving a more distinctive description to match.

While many photogrammetric systems have achieved satisfactory results for particular domains, the correspondence problem forces those systems to compromise generalisation for automation. Traditional stereo geometry requires correspondence to be solved, and that is the limiting factor in current techniques.

Computational Stereo

The field of computational stereo can be divided into three distinct areas: the human visual system, robot vision and image metrology. While all three areas share the fundamental framework of stereo vision, each has a unique goal, and the methods employed to reach those goals are very different. In the study of the human visual system, computer scientists attempt to model parts of the system, and use the computer models to help gain a better understanding of how human vision works. The emphasis is placed on biological and psychophysical accuracy, in order to ensure results remain relevant to the human model. Meanwhile robot vision is aimed at achieving the more practical goal of navigation and simple object identification. In the case of navigation, the important result is to avoid obstacles while moving in a "real world" environment. Industrial robots are also required to inspect objects, identify and sort them into appropriate groups. Both these tasks require some calculation of three dimensional position, but more emphasis is placed on recognising structure and shape. The final area of computational stereo, image metrology, closely parallels photogrammetry. The goal is to extract measurement of an object, by determining the three-dimensional coordinates of a series of points on its surface.

Barnard [Barn82] claims that the process of automatically matching features between a stereo pair of images is “the hardest and most significant problem in computational stereo”. This is certainly the case in image metrology research, and the variety of methods used fall into two broad categories: pixel- and feature-based matching. The former method attempts to match individual pixels or small clusters of pixels all over the images, creating a dense depth map, while the latter uses image processing techniques to identify sparse significant features to match between the images.

Pixel-based methods, such as those proposed by Cappellini *et al.* (1987), Trivedi (1985) and Zheng *et al.* (1990), rely on comparing the intensity value of individual pixels or small clusters of pixels. Arnold and Binford (1980) noted that pixel-based intensity matching was prone to variation from camera settings, lighting and so on. Instead, they proposed a simple feature-based method, identifying edges, which was more robust under the same conditions. Features are extracted by filter functions and are often edge pixels, line segments, curves or junctions. Marr and Poggio (1982) proposed the Laplacian of Gaussian (LoG) filter, which had a very similar effect to its biological equivalent in the human eye. The LoG filter detects edges by first smoothing in a circular, Gaussian fashion, then taking the second derivative of the smoothed intensity.

To reduce the extracted feature set and improve the uniqueness of each feature, the images can be filtered and features matched hierarchically. For example, Jin and Li (1988) proposed an edge-based hierarchical matching scheme. Edges are extracted by a selectable-resolution filter function. Matching starts with low resolution, gross features, and is continually refined to higher resolutions, using the previous resolution’s results to segment the image and provide depth estimates.

There are clearly more similarities than differences between the fields of photogrammetry and computational stereo, particularly when considering metric photogrammetry and image metrology. One notable distinction between the fields is the use of targets to identify feature points. This is common practise in industrial photogrammetry, and is often combined with controlled lighting to ensure the targets are easily identifiable in the images. Conversely computational stereo approaches aim for generalised solutions, extracting natural feature points or matching clusters of pixels. Accuracy is seldom quantified, while automation is a key goal.

Shape from Motion

Another popular method for obtaining three dimensional data from visual images is *shape from motion*. In such methods, two or more images are captured as the camera—or the scene—moves. The images are captured in rapid succession, so features within the scene will only move a few pixels from one image to the next. Basic principles of perspective dictate that features close to the camera will move differently from those further away. The essential problem is to determine optical flow—the differing motion of various regions within the image sequence. Once motion vectors have been calculated, they are used to determine the slope of surfaces in the scene.

Many techniques have been proposed to calculate optical flow, for instance differential methods, region matching, energy–based and phase–based techniques. A detailed description and comparison of examples of these methods can be found in a thorough review by Barron *et al.* (1992). All of these methods contain the same broad stages of processing. Generally the first step involves filtering or smoothing the images, followed by the extraction of basic measurements of motion. These measurements are then integrated into an optical flow field, which itself is smoothed before being used to determine the surface gradients within the scene.

Monocular Methods

While photogrammetry and computational stereo use stereo pairs of widely spaced images, and motion techniques use streams of closely spaced images, another class of methods for calculating depth use single images. *Shape from shading* analyses the intensity variations within an image to determine the shape of surfaces. The simplest form of shape from shading uses a single image, and makes assumptions about the nature and direction of the light source, as well as the albedo or reflectance properties of the object itself. The lighting graduations on the surface are then used to determine the curvature, and consequently the relative depth can be calculated. Bruckstein (1988) provides an excellent overview of the problem, and a simple recursive procedure to determine equal-height contours of a surface given a tightly constrained experiment.

Due to the number of assumptions required, single image shape from shading techniques are numerically unstable and their results may not be reliable. So the method was extended to *photometric stereo* where images are captured from the same viewing position under different lighting conditions. Lee and Kuo (1992) describe two methods for combining the multiple images: parallel and cascade schemes. The parallel scheme uses all images at the same time, formulating the intensities into a linear system of equations. The solution of these equations gives a height field for the scene. By contrast, the cascade scheme uses a single image shape from shading method for each image sequentially, where the result from one image is used as a starting condition to constrain the next.

Another approach uses photometric stereo with structured lighting. For example, instead of a conventional light source, Klette *et al.* (1995) use a single stripe of laser light projected into the scene. When the stripe of light hits an object it forms a contour line on the surface, which can be easily extracted from an image. By applying this technique in many places over the surface of the object, a very accurate model can be made.

