

SPRING: INTEGRATING REMOTE SENSING AND GIS BY OBJECT-ORIENTED DATA MODELLING

GILBERTO CÂMARA, RICARDO CARTAXO MODESTO SOUZA,

UBIRAJARA MOURA FREITAS, JUAN GARRIDO, FERNANDO MITSUO II

Image Processing Division (DPI),

National Institute for Space Research (INPE), Brazil

1. INTRODUCTION

Brazil, being a large and diverse country, faces a great number of issues related to the handling of its natural resources. Ranging from deforestation of tropical forests in the North to the impacts of rapid urbanisation in the Southeast, and including drought assessment and management of water resources in the Northeast, country planners face challenges at a continental scale.

The need for GIS technology in Brazil is best enlightened by considering the problem of monitoring deforestation on the Amazon forest. Arguably, no single environmental issue has captured so much international attention in recent years. The process of land use and cover change detection in the region, carried out by Brazil's National Institute for Space Research (INPE), relies on Remote Sensing and GIS

technology as the only possible way to cover more than 5,000,000 km² (the area of Legal Brazilian Amazonia).

These environmental issues have motivated significant investments in the use and development of Remote Sensing and GIS technology at INPE. The SGI/SITIM system, developed by INPE for IBM Pcs from 1982 to 1991, is currently in use in the majority of the country's Remote Sensing laboratories. In 1991, we started the development of SPRING, a software designed to meet Brazil's challenges on natural and human resources monitoring, with the following design objectives:

- Operate as a *seamless geographical data base*, with a large volume of data, without being limited by tiling schemes, scale and projection. Object identity should be maintained on the whole data base.
- Support both raster and vector data geometries and *integration of remote sensing data into a GIS*, with functions for image processing, digital terrain modelling, spatial analysis and data base query and manipulation.
- Achieve *full scalability*, that is, be capable of working with full functionality from desktop PCs running Windows or OS/2 to high-performance UNIX workstations.
- Provide an easy-to-use, yet powerful environment, with a combination of menu-driven applications and a spatial algebra language.

This paper describes the general features of the SPRING software and is organised as follows. Section 2 includes a brief discussion of previous work, showing the main areas of innovation of the project. The object-oriented data model used in SPRING

is presented in Section 3 and the LEGAL query and manipulation language is part of Section 4. Section 5 contains a review of the functionality available in version 2.0 of the systems and Section 6 describes the project status.

2. RELATION TO PREVIOUS WORK

There is a general consensus on GIS literature about the two broad classes of models of geographic information: field-based and object-based [7]. While the former deals with spatial distributions over a geographical region, the latter deals with discrete, identifiable entities on the geographical space [8]. These two approaches have led to two separate areas of research: definition of spatial operations on discrete objects [9], and definition of operations on fields (the “map algebra” of Tomlin [10]). Furthermore, most current GIS implementations provide different sub-systems for map algebra, spatial queries and image processing.

To improve upon this situation, there is a need for a comprehensive approach which unifies the field-based and object-based models, and provides the basis for answering questions such as: How can a set of discrete objects be converted into a field ? Are there operations which generate a set of objects from a field ? What are the acceptable formats for fields and objects? How to design a general query and manipulation language?

Another important issue is the integration of remote sensing data into a GIS. It is necessary to include such data into a general GIS data model, and to provide suitable algorithms for information extraction on images.

In order to answer these questions, our approach was to derive, initially, an object-oriented data model for GIS data which combines the ideas of “fields” and “objects” [1]. Based on the model, we investigated the operations over *geo-fields* and over *geo-objects*, and the *transformations between geo-fields and geo-objects* [2]. The definition of operations was used to derive a query and manipulation language, an user interface and the different functions which comprise SPRING. Innovative algorithms for spatial indexing [3], image segmentation [4], region-based classification by neural networks [5] and TIN generation [6] enable improved performance on environmental assessment problems and the simplify the integration of remote sensing data into the GIS environment.

3. AN OBJECT-ORIENTED DATA MODEL FOR GIS DATA

In order to derive an object-oriented model, it is necessary to distinguish between the *conceptual level* (where the abstract entities are defined) and the *implementation level* (where the graphical representations are constructed).

