

An Extended Representation Model for Geographic Data

J. Raul Ramirez, Ph.D.
The Ohio State University Center for Mapping
raul@cfm.ohio-state.edu

Abstract

This paper extends the traditional representation model for geographic data. The traditional representation model is based on a two-dimensional, static, time-independent world, where only well defined or continuous objects exist. This of course is a very limited representation of the real world. The extended model considers a multi-dimensional, dynamic, time-dependent world where objects can be well-defined (with crisp boundaries), poorly defined (with fuzzy boundaries), continuous or discontinuous surfaces. The extended model goes beyond the traditional idea of maps as schematic representations of the environment, and provides the framework for realistic multi-media visualization tools: the maps of the future.

1. Introduction

How we perceive and represent the environment is an issue of great interest in Geographic Information Sciences. Burrough (1996) presents a version of the traditional representation model (Figure 1). This model assumes a two-dimensional, static, time-independent world and limits the representation of our perception of reality to two cases: the exact object, and the continuous field. The exact object has been implemented using the graphic model of vector structures. The continuous field has been implemented using the graphic model of raster structures. The result is computer based visualization tools that, at best, show data similarly to paper maps.

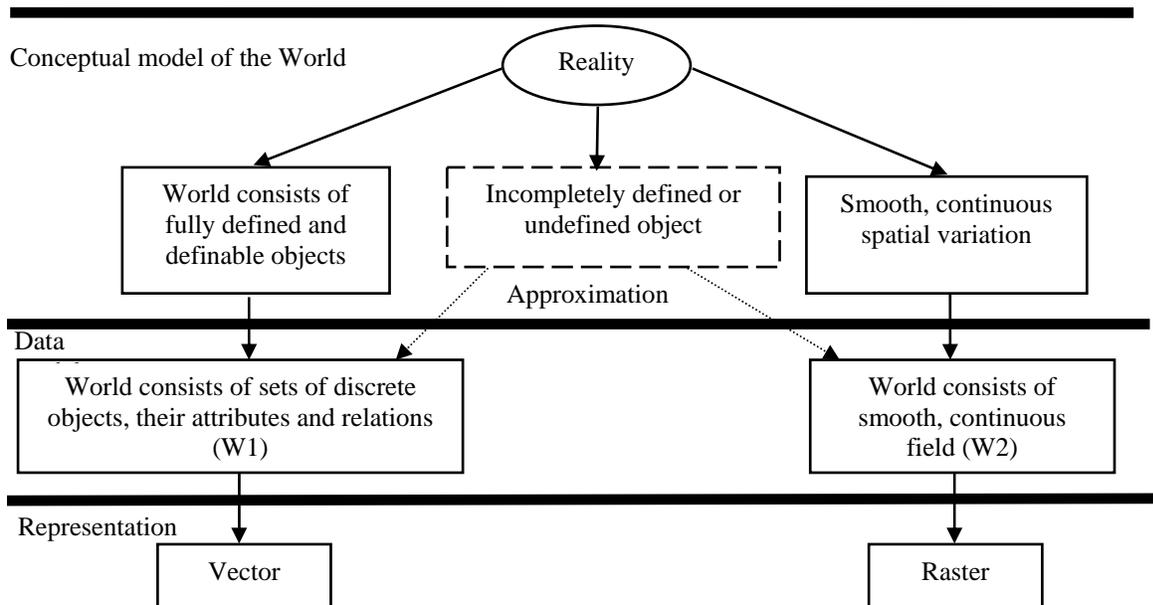


Figure 1. Main Steps from Perception of Reality to GIS (adapted from Burrough, 1996).

In the traditional model there are three types of world entities: fully defined objects, smooth field variations, and incompletely defined and/or undefined objects (for example, the fuzzy outline of a road) but only two representations: vector (for fully defined objects), where edges and corners are crisp and well defined, and raster (for smooth field variations), where continuous fields can be represented. If the phenomenon belongs to neither, then there are really only two options: vector or raster representation. Therefore, any undefined or incompletely defined object must be approximated as a well-defined object or as a field.