Shape from contour uses a technique quite similar to the structured lighting method described above. Instead of using a laser stripe to illuminate a contour line, the silhouette of an object is extracted from many images captured as it rotates in front of a camera. The series of silhouettes can then be assembled and analysed to determine the object's shape. Unfortunately this algorithm fails with all but the simplest objects, because concave areas of a surface can be occluded in all of the silhouettes.

A more robust approach is described by Zheng (1994), which identifies an area of concavity as a discontinuity in the motion of a contour. The details of Zheng's method are beyond the scope of this paper, but one important aspect is the detection of occluded regions in an epipolar plane image, which is formed by extracting a single selected line from many images, and stacking them all together. Such a process is similar to that used, albeit for a different purpose, by the method proposed in this paper. Other shape from contour methods extract structure from a series of contours by recognising junctions or generalised cylinder, Ulupinar and Nevatia (1992/95).

Another type of single image method determines shape by recognising scene structure. These model based approaches, reviewed by Chin (1986), while particularly important in industrial applications, are not related to the method proposed in this paper, and are mentioned only for completeness.

All of the methods described above have proven to be successful in particular cases. All have constraints, however, including adding reference to the scene, controlling the lighting or making assumptions about the reflective properties of the surface.

SPATIAL DATA CAPTURE FROM FUNCTIONAL MOTION

An alternative approach for calculating depth from images is based upon the principles of photogrammetry, while circumventing the correspondence problem. In doing so many restrictions of other methods are avoided: no reference points are required, nor special markers on the object. Where photogrammetry requires a widely spaced pair of images, the new method uses a dense image set. The motion of the camera relative to the scene is determined by a simple function. Data from the camera is treated as a three dimensional block, consisting of X and Y spatial dimensions, and a temporal dimension. Figure 3 illustrates this *spatiotemporal* volume. Interesting features are tracked as they move within this space; the function of their motion determines the three dimensional location of each feature in space.

Bolles, Baker and Marimont [Boll86 and 87] first proposed slicing through the XT plane of a spatiotemporal volume to form an *epipolar plane image* (EPI). They noted the relationship between the path of a feature within the spatiotemporal volume, and the three dimensional coordinates of that feature in the scene. The aim of their research was a robot vision system, where a measurement of free-space within a scene was more important than measuring individual objects. Hence the purpose of their EPI analysis was to calculate depth and determine occlusions, which proved to be successful. However, they did not establish the relationship between the EPI analysis and the process of traditional photogrammetry.