3.1 The conceptual level

We take the perspective that an environmental GIS should consider both *fields* and *objects* as abstract data types, and we have tried to find useful generalisations and specialisations from these basic notions. The conceptual level of the model is shown in Figure 1.

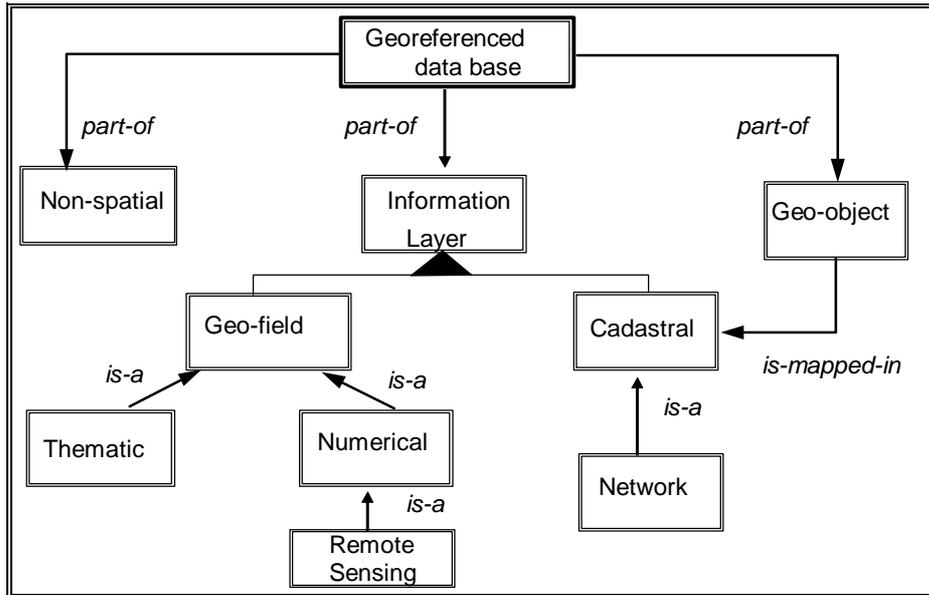


Figure 1 - The conceptual level of SPRING's data model.

In our definitions, we consider a geographical region R to be a part of the surface of the Earth, represented in a suitable cartographical projection. Given a geographical region R , we define a *geographical field* (or *geo-field*) as an object $f = [R, \lambda, V]$, where λ is a mapping between locations in R and values in V .

A *geo-field* represents a geographical variable which takes values over all locations of a region of the Earth. We shall denote the class of geo-fields by GEO-FIELD¹, which can be specialised into the classes THEMATIC (when V is a finite denumerable set - the *themes* of the map) and NUMERIC (when V is the set of real values). To integrate remote sensing data into the GIS environment, we define the REMOTE SENSING class as a

¹Throughout the text, model and derived classes are indicated in SMALLCAPS.

specialisation of the NUMERIC class, when V is a set of discrete values obtained by quantization of the response obtained by an active or passive sensor.

Given a set of geographical regions R_1, \dots, R_n , and a set of attribute domains A_1, \dots, A_m , we define a *geographical object* (or *geo-object*) as an object $go = [a_1, \dots, a_m, r_1, \dots, r_n]$, where $a_i \in A_i$ is the value of its attribute in the attribute domain A_i and $r_i \in R_i$ indicates the geometrical representation associated to the geo-object in the geographical region R_i . Therefore, a geographical object is a unique element that has descriptive attributes and which can be represented on many different maps.

To support the idea of multi-scale and multi-tile representations assigned to the same geo-object, we introduce the concept of *cadastral objects*. Given a geographical region R , a set of geo-objects GO , a *cadastral object* is an object $co = [R, GO, geo]$, where geo is a mapping which assigns, for each geo-object $go \in GO$, a representation $geo(go)$ in R .

A cadastral object a is complex geo-object which groups together the mapping of geo-objects for a given geographical region. The relation between a geo-object and a cadastral object is IS-MAPPED-IN, which defines, for a given geo-object, its mapping in the geographical region R .

