Some Core Issues in Design for a General Purpose Three Dimension Spatial Information System

Yuxiao Li[1], Peter Whigham[2], George L. Benwell[3] and Nick Mulgan[4]

[1]Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
Phone: +64 3 479 8301 Fax: +64 3 479 8311
Email: yli@infoscience.otago.ac.nz

[2]Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
Phone: +64 3 479 7391 Fax: +64 3 479 8311
Email: pwhigham@infoscience.otago.ac.nz

[3]Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
Phone: +64 3 479 7716 Fax: +64 3 479 8650
Email: gbenwell@nimrodel.otago.ac.nz

Compudigm International Ltd
P.O Box 10 703, Wellington, New Zealand
Phone: +64 4 499 9881 Fax: +64 499 9853
Email: nick.mulgan@compudigm.co.nz

Presented at SIRC 99 – The 11th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 13-15th 1999

ABSTRACT

The aim of this paper is to explore some core issues in design for a general purpose three dimension spatial information system (3D SIS). Space and time as structure for the existence of both entity and phenomenon is well accepted in scientific research work. Discovering the state, process and spatial relationship of these entities and phenomena constitute the main study activities in various domains. Usually, Euclidean geometry or set based geometry is employed to work as the formal abstraction of the space, e.g. in spatial information system (SIS) and computer aided design (CAD) packages. Spatial objects are represented as either discrete spatial entities distributed in the space or continuously changing phenomena filling the whole study area.

A general purpose three dimension spatial information system is a result of expansion and diffusion of SIS to other space related research domains, from megascope to microscope. A successful general purpose SIS should be built by considering the needs and requirements from various potential application disciplines, while being independent of any specific one. A SIS should facilitate individual research activities in various space-related domains by producing tools for spatial data management, manipulation, analysis and visualization. This paper addresses the core issues around the design of such a system, from conceptualization of space, spatial data model based on the conceptual model of space, to further potential spatial data analysis tools in general. Future works on feature-based spatial data structures and analysis tools built around features are suggested.

Keywords and phrases**:** conceptual model of space, spatial data model, abstraction and generalization of spatial representation, feature-based spatial data structure, and spatial data analysis

1.0 INTRODUCTION

The conceptualization of space and time can be traced back to the research of the philosophy of science (Couclelis, 1992). The concept of space and time works as a structure for ordering the spatial *variables*, entities or phenomena has a long history in scientific research in the Western tradition. How to represent the space and time structure as well as the *variables* framed by them is the central issue in the development of spatial information techniques. Associated with spatial information techniques is the issue of what kind of role do space and time play in the *state*, *relationship* and *process* of entities or phenomena. The embodying of these two issues, generalized as spatial representation and spatial data analysis based on the abstraction of reality, will unavoidably face the suitability and accuracy question. Answers to such a question are highly related to specific application domains. In fact, scientific representation of spatial objects, spatial analysis, information and knowledge retrieved can only be evaluated when they are built on and communicated within a specific representation requirement, measured data and manipulations on these measurements. Data representation for spatial entities or phenomena and further analysis operations on them will generate meaningful results if and only if the data representation and operation reflect and agree with reality.