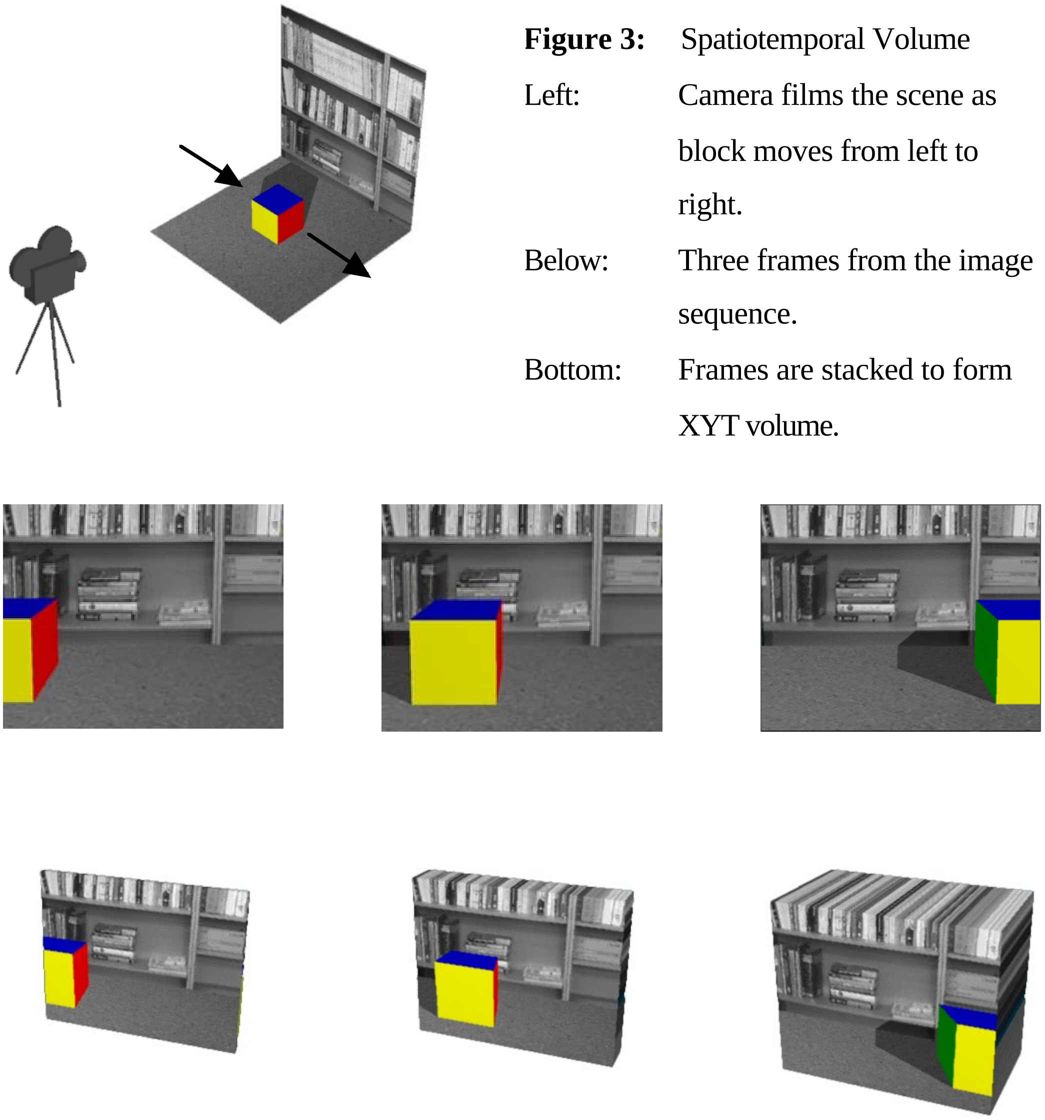


Figure 3: Spatiotemporal Volume
Left: Camera films the scene as block moves from left to right.
Below: Three frames from the image sequence.
Bottom: Frames are stacked to form XYT volume.

Since the relative motion between the camera and the scene is known, the general form of the motion of points within the spatiotemporal volume can be determined. Furthermore, the precise motion of every feature point is a parametrised form of the general case, where the parameters relate directly to the three dimensional position of each point in the scene. Hence the mapping from the given XYT coordinates to the desired XYZ coordinates becomes a task of fitting data to a function by adjusting these parameters.

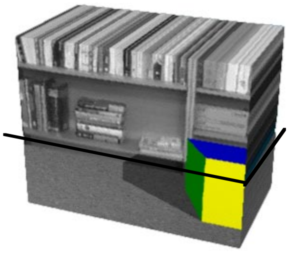
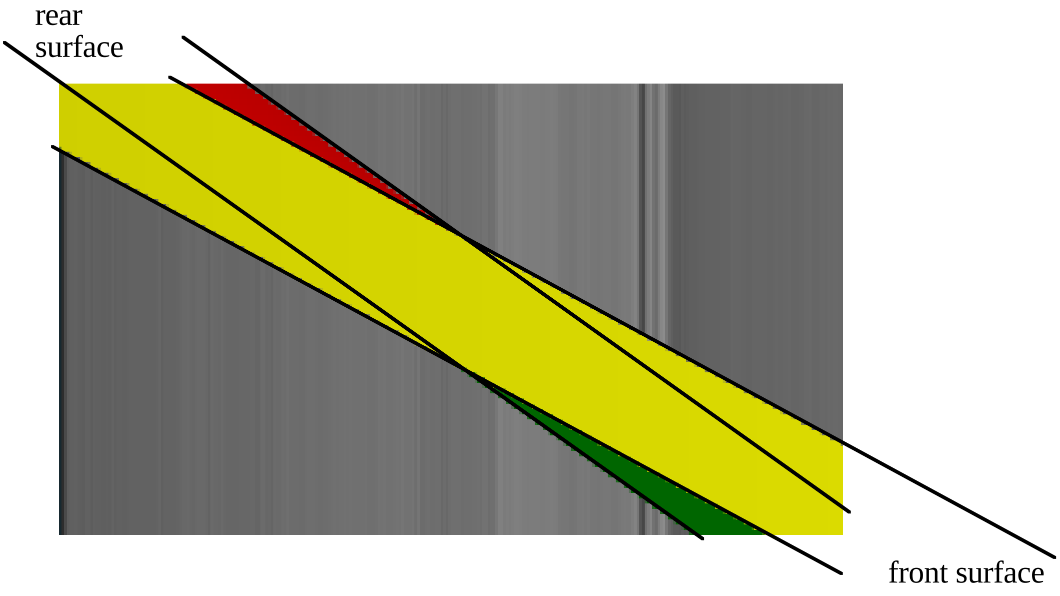


Figure 4: Epipolar Plane Image

This is an epipolar plane image taken from the spatiotemporal volume in Figure 3. Time is shown on the vertical axis. The front and side surfaces of the cube have formed bands through the EPI, where the gradient of each edge is proportional to its position in XYZ space. The front edges are parallel to each other and at a different angle to the rear edges.



Linear Motion Example

The simplest case of the proposed method is linear motion. This will arise, for example, when the camera is stationary while an object moves past it. As the object moves, frames from the camera are captured to form a dense spatiotemporal volume. This volume can be segmented into XT slices: each image plots the paths of points on the object at one height, or Y value, in the XY frames. By considering the apparent motion of individual points over time, it can be observed that points at different distances from the camera move at different apparent velocities. All points move in straight lines, but those closer to the camera appear to move faster than those farther away. Imagine looking out the window of a moving train: trees close to the railway line appear to rush past in a flash, while mountains in the distance appear to move very slowly. When viewing this motion as an epipolar image, objects at different distances from the camera appear as lines of different angles. Measuring the angle between the lines allows a direct calculation of the difference in distance from the camera, and hence the size of the object.

For example, Figure 4 shows an EPI from the spatiotemporal volume shown in Figure 3. Note in this case the object moved through the scene while the camera stayed still. Hence the stationary background forms vertical stripes, while the object itself forms diagonal bands. From the epipolar plane image, it is easy to see the relationship between the path of a point in XYT and its location in XYZ. In this case, the function is a straight line, where the gradient is proportional to the depth of the point from the camera.