To illustrate the concept, consider figure 2, which illustrates a data base for rivers of the Brazilian Amazonia. The location of two rivers (Amazonas and Xingu) is indicated, for discussion purposes only. Since the region covers a very large area, a geographical data base in the 1:250000 scale (on UTM projections) will consist of several non-overlapping tiles. However, we would like that the object denoted by

“Amazon river” is seen by the data base as a single object, even though its representation may span several maps. To that end, we associate each UTM partition to a cadastral object which includes a mapping for all which are included in the geographical area of the partition.

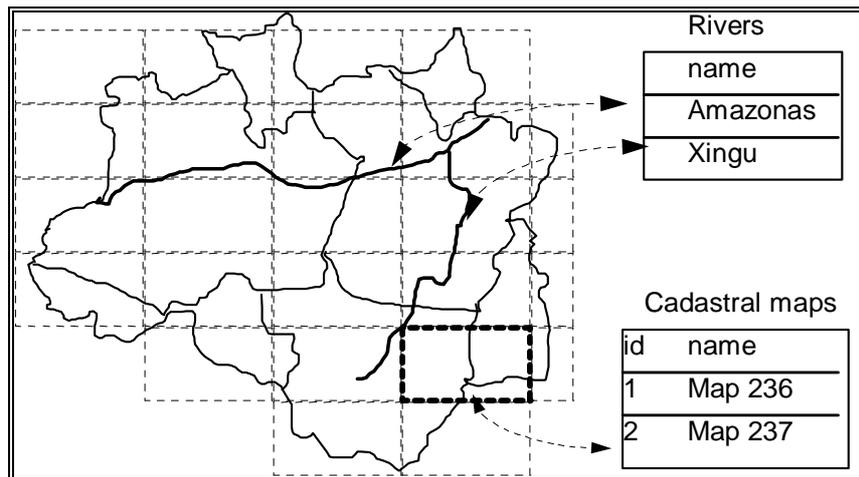


Figure 2 - Geo-Objects and Cadastral objects.

We consider one specialisation of the CADASTRAL class: the NETWORK class, whose instances store the arc-node topology and correspond to the linearly connected structures of a network.

Since *geo-fields* and *cadastral objects* have in common the property of geographical location and projection, it is useful to define a *container* element, an abstract entity which simplifies the user interface. To that end, we define the information layer class (or INFOLAYER for short) as a generalisation of the classes GEO-FIELD and CADASTRAL. An *infolayer* represents, for a given geographical region, the values of a *geo-field* or the geographical location of a set of *geo-objects*. Our model also includes

the NON-SPATIAL OBJECT class, which caters for non-georeferenced data which may be aggregated to a GIS.

The process of obtaining the conceptual schema of a geographical database consists of *specialising* the basic model data types. As an example, consider the definition of the database schema for a *land use information system*, which might contain: (i) a FARM class, specialisation of the GEO-OBJECT class, which may be further specialised into classes such as CATTLE RANCH, AGRICULTURAL FARM and NON-PRODUCTIVE PROPERTY; (ii) a FARM MAP class, specialisation of CADASTRAL, which includes the graphical representations for the objects of the FARM class and its specialisations; (iii) a SOIL MAP class, specialisation of THEMATIC, whose instances are geo-fields that store the different soils map for the study area; (iv) the classes TOPOGRAPHIC MAP, SLOPE MAP and ASPECT MAP, specialisations of the NUMERIC class, whose instances are geo-fields that store topography, slope and aspect maps for the study areas. (v) a LANDSAT TM class, specialisation of the REMOTE SENSING class, whose instances store satellite images over the study area.

3.2 The implementation level

The model provides, for system implementation, the classes VECTOR REPRESENTATION and RASTER REPRESENTATION, and enables the association of both representations to the same information layer. We consider the following specialisations of VECTOR REPRESENTATION: PLANAR SUBDIVISION (a set of arc, nodes and polygons used for thematic fields in vector format), ORIENTED GRAPH (set of arcs and nodes connected on a graph), CONTOUR LINES (set of lines, each associated to a z value), 3D

POINTS and TRIANGULAR GRID. We also specialise the RASTER REPRESENTATION class as follows: REGULAR GRID, GREY-LEVEL IMAGE, THEMATIC IMAGE (raster representation of a THEMATIC field). Table 1 shows the relation between the conceptual and implementation levels. Note that the user will mostly deal with conceptual level classes, whereas the system developers will use the implementation level classes.