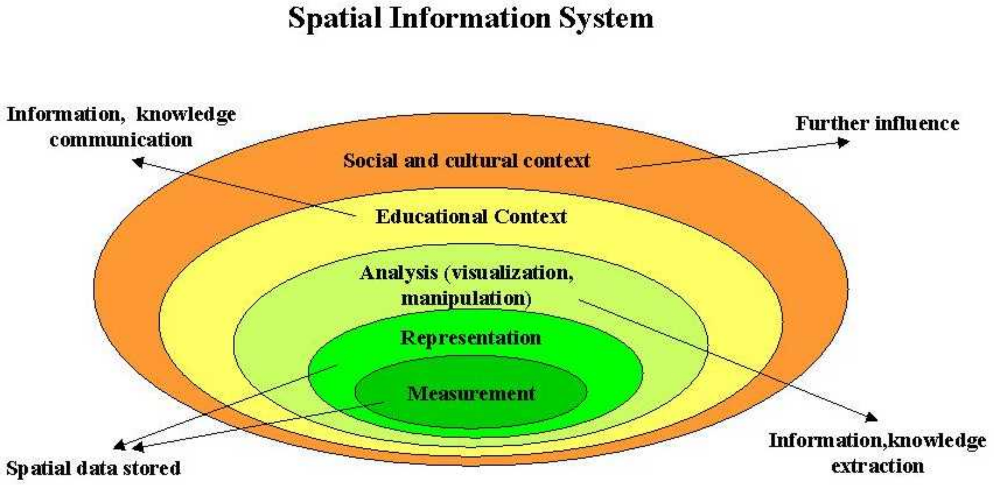

Figure 1: Spatial information system, redraw from Chrisman's sphere on organizing activities around geographical information system (Chrisman, 1996)

This paper is driven by the objective for a design and implementation of a general-purpose three-dimension spatial information system (3D SIS). Expansion and diffusion of ideas developed and accumulated in spatial information system (SIS), i.e. integrating spatial data input, storage, management, visualization and analysis in one system, to other space-related study have their practical meaning (see Figure 1). These research domains include material science industry, medical imaging, civil engineering, fluid dynamics and engineering, to a broad sensed geoscience, cover from microscopy to macroscopy and even megascopy (see Figure 2). A well-designed system can facilitate specific research work, supply new methods, help further understanding and interpretation of the research problem and assist knowledge accumulation. (McCormick *et al*, 1987).

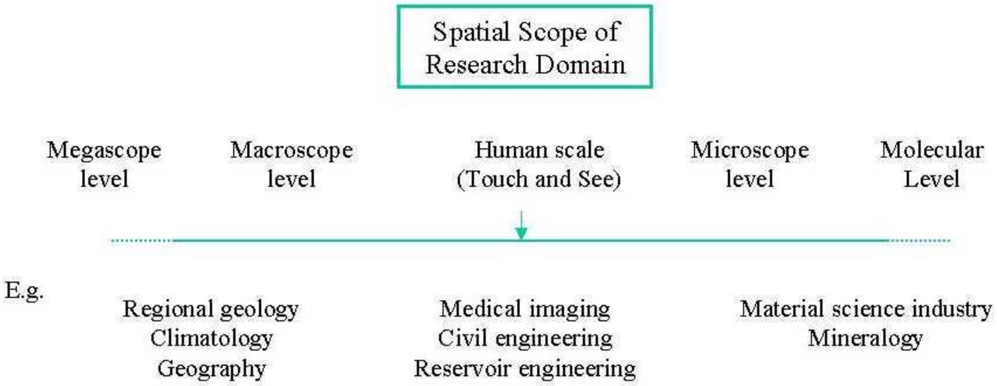

Figure 2: Scopes of space-related research domain with examples

The issue of suitability and accuracy of the spatial representation based on a conceptual model of space shows it must be addressed before considering system design and implementation. As the objective of a 3D SIS is to facilitate various space-related scientific works efficiently, the task is to examine the various requirements of spatial representation and spatial reasoning beyond a superficial complexity.

The paper organises these issues in the following structure. In section 2.0, the issues regarding a conceptual model of space are given. Two conflicting views of space are described and the weakness of each single model is emphasized. Section 3.0 considers the spatial data model. Emphasis is placed on the characterization of the abstraction of spatial representation and the appropriateness of an object-based spatial data model. Section 4.0 describes the potential analysis tools including visualization, spatial query and manipulation. Analysis tools are approached by considering the features to be dealt with, i.e., single object, group of objects with the same class or different class and subset of a thematic scene. The final section 5.0 draws conclusions of this study.

2.0 CONCEPTUAL MODEL OF SPACE

Conceptualization of space as a fundamental problem in philosophy of science effects and accompanies the development of SIS techniques. There exist two conflicting hypotheses about the world. The first one views the world like “there exist things in time and space which have (known and unknown) attributes” (Couclelis, 1992). The second one views the world like “the spatio-temporal clusters of known attributes are the things” (Couclelis, 1992). The conceptual models of space in a SIS are parallel to the two hypotheses of world with Euclidean geometry or set based geometry as the space structure. This was recognized early in 1970s by Chrisman (Worboys, 1995). Spatial objects in a SIS are either described as discrete entities or continuously changing phenomena. The entity has attributes that make it distinguished from the environment, while for the alternative approach the phenomena itself forms the attributes field.