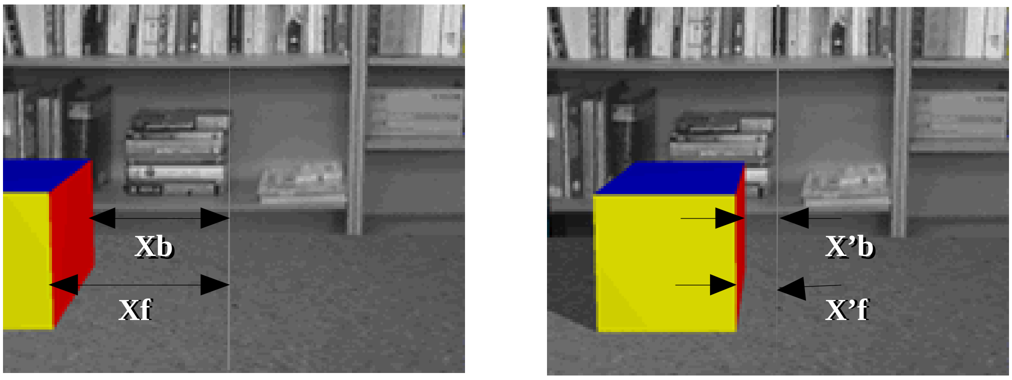


Figure 5: Parallax measurement in Photogrammetry
The distance is measured from the centre-line to the point of interest in each image. Parallax is then calculated as the measurement difference between one image and the other.

Now consider this case in traditional photogrammetric terms. Instead of a continual sequence of images, a stereo pair is selected (Figure 5). To calculate the depth of the cube, a parallax measurement is required for the front and rear surfaces.

$$\begin{aligned} \mathrm{Pf} &= \mathrm{Xf} - \mathrm{X'f} \\ \mathrm{Pb} &= \mathrm{Xb} - \mathrm{X'b} \end{aligned} \tag{1}$$

Given the camera baseline and focal length, depth can be calculated:

$$\mathrm{D} = \mathrm{B}f\left(\frac{1}{\mathrm{Pf}} - \frac{1}{\mathrm{Pb}}\right) \tag{2}$$

Figure 6 demonstrates how the stereo pair images form two lines of the EPI, allowing the parallax measurement to be made directly. The ratio of spatial to temporal resolution in the XYT volume is 1:1, so the camera baseline is equivalent to the vertical distance between the stereo pair slice in the EPI. Hence equation (2) can be expressed as:

$$\mathrm{D} = f\left(\tan\theta_{\mathrm{f}} - \tan\theta_{\mathrm{b}}\right) \tag{3}$$

For the linear motion case then, the proposed method can be expressed in terms of traditional photogrammetric approaches.

Rotational Motion Example

This method extends to rotational motion. As an object rotates before a camera, or the camera rotates around the object, individual points on its surface form sinusoidal paths in an epipolar image. Consider watching a glass of water placed at the edge of the turntable inside a microwave oven. The glass appears to move back and forth periodically: plotting this motion against time produces a sinusoidal curve. Now consider a second glass in the microwave, closer to the centre

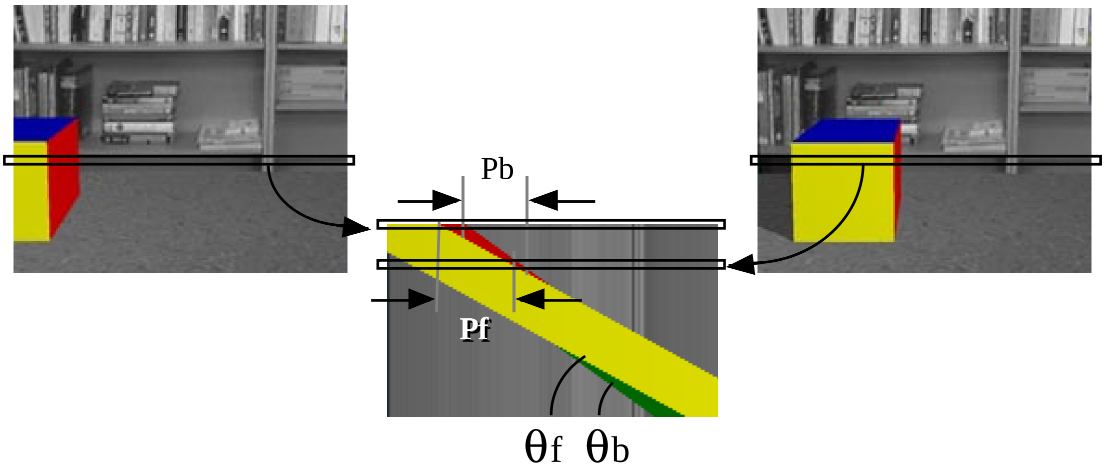


Figure 6: Stereo Pairs in an EPI
The stereo pair from Figure 5 can be considered as two lines within the EPI. Consequently the parallax can be measured directly from the EPI, and is proportional to the angle of the bands.

than the first. The epipolar image will contain two sinusoids, one with a smaller amplitude than the other. If the two glasses are in line with the centre, the sinusoids will be in phase. Shifting the first glass around the edge of the turntable will give the large amplitude sinusoid a different phase to the small amplitude one. By extracting the phase and amplitude of each sinusoidal curve in an epipolar image, the position of each tracked point relative to the camera can be calculated. Again this allows the size of the object to be determined.