TABLE 1

REAL-WORLD	CONCEPTUAL LEVEL	IMPLEMENTATION LEVEL
Geological Map	Thematic Geo-Field	Vector (planar subdivision) or Raster (thematic image)
Altimetry	Numeric Geo-Field	Vector (contour map, triangular grid, 3D points) or Raster (regular grid)

RELATION BETWEEN CONCEPTUAL AND IMPLEMENTATION LEVEL

LANDSAT Image	Remote Sensing Geo-Field	Raster (grey-level image)
Urban Parcels	Geo-objects	
Parcel Map	Cadastral	Vector (arc-node-polygon)
Electricity Network	Network	Vector (oriented Graph)

4. LEGAL - A QUERY AND MANIPULATION LANGUAGE

LEGAL is a general spatial language which includes *boolean*, *mathematical*, *local* and *zonal* operations over geo-fields, *spatial selection* and *spatial join* operations over geo-objects, transformations from geo-objects to geo-fields (such as *attribute reclassification* and *buffer maps*) and from geo-fields to geo-objects (*spatial interpolation and identification*). Its syntax is based on the OQL (“Object Query Language”) standard proposal for object-oriented data base management systems [11]

In LEGAL, *spatial queries* over geo-objects are implemented using OQL extensions, and *manipulation operations* over geo-fields and *transformations* between geo-fields and geo-objects are implemented by operators at the same semantic level as OQL statements.

LEGAL is *strongly typed*. Each data type belongs either to one of the basic data classes (THEMATIC, CADASTRAL, REMOTE SENSING, NUMERIC, NETWORK, OBJECT), to a specialisation of these classes, or to a complex object type which stores the result of a spatial query (the COLLECTION class).

To indicate how the language operates, we present here an example of a site selection study for a waste disposal facility based on three inputs: soil type, topography and the hospital location. The operation requires a spatial query (“select all hospitals located in the borough of Green Village”), a transformation operation (“calculate a distance map based on the hospitals”), a numerical formula (“determine the slope of a given topography”) and a boolean analysis (“suitable areas have non-permeable soils, with less than 5% slope and are located at more than 10 km from a hospital”). This

operation in LEGAL is shown below, assuming a geographical data base schema with the following classes: HOSPITAL, BOROUGH (specialisations of OBJECT), ROAD_MAP (specialisation of CADASTRAL), SOILS_MAP, SUITABILITY_MAP, DISTANCE_MAP (specialisations of THEMATIC) and TOPOGRAPHY and SLOPE_MAP (specialisations of NUMERIC).

```

COLLECTION      hospitals_SJC      HOSPITAL;
CADASTRAL       roads_SJC         ROAD_MAP;
THEMATIC        soil_SJC          SOILS_MAP;
THEMATIC        suit_map         SUITABILITY_MAP;
THEMATIC        distance_hosp     DISTANCE_MAP;
NUMERIC         topo_map         TOPOGRAPHY;
NUMERIC         slope_sjc        SLOPE_MAP;

roads_SJC = (SELECT map FROM map in ROAD_MAP
             WHERE map.name = "Sao Jose dos Campos");
hospitals_SJC = (SELECT hosp
                FROM hosp IN HOSPITAL ON MAP roads_SJC,
                borough IN BOROUGH ON MAP roads_SJC
                WHERE borough.name = "Green Village"
                AND hospital.geometry IN borough);
distance_map = DISTANCE (hospitals_SJC,
                        CLASSES = {[0-5 km], [5-10 km], [>10 km]});
topo_map = (SELECT map FROM map in TOPOGRAPHY
            WHERE map.name = "Topography_SJC");
soil_map = (SELECT map FROM map in SOIL_MAP
            WHERE map.name = "Soils_SJC");
slope_map = SLOPE (topo_map);
suit_map = NEW (NAME = "WasteDisposal")
            WITH RASTER REPRESENTATION (RES_H = 50, RES_V = 50);
suit_map = SWITCH (
    "Good": soil_map.class = "non-permeable"
            AND slope < 5%
            AND distance_map.class = "[> 10 km]";
    "Bad": OTHERWISE);