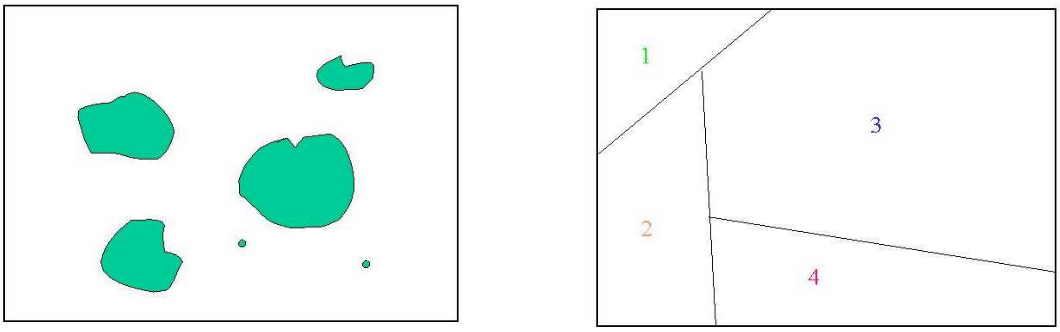

Figure 3: The space system or subset of it

Distributed with discrete entities or filled with continuously changed phenomena

Both conceptualizations of space-time get their evidence from the advancements in modern physics. From most textbooks of physics, the fundamental particle is described as having a field property. It is shown that both the *atomic* and *plenium* space, or *wave* and *particle* exist in the quantum world.

More or less affected by its cartographic origin, the world in vector SIS is represented by points, lines and polygons that seldom exist in real nature except for man-made artifacts. Like what Couclelis wrote in her article:

> *Objects in a vector GIS may be counted, moved about, stacked, rotated, coloured, labelled, cut, split, sliced, stuck together, viewed from different angles, shaded, inflated, shrink, stored, and retrieved, and in general, handled like a variety of everyday solid objects that bears no particular relationship to geography.* (Couclelis, 1992).

Vectorized spatial representation in most case restricts our understanding and modeling of relationships between different attributes of behaviours (or called variables) as these attributes typically do not have a sharp boundary with them except for the man-made artefacts.

The development of raster SIS is strongly related to the view of earth scientists, such as geologist, ecologist, hydrologist, etc., and encouraged greatly by the rapid growth of remote sensing images. Here:

> *…the phenomenon of interest is blithely bisected by the image frame… for the mindless mechanical eye everything in the world is just another array of pixels.* (Couclelis, 1992).

This absolute raster SIS creates a different set of problems. Sometimes the world needs to be abstracted by point, line polygon or polyhedron which carrying and transferring the information for the purpose of better understanding and knowledge acquisition.

More and more SIS researchers recognize the weakness of any single conceptual models. This weakness even hinders the development of SIS. A new approach to the construction of a conceptual model of world is required. This is further discussed in section 3.0.

3.0 SPATIAL DATA MODEL

Data models are the formal abstractions of reality (Burrough and McConnell, 1998). Precise, analogic, photographic representation of spatial data and information with infinite detail is expected but impossible, impractical, and sometimes useless because of the limitation of techniques and the restriction of human perception and cognition. Abstraction and generalization of any spatial feature for a specific accuracy is necessary for information interpretation, communication and knowledge acquisition. In present SIS, spatial entities or phenomena are abstracted and represented by an aggregation of points, lines, polygon and polyhedrons or array of pixels and voxels. The characteristics of the abstraction will be discussed in the following section, beginning from the consideration of data flow inside the system.

3.1 Data Flow inside the System

Spatial information systems mean different things to different people (Burrough and McConnell, 1998, Demers, 1997). In a broad sense, i.e. a system with functions of data input, storage, manipulation, analysis, visualization and output of spatial data and related non spatial data (see Figure 1).