In Figure 7 a cube is rotated around its centre and the XYT volume is formed and sliced through an XT plane. Figure 8 shows the resulting EPI, with one edge highlighted as it moves through the spatiotemporal volume. In this example, because the cube was rotated around its centre, all four edges form sinusoids with the same amplitudes, but different phases. The amplitudes and phases of the sinusoids are extracted and used to

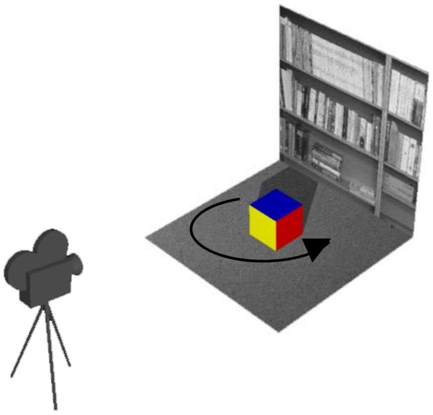
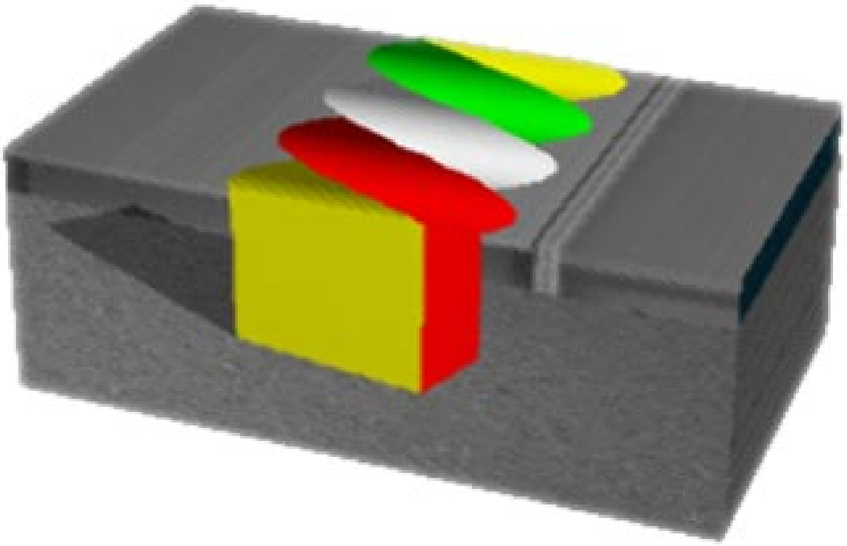


Figure 7: Spatiotemporal Volume

Left: Camera films the scene as block rotates about its centre.

Below: XYT volume sliced to show an EPI.



locate the edges in three dimensional space, in polar coordinates (Figure 9). Since all the sinusoids have the same amplitude, the edges are all the same distance from the centre. Similarly the phase of the sinusoids are at 90° intervals, as are the edges of the cube in polar coordinates.

Also note that the sinusoids are occluded for half of the EPI: the functional nature of each point's locus means once the parameters have been determined, the point can be "found" when it reappears after occlusion.

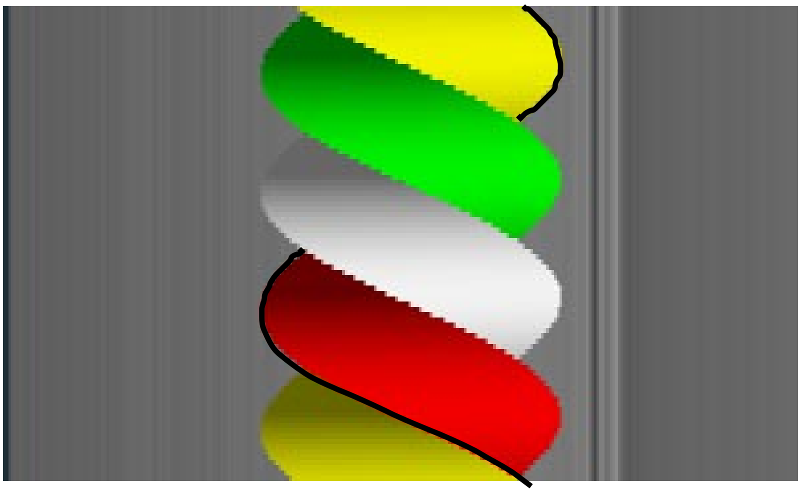


Figure 8: EPI of Rotational Motion

The motion path of each edge is sinusoidal. Time is shown on the vertical axis. Note the highlighted edge becomes obscured then reappears in an expected location.