```

5. SYSTEM FUNCTIONALITY

5.1 *User Interface*

Use of the model concepts has enabled the design of an user interface which allows manipulation of geographical data at an abstract level. The interface allows *definition of the geographical database schema* and *selection of spatial analysis operations*. When a user selects an *infolayer* which has been specialised from one of the classes of the database schema, only the operations available for that specific data type are made available to him. This approach reduces to a large extent uncertainty in the choice of valid functions.

5.2 *Spatial Data Base Management*

All the descriptive attributes of the geo-objects and geo-fields are stored on a data base management system. The data base interface supports commercial systems, such as Ingres and ORACLE. SPRING manages 12 cartographic projections. Facilities for data management, projection conversion, and raster and vector mosaicking images are available.

5.3 *Data Entry*

Vector maps can be digitised and edited on tablets or on the screen, with automatic creation of topological information. *Digital terrain models* can be created by digitising irregularly-spaced points or by sampling contour lines, with support for both regular and triangular grids. *Remote sensing image* geocoding can be made by means of ground control point location, and images can be registered with maps or with other

images. *Raster-to-vector* and *vector-to-raster* conversions enable the mapping between the available formats. SPRING also *imports* and *exports* data from a number of formats, including: ARC/INFO, DXF, SPANS, TIFF, ERDAS, PCI, MaxiCAD and SGI/INPE.

5.4 Image Processing

Facilities for digital image processing include contrast enhancement, spatial filtering, radiometric correction, arithmetic operations, image statistics, maximum-likelihood (statistical) and segmentation classifiers and a specific module for radar images. SPRING uses a *segmentation* algorithm which is based on region-growing methods [4]. After the segmentation is performed, various region classifiers can be executed: clustering, maximum-likelihood and neural networks [5]. Results of projects conducted at INPE have shown a marked improvement of region-based classifiers over traditional methods for environmental tasks such as tropical forest deforestation mapping [12]. Colour Plate 2 shows the results of the segmentation on a LANDSAT image (bands 3,4 and 5) over an area of the Amazon forest (the city of Rio Novo).

5.5 Geographical Analysis and Digital Terrain Modelling

The analysis is by means of the LEGAL language or by using in-built functions, which include *calculation of area, perimeter, distances and angles, generation of buffer maps*. *Spatial data base query* operations can be expressed in LEGAL or by means of an interactive interface. Digital terrain models (DTMs) are stored in *regular* and *triangular grids*, and can be shown as contour lines, grid point values and by 3D visualisation. DTMs can be analysed, including calculation of *slope* and *aspect maps*, and transformation into thematic maps or images.

5.6 Map Composition and Plotting

This module enables interactive map composition and plotting, with a WYSIWYG interface. A symbol library is maintained in the DXF format, which enables easy addition by the user. Complete control over graphical elements is possible (size, position, colour and slant). Output devices supported include Postscript and HPGL/2, two standards widely accepted by the industry. Colour Plate 2 shows the result of a map generation.

6. PROJECT STATUS

SPRING is being developed by INPE (the Brazilian National Space Institute), in co-operation with EMBRAPA (the main agricultural research organisation in Brazil) and the IBM Rio Scientific Centre. The development language for is C++, and the system works on UNIX, OS/2 and Windows platforms. The SPRING project was initiated in 1991, the first version was delivered on April 1994, and the second version is scheduled for April 1996. The current version runs to 200,000 LOCs in C++ and over 700 pages of documentation [13]. More than 100 man-years have been so far invested in SPRING.

The algorithms for spatial indexing [3] and classification by neural networks [5] have been developed by the IBM Rio Scientific Center.

Researchers interested in a working copy of SPRING for evaluation purposes may write to Gilberto Câmara, head, Image Processing Division, INPE (email: gilberto@dpi.inpe.br). Further information on SPRING will be found on WWW at the address *<http://www.inpe.br/spring/home.html>*.