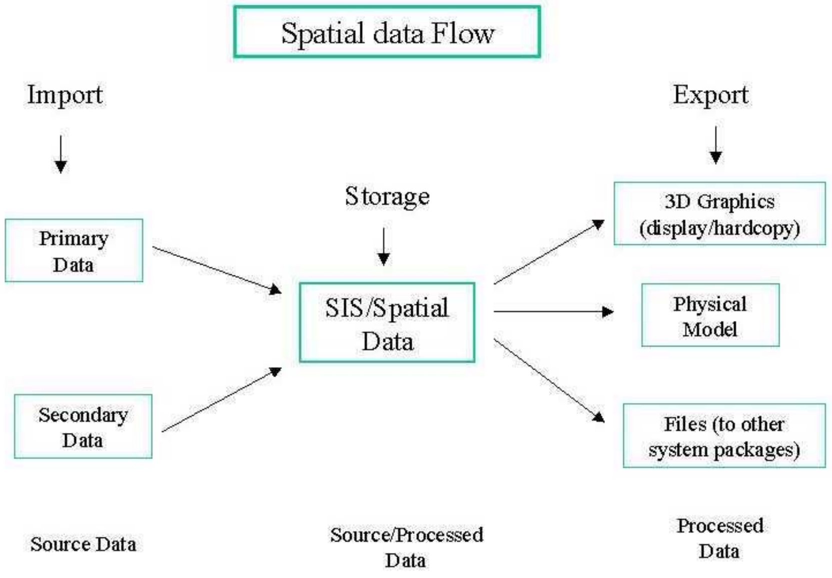

Figure 4: Data flow inside the system

The data fed into the system from all disciplines (except in CAD design, where the data is created in the design stage. Only after the design is implemented and the project is stored for management purpose, can it become the data in a SIS) for storage and analysis are generally divided into two categories: the primary data and the secondary data. The primary data category includes the data from the direct measurement, surveying data or digital images, such as images from remote sensing and medical image modalities. They are the raw data in the sense that they do not experience further processing or manipulation. But they are abstracted in the sense that they are the measured data for spatial representation, which is based on specification of *variables* of our research problem. For example, raw data from field sampling, the *variable* of elevation, will be used to build the surface of earth.

The secondary data includes any other data, which can be archived digital representation of reconstructed, modeling and simulating results. It also includes the data from digitizing the legacy map. The abstraction and generalization in this type of data is much clearer. For example, spatial data for road representations are recorded as connected series of arcs, spatial data for cities are recorded as polygons. Features extracted from medical images are represented by asymptotic triangular-net surface. Mineral grades in ore deposit are recorded in secondary data form are isosurfaces.

As both the primary data and secondary data will be the input source data for storage, further analysis, visualization and output as a physical model (in 3D) or to other system (see Figure 4), the characteristics of the abstraction is worth further consideration.

3.2 Abstraction of spatial object

3D, 2D or no space at all?

The answer to such a question, i.e. the conceptualization of space, is totally determined by the specification of *variables* for the problem in the application field. A suitable abstraction of spatial object for the *variable* is application-driven. Abstraction of variables from different research domains can differ greatly. In geographical space, the variables can be correctly represented in the 2D or 2.5D space, while 3D data representations and manipulations are needed in medical imaging, local geology, reservoir engineering, climate study and robotics (see Table 1 for the examples of various space-related domains). To the same variable, the spatial extension of study area also effect the abstraction and hence to its representation. For example, rock properties can be abstracted and represented using 2D polygons in a regional map, but in a reservoir engineering simulation study, this 2D representation is unacceptable as the characterization of subsurface rock properties are of fundamental importance.

Table 1: Example research domains with space contents

Research Directions	Geosciences	Civil Engineering	Medicine
Mega-level	tectonic geology global climate geography/ecology	road network management, urban planning	anatomy histology medical imaging
Macro-level	reservoir engineering ore deposit	architecture	physiology pathology
Micro-level	mineralogy petrology	material science & industry	cytology
Molecular-level	crystallography	chemistry/physics	biochemistry
Size of space system	approx. the earth or larger	smaller than the earth	size of human body
Coordinate system	Geo-coordinates	Geo-coordinates Cartesian	Cartesian

Inheritance and propagation of spatial abstraction

An accurate spatial representation depends on the abstraction of the spatial object, specifically speaking, the specification of attributes and relationships, their measurement and manipulation. Once the abstraction of a spatial object is built, the representation of it will inherit the concepts and transfer them to the results for further analysis. For example, a river abstracted as a line can not restore its real shape, nor is it possible to get detailed population information when it is abstracted to a city polygon. Geometric attributes of fractures for fluid flow transport in a geothermal reservoir simulation study are unavailable when the medium is treated like a continuum block if no other information is supplied. Errors associated with the measurement and manipulation of raw data also cause problems for the spatial representation as they are propagated during analysis.