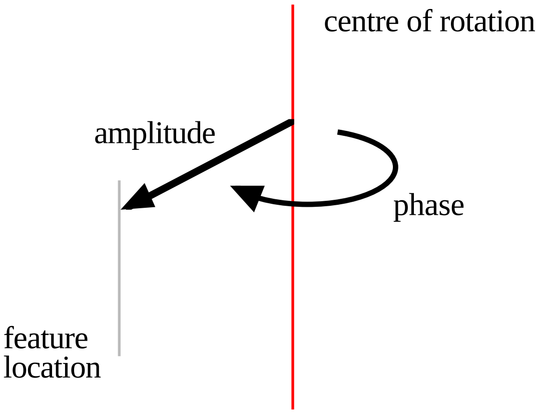
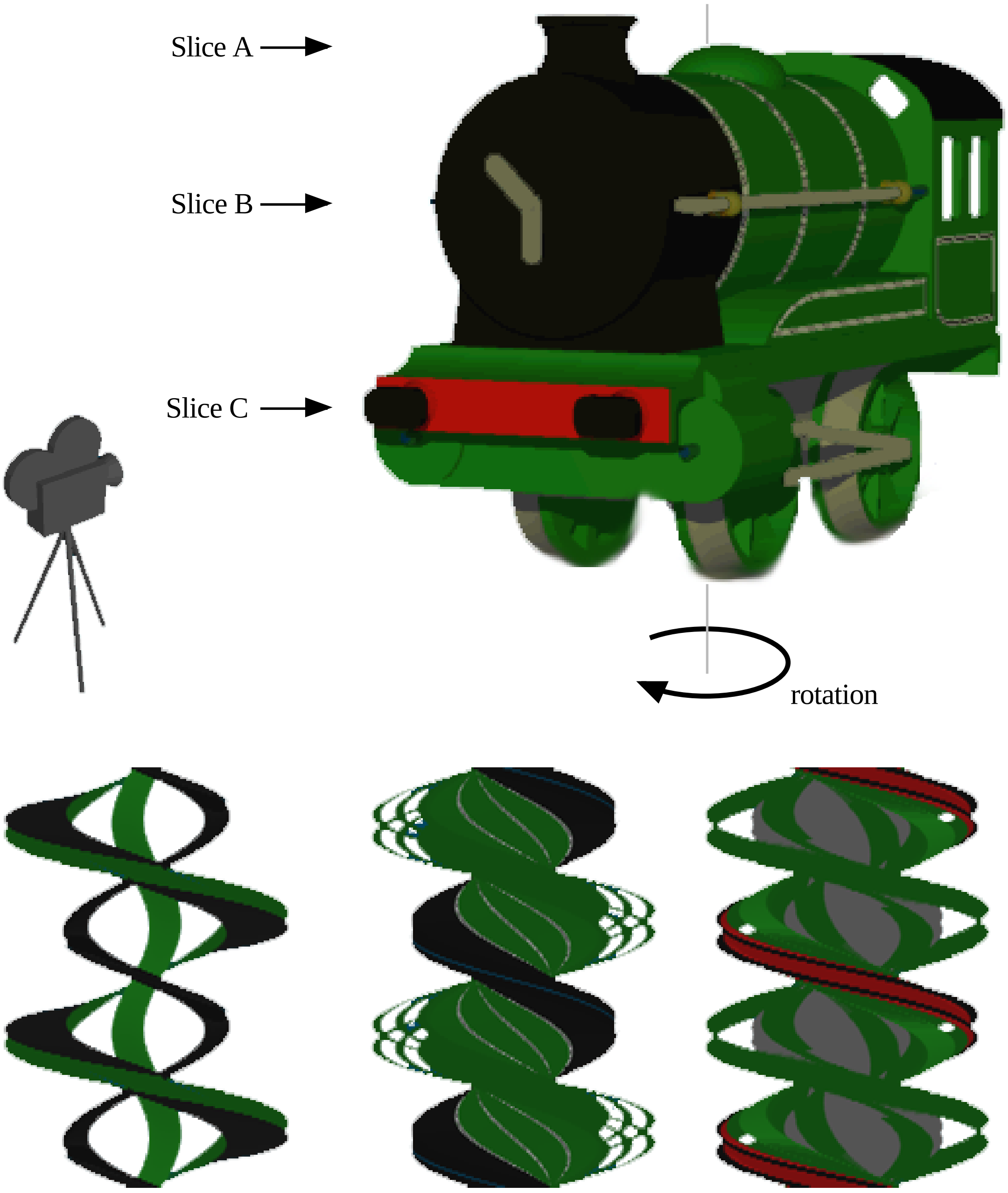


Figure 9: Location using Polar Coordinates

To locate the edge highlighted in Figure 8, the extracted phase determines the rotation and the amplitude determines the distance from the centre.



Slice A Slice B Slice C

Figure 10: Complex Example of Rotational Motion

Top: The locomotive is rotated about its centre as the images are captured.

Bottom: Three EPIs from different positions within the spatiotemporal volume. Slice A shows the smoke-stack, bell top and cab roof, with clearly distinguishable phases and amplitudes. Slice B shows the boiler and cab. Slice C shows the front and rear buffers, including overhang at the rear.

Figure 10 shows a more complex example of rotational motion. The locomotive has been rotated about its centre, and the captured images formed into a spatiotemporal volume. Three EPI slices are shown from various heights within the volume. The first slice contains three sinusoidal bands, representing the smoke stack, bell top and cab roof. These bands clearly show the relationship between the phase and amplitude of each sinusoid, and the position of the feature in polar coordinates. Similarly the other slices show identifiable features as sinusoidal stripes.

Figure 11 shows an example of real video data, using the toy locomotive simulated in Figure 10. Klette *et al.* noted the transition from noiseless simulated data to noisy real data can be catastrophic for shape from motion techniques. Recall that such techniques, however, rely on calculating optical flow fields between consecutive frames, and use iterative methods to smooth the fields and calculate surface gradients. It is clear from the EPI slices in Figure 11 that tracking points through a dense spatiotemporal volume is more robust in noisy data than methods using only a few images.

Hierarchical Filtering Methods

To extract the parameters of the functional motion of each point, it is necessary to track points through the spatiotemporal volume. As Bolles *et al.* pointed out, the consistent image density of the spatial and temporal dimensions allows the data to be treated as a purely spatial image. Hence normal filtering and edge detection techniques work in all dimensions; for example a LoG operator can be applied to an EPI to find edges in the motion paths.