Burroughs (1996) recognizes that “as knowledge about natural phenomena has progressed it has become clear that these simple ‘object/field’ models of reality are often inadequate to convey the detailed information that is required to support investigations of how different aspects of the environment interact.” For example, boundaries of forests or soil units are not necessarily homogeneous, man-made structures are mostly three-dimensional, water in rivers is in continuous movement, etc. The traditional models hardly account for these characteristics of geographic data.

We believe that the framework for the traditional model was based on the way geo-spatial data was acquired: by surveying techniques (vector) or by remote sensing techniques (raster). Both were done separately, for an instant of time. Today, we are collecting multi-dimensional, multi-spectra, multi-quality, dynamic, and other anecdotal types of geo-spatial data in an integrated fashion. Therefore, current representation models need to be expanded to accommodate the new capabilities and hence the capability to store and retrieve those data types. We propose to develop an expanded representation model and to assemble the appropriate mathematics and cartographic principles to support its implementation.

2. The Concept of an Extended Model

At the Center for Mapping we have worked on an expanded model that we believe better fits how the environment is perceived. This unpublished model is shown in Figure 2 (W1 and W2 were defined in Figure 1). This model assumes that representational models are not limited to raster and vector but are part of a continuum whose extremes are raster and vector representations. We believe that raster or vector alone does not allow, in general, a realistic representation of the environment, and that a combined representation is a better approach. In some ways, this agrees with modern capabilities of data collection. Today, collection systems, such as Mobile Mapping, are collecting raster and vector data, simultaneously, in real time. Airborne, multi-sensor platforms are other examples of the simultaneous collection of datasets of different natures.

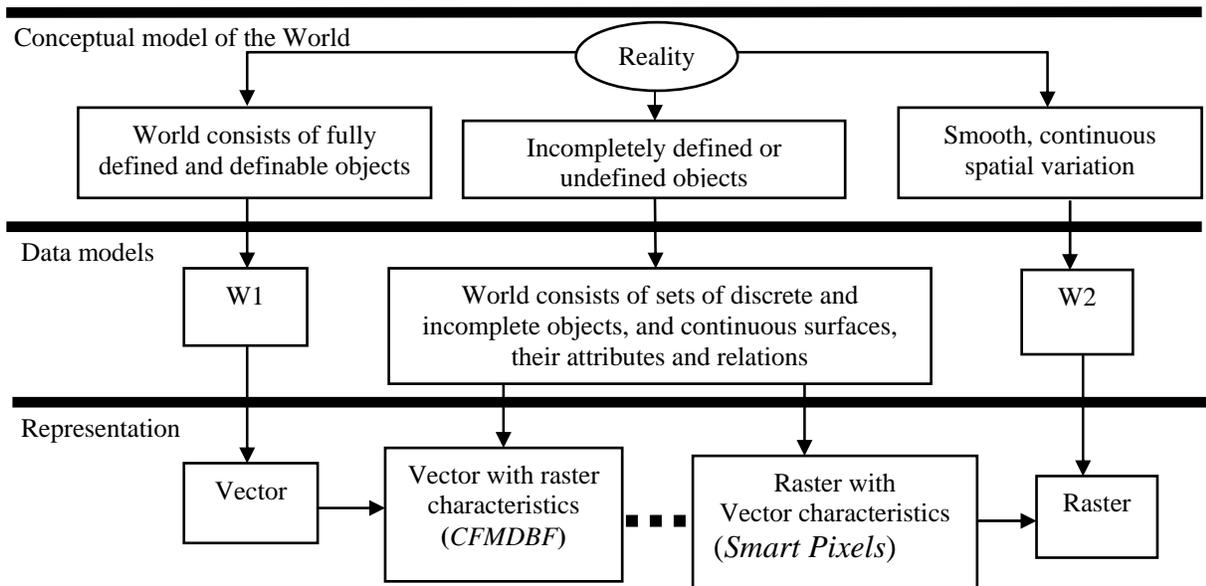


Figure 2. Main Steps from Perception of Reality to GIS Representation.

As part of our research at The Ohio State University Center for Mapping, out of the continuum of possible data models, we have concentrated on the development and implementation of two:

- (1) A vector-based model with some raster characteristics: the Center for Mapping Database Form (CFMDBF), and
- (2) A raster format with some vector characteristics: the Smart Pixel format. The development of the smart pixel format is an ongoing effort.

In the CFMDBF, features are stored as attributed vectors, plus they are rasterized using a fixed-