3.3 Data Model and Data Structure

Allowing for various potential users from different applications all with their available data source (see Figure 2 and Table 1), attention must be paid to a choice of flexible and suitable spatial data model. This is enforced by the characteristic in the abstraction of spatial representation, as it has direct effect on the following manipulation and accuracy of analyzing results. A single vector or raster sometimes is not sufficient, as both data structures may be needed by different users. An object-oriented system is an alternative approach with some merit.

A feature-based data model suggested by several researchers (see Worboys, 1995b, Raper and Livingstone, 1995; Tang, 1996) as an alternative seems to offer some promise. The conceptual model of space is highly dependent on the research issue in the application domain. In a feature-based SIS, “…fields may not always have to be represented as data layers; object need not always be Euclidean points, lines, areas and polygons” (Couclelis, 1992). But, “…in the object-oriented approach the environment scientist must declare the nature of the real world entities identified first: their characteristics and behavior structure the spatial representation” (Raper and Livingston, 1995).

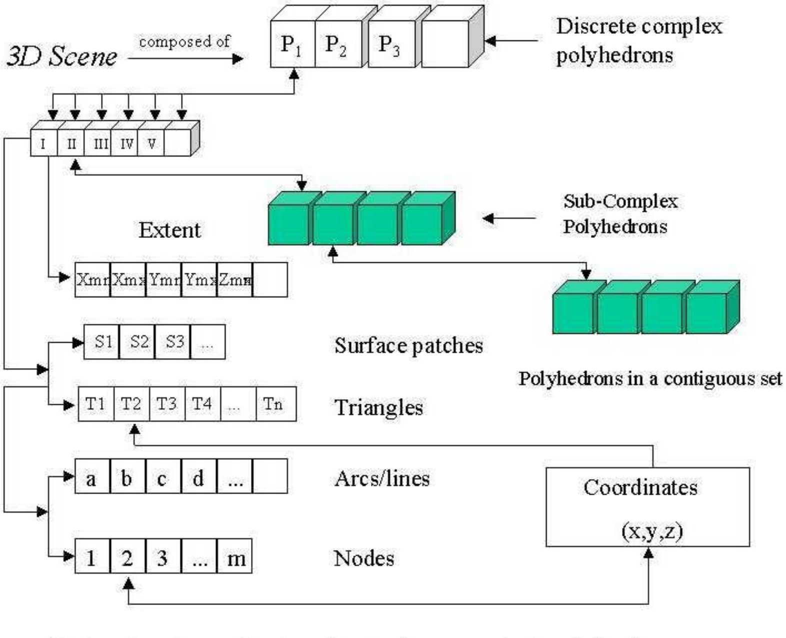

Figure 5: A possible 3D data structure of topological connected polyhedrons

Research on a hierarchical structured feature-based spatial data model with topological information is worthy of furthering consideration and testing. There are some advantages of such a system. As the definition of the feature is application specific, only the relevant characteristics are collected, stored and further analyzed. Multi-scale abstraction and generalization in spatial representation is possible from hierarchically recorded objects based on scale requirements on output and available data resolution. Spatial data analysis functions are encapsulated and independent of representation and specific application. The following section gives the description of potential analysis based on a consideration of the types of spatial features to be dealt with.

4.0 SPATIAL DATA ANALYSIS

Adoption of analysis techniques in individual research is absolutely determined by the research purpose and available source data. Data dictates in most cases the tools, like an old Chinese saying: *no matter how skillful, no woman can make meals without rice*. Techniques for spatial analysis are pointless without data.

Potential spatial data analysis tools from various application domains are generalized in the following sections with visualization, spatial query and analysis included.

4.1 Visualization

There are several issues related to visualization.

4.1.1 Co-ordinate system

A tool to easily convert coordinates system is required since users adopt different coordinates system for their spatial structure. This includes transformation between Cartesian coordinates system and geo-coordinates and between several different geo-coordinate systems, such as the geodetic coordinates system, geocentric coordinate system and Universal Transverse Mercator.