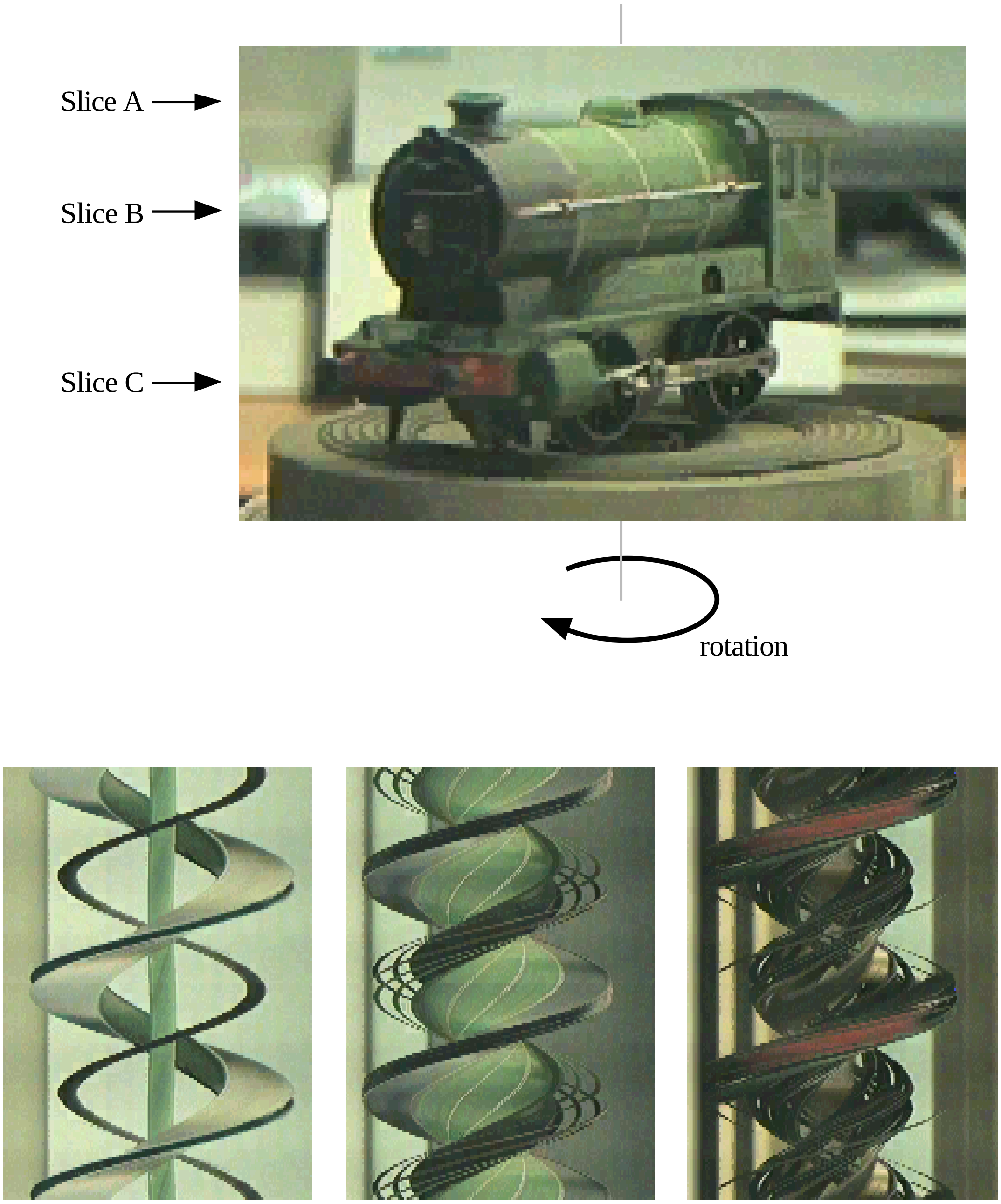
A *data crawler* can be used to step through the XYT volume, attempting to follow a feature point. Each step is determined by a set of constraints: firstly the neighbouring pixels are examined to find the candidates for a closest match given some standard image processing criterion. This search is constrained by the knowledge that all the points' motion is planar within the spatiotemporal block, and furthermore the motion will be functional. Hence the search space is reduced to a region fitting the general profile of the motion. As the motion path of a point is extracted, the error bounds are gradually reduced, so the search constraints can be tightened.

Given the density of the spatiotemporal volume, it would be impractical to attempt to follow every pixel individually. Hence a hierarchical filtering approach is used, similar to Rosenholm's multi-resolution filtering for feature matching. Initially a very broad filter extracts only the most distinct features, which are easily tracked. Once the motion paths of the top level features are determined, they can be used to segment the image for further processing. Multiple passes using progressively finer filters can then add detail to the model, using structure determined at higher levels to constrain the search space.

CONCLUSION

A new method for extracting spatial data from motion video has been proposed. The theory behind the linear and rotational motion has been established, and the methods for extracting parameters for the generalised functional motion of points have been determined. Examples of simple and complex cases with simulated and real data demonstrate that this method shifts the difficult task of solving correspondence to the simpler domain of tracking edges and fitting data points to functions.

The new method requires functional motion between the camera and scene, currently demonstrated for the first and second order cases. Ideally it would be possible to walk around any object with a hand held camera, and then have a computer system automatically produce a three dimensional model of that object. This research is an important step along the way.