7. ACKNOWLEDGEMENTS

SPRING is the result of a team effort. The system architects are Ricardo Cartaxo Modesto de Souza and Ubirajara Moura Freitas and the project manager is Juan Carlos Pinto de Garrido. The data model and the LEGAL language have been designed by Gilberto Câmara, Ubirajara Moura Freitas from INPE and Marco Antônio Casanova and Andrea Hemerly from IBM Brazil. SPRING's development team includes:

At INPE: Ana Paula Dutra de Aguiar (1992-1994), Carlos Felgueiras, Cláudio Clemente Barbosa, Eduardo Camargo, Fernando Mitsuo Ii, Fernando Yutaka Yamaguchi, Guaraci José Erthal, Eugenio Sper de Almeida, João Argemiro de Carvalho Paiva, João Pedro Cordeiro, João Ricardo Freitas Oliveira (1991-1993), José Cláudio Mura, José Zamith (1993-1994), Júlio Cesar Lima D'Alge, Laércio Namikawa, Lauro Hara, Leila Garcia, Leonardo Bins (1991-1995), Marina Ribeiro, Marisa da Motta, Mary Minamoto, Misae Yamamoto, Silvia Shizue Leonardi, Sergio Rossim and Virginia Correa. Silvana Amaral, Flavia Nascimento, Regina Bruno, Maycira Costa (1992-1994) and Lygia Mammana (1992-1994) have assured user support, quality control and user documentation. Diógenes Alves, José Carlos Moreira, Eliana Kalil have conducted extensive applications of SPRING for land use and cover change surveys in Amazonia.

At IBM Rio: Marco Casanova, Andrea Hemerly, Paulo Souza, Alexandre Plastino (1991-1992), Mauricio Mediano (1992-1994), Claudia Tocantins (1993-1995).

At EMBRAPA: Ivan Lucena, Moacir Pedroso.

The Brazilian National Research Council (CNPq) has also provided important support for SPRING, through the programs RHAЕ and ProTem/CC (“GEOTEC project”).

8. REFERENCES

1. G. Câmara, *Models, Languages and Architectures for Geographical Information Systems*. Ph.D. Thesis in Applied Computer Science (in Portuguese). São José dos Campos, INPE (1995).
2. G. Câmara, U. Freitas and M.A. Casanova, Fields and Objects Algebras for GIS Operations. *Proceedings of III Brazilian Symposium on GIS*, 407-424 (1995).
3. M. Mediano, M.A. Casanova, M. Dreux, V-tree - a storage method for long vector data. In: *Proceedings of the 20th. Conference on Very Large Data Bases*, Santiago, Chile (1994).
4. L.Bins, G.J. Erthal and L.M.G. Fonseca, A method for non-supervised region classification. In: *Proceedings of VI Brazilian Symposium on Computer Graphics and Image Processing*, 65-68 (1993). (in Portuguese).
5. V. Barbosa, R. Machado and F. Liporace, *A neural system for deforestation monitoring on Landsat images of the Amazon region*. IBM Rio Scientific Center, Technical Report CCR-157 (1993).
6. C.A.Felgueiras and M. Goodchild, *An incremental constrained Delaunay triangulation*. NCGIA Technical Report 95-2 (1995).

7. M. Goodchild, Geographical data modeling. *Computers & Geosciences*, 18(4): 401-408 (1992).
8. M.F. Worboys, *GIS: A Computing Perspective*. Taylor and Francis, London (1995).
9. M.J. Egenhofer and R.D. Franzosa, On the equivalence of topological relations. *International Journal of Geographical Information Systems*, **9**(2), 133-152 (1995).
10. D. Tomlin, *Geographic Information Systems and Cartographic Modelling*. Prentice-Hall, Englewood Cliffs, NJ (1990).
11. R. Catell (ed.), *The ODMG-93 Standard*. Morgan Kauffmann, San Francisco (1993).
12. G.T. Batista, J.C. Moreira and E.K. Mello, A new method for Amazon deforestation mapping. *Proceedings of ISPRS Comission VII Symposium*, Rio de Janeiro (1994).
13. G.Câmara and A.P. Aguiar (eds.), *SPRING Programming Manual*. São José dos Campos, INPE (1996).