4.1.2 Visual examination or visual exploration

Visualization of results or visual exploration analysis is an issue that needs to be addressed. The choice of a static visualization or dynamic visualization forms current research objectives. Dynamic visualization means two things. One is a movie-like visualization. This type of visualization is suitable in 3D reconstruction of interested features in medical imaging application domains. The other is interactive visualization to permit the input of parameters and to check the changing of modelling and simulating results. This is more expected from computational scientists and engineers, while it is more difficult to reach, especially when manipulating a very large data set, where huge computation of analysis will need to perform before getting the computer graphics. Gahegan outlines the four barriers to this visual exploration of SIS data (Gahegan, 1999).

4.1.3 Speed and quality issue

Speed and quality of 3D graphics rendering are restricted by computer hardware and software. Multiscale based graphic reconstruction has more potential when facing very large spatial data sets. In this approach, graphics is generalized and rendered based on user specified scale for output. Data resampling considering output scale with acceptable accuracy for 3D graphics reconstruction and representation, volume or surface rendering techniques employed, can improve rendering speed through reducing the dataset need to be handled on. Visualized 3D objects may look different for different scale specifications because of the effect of generalization in the output spatial representations.

4.1 Spatial Query

Query on spatial position plus non-spatial attributes for individual objects should be available in all spatial features dealt with while accuracy will be affected by the scale based retrieving, especially near the boundary.

Distance searching is useful for visual exploration. The searched neighborhood objects can be from the same class or different classes.

As spatial query on individual 3D spatial objects is not as accurate as those to the 2D graphics because of rendering 3D graphics on a 2D screen with a point on screen related to varying values in the third dimension. Improvement on 3D query is an important research issue. It may be possible with the scale based 3D graphics display under the help of pan, zoom and rotation operations.

4.2 Spatial Data Analysis

Different research disciplines have different emphases on what kinds of tools are appropriate. Some concern single spatial feature, for example, in medical application on 3D reconstruction. Some concern spatial distribution patterns of a group of objects of the same type, such as the non-destructive evaluation in product quality control in forestry industry. Some concern the aggregation of objects from different classes. Some research directions concern the processes of the spatial feature, e.g. the modeling a petrol and oil reservoir, or the underground hydrogeology.

4.3.1 Single object

Analysis on a single feature may simply include the calculation of size, shape, orientation and shape complexity, such as the calculation of surface area and volume in 3D, perimeter and area in 2D. It is not a simple task when high accuracy is required. For example in medical imaging, these parameters based on 3D reconstruction are meaningful for diagnostic purpose. Inside components and texture, outside environment and structure may also be of interest if these are represented.

4.3.2 Group of objects from the same class

The analysis of this type of spatial feature is usually concerned with spatial distribution patterns. It is noted that the existence of pattern is highly scale based. The patterns appear only in a limited spatial extent. The techniques for such analysis include distance-based techniques and density based approaches, such as nearest neighbour distance analysis (NND) and Repley’s K.

4.3.3 Aggregation of objects from different classes

The analysis techniques in this category are concerned with spatial connectiveness and adjacency. Spatial autocorrelation analysis may also be included here. These techniques are well presented in ecological study, such as patch and patchness analysis and the calculation of indices of spatial association. It is also seen in mineralogical study. An example can be found in work of Sardini *et al* (Sardini, *et al*, 1999).

4.3.4 3D Thematic scene

Spatial data analysis in this category composes the largest, most attractive and complex part. It includes both 3D reconstruction (state of spatial object) from image (the image is seen as density thematic), sampling data and other modelling and simulating approaches which try to capture the developing process of spatial entity or phenomena. 2D surface can be considered as the special case with the value changing in the third dimension as zero.

In data rich applications, such as remote sensing and medical imaging, data analysis tools required are basically two types: image registration and feature segmentation (or classification in 2D environment).

In data limited applications, such as subsurface characterisation of rock properties, interpolation techniques are needed, including global trend surface interpolation, local deterministic method (Voronoi diagram), inverse distance interpolation, splines (Hutchison, 1995), and optimal geostatistic method (kriging) (Burrough and McDonnell, 1998).