Slice A Slice B Slice C

Figure 11: Real Video of Rotational Motion

Top: The locomotive is rotated on a turntable as the images are captured.

Bottom: Three EPIs similar to those in Figure 10. Note the sinusoidal bands are still clearly visible in the noisy video data.

REFERENCES

Arnold, R.D. and Binford, T.O., 1980. Geometric Constraints in Stereo Vision. SPIE Vol 238 Image Processing for Missile Guidance (1980), 281—292.

Baker, H.H., Bolles, R.C. and Marimont, D.H., 1986. A new technique for obtaining depth information from a moving sensor. Proceedings of the ISPRS Commission II Symposium on Photogrammetric and Remote Sensing Systems for Data Processing and Analysis, 1986.

Barnard, S. and Fischer, M., 1982. Computational Stereo. Computing Surveys, Vol 14, Number 4, 553—572.

Barron, J., Fleet, D., Beauchemin, S. and Burkitt, T., 1992. Performance of Optical Flow Techniques. IEEE CV&IP 3, 1992, 236—242.

Beyer, H., 1992. Automated Dimensional Inspection with Real-Time Photogrammetry. International Archive of Photogrammetry and Remote Sensing, 1992, 722—727.

Bolles, R., Baker, H. and Marimont, D., 1987. Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion. International Journal of Computer Vision, 1, 1987, 7—55.

Bruckstein, A., 1988. On Shape from Shading. Computer Vision, Graphics and Image Processing 44, 139—154.

Cappellini, V., Mecocci, A., Renault, S., 1987. Digital Processing of Stereo Images and Three Dimensional Reconstruction Techniques. Time-Varying Image Processing and Moving Object Recognition, 105—115.

Chin, R. and Dyer, C., 1986. Model-Based Recognition in Robot Vision. Computing Surveys, Vol 18 No 1, March 1986, 67—108.

Clarke, T.A., Cooper, M.A.R., Chen, J. and Robson, S., 1995. Automated Three Dimensional Measurement Using Multiple CCD Camera Views. Photogrammetric

Record, Vol 15 No 85, April 1995, 27—42.

Erez, M.T. and Dorrer, E., 1984. Photogrammetic Data Acquistion using an Interactive Computer Graphics Systems. Photogrammetric Engineering and Remote Sensing, Vol 50, February 1984, 183—188.

Fraser, C., 1988. State of the Art in Industrial Photogrammetry. International Archive of Photogrammetry and Remote Sensing, 27(B5), 166—181.

Fraser, C. and Shortis, M., 1994. Industrial Inspection Using a Still Video Camera. Proceedings Resource Technology '94, Melbourne, Australia, 362—375.

Klette, R., Mehren, D. and Rodehorst, V., 1995. An Application of Surface Reconstruction from Rotational Motion. Real-Time Imaging, 1 (1995) 127—138.

Koch, R., 1992. Model-Based 3D Scene Analysis from Stereoscopic Image Sequences. International Archive of Photogrammetry and Remote Sensing, 1992, 427—433.

Kölbl, O., Chardonnens, B., Gilliéron, P.-Y., Hersch, R.D. and Lutz, S., 1991. The DSR15T A System for Automatic Image Correlation. Proceedings of ASPRS/ACSM/ AUTO-CARTO 10, Baltimore 1991.

Lee, K. and Jay Kuo, C.-C., 1992. Shape Reconstruction from Photometric Stereo. IEEE CV&IP 3, 1992, 479—484.

Marr, D., 1982. Section 3.3: Stereopsis. Vision, 111—159. WH Freeman 1982.

Peipe, J. and Schneider, C.-T., 1995. High Resolution Still Video Camera for Industrial Photogrammetry. Photogrammetric Record, Vol 15 No 85, April 1995, 135—139.

Rivett, L.J., 1983. The Assessment of Post-Operative Swelling by Analytical Photogrammetry. Australiasian Physical & Engineering Science in Medicine (1983), Vol 6 No 3, 138—142.

Rosenholm, D., 1987a. Least Squares Matching Method: Some Experimental Results. Photogrammetric Record, 12(70): 493—512 October 1987.

Rosenholm, D., 1987b. Empirical Investigation of Optimal Window Size Using the Least Squares Image Matching Method. Photogrammetria 42 (1987), 113—125.

Sheng Jin and Yunming Li, 1988. Acquiring Range Data in the Binocular 3-D Computer Vision using Edge-Based Hierarchical Matching. IAPR Workshop on CV, October 12-14, 1988, Tokyo.

Trivedi, H., 1985. A Computational Theory of Stereo Vision. IEEE 1985 277—282.

Ulupinar, F. and Nevatia, R., 1995. Shape from Contour: Straight Homogeneous Generalized Cylinders and Constant Cross Section Generalized Cylinders. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 17 No 2, February 1995, 120—135.

Ulupinar, F. and Nevatia, R., 1992. Recovery of 3-D Objects with Multiple Curved Surfaces from 2-D Contours. IEEE CV&IP 3, 1992.

Zheng, J. Y., 1994. Acquiring 3-D Models from Sequences of Contours. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 16 No 2, February 1994.

Zheng, Y., Jones, D.G., Billings, S.A., Mayhew, J.E.W. and Frisby, J.P., 1990. SWITCHER: A Stereo Algorithm for Ground Plane Obstacle Detection. Image and Vision Computing, Vol 8 Num 1, February 1990.