When having more than one *variable* for a problem of study at hand, several thematic scenes need to be handled. Usually, deterministic and probabilistic modeling approaches which combine prior knowledge on the relationship of these thematic scene are also required in some research application domains (Cox *et al*, 1997). While complex processes, e.g. in study of the complex fluid dynamic in reservoir engineering, stochastic process simulation techniques are required. Although specific data analysis functions will be developed under cooperation with specialist and experts in specific domains, general tools for feature extraction in image processing, deterministic modeling and regression analysis can be considered by supplying the cell-based basic arithmetic /trigonometric operations. Interface in the SIS for integration with those well developed modeling packages may offer some promises. (Turner, 1989)

5.0 CONCLUSION

The goal that the SIS as a tool to facilitate various scientific researches with appropriate analysis functions requires a correct development direction. A successful system needs to comprise the needs and requirements from individual research domains while still allowing useful generalized operations. Different requirements on spatial representation and spatial data analysis create the difficulty in the system design, while they also join these various domains together. A suitable spatial data model, which facilitates spatial representation under different abstractions of space is a key issue, especially when dealing with the requirements from various application fields. A spatial data model built on an appropriate conceptual model will earn a long life for the system. Feature-based spatial data structure seems to offer promise and will be further researched.

Powerful, easy to use spatial data analysis techniques based on features recorded will also give the system a broader range of possible applications. Improvement of performance of the system will benefit from adoption and absorption of advance from computer science, e.g. the computer graphics and spatial database management and spatial data analysis techniques.

ACKNOWLEDGEMENTS

We thank the Foundation for Research, Science and Technology (FRST), Technology New Zealand and Compudigm International for their assistance in supporting this research. A special thanks is also given to Dr. Jochen Albrecht, University of Auckland for his earlier assistance in developing some of these ideas.

REFERENCES

AGCRC, (1997), Homepage of AGCRC - Australian Geodynamics Cooperative Research Centre. Accessed 10 September, 1999, <http://www.agcre.csiro.au>

Burrough, P. and McConnell, R.A, 1998, *Principles of Geographical Information Systems,* Oxford University Press

Chrisman, N., 1996, *Exploring geographic information systems,* John Wiley & Sons, Inc., New York

Couclelis, H, 1992, People manipulate objects (but cultivate field): beyond the raster-vector debate in GIS, *Theories and methods of spatio-temporal reasoning in geographic space*, Frank, A.V, Campari, L and Formentini, U. (eds), Springer Verlag, Berlin, pp65-77

Demers, M. N, 1997, *Fundamentals of geographic information system*, John Wiley & Sons, Inc, New York

Gahegan, M, 1999, Four barriers to the development of effective exploratory visualisation tools for the geosciences, *Int. J. Geographic Information Science*, Vol. 13, No. 4, pp289-309

Hutchinson, M.F., 1995, Interpolating mean rainfall using thin plate smoothing splines, *Int. J. Geographical Information Systems*, 1995, Vol., 9, No. 4, pp385-403

McCormick, B.H.*et al*, 1987, Visualization in scientific computing, *Computer Graphics*, Vol. 21, No. 6

Sardini et al, 1999, Primary mineral connectivity of polyphasic igneous rocks by high-quality digitisation and 2D image analysis, *Computer & geosciences*, 25, (1999), pp599-608

Raper, J and Livingstone, D, 1995, Development of a geomorphological spatial model using object-oriented design, *Int. J. Geographical Information Systems*, Vol. 9, No 4, pp359-383

Tang, A.Y. *et al*, 1996, A spatial model design for feature-based geographical information systems, *Int. J. Geographical information systems*, Vol 10, No 5, pp643-659

Turner, A.K, 1989, The role of three-dimensional geographical information systems in subsurface characterization for hydrogeological applications, in: *Three dimensional application in geographic information systems*, Raper, J. (eds), Taylor & Francis, Chapter 8, pp125-127

Worboys, M.F., 1995, *GIS: A computing perspective*, Taylor & Francis, London

Worboys, M.F., 1995b, Object-oriented approaches to geo-referenced information, *Int. J. of Geographical Information System*, Vol. 8, No.4, pp385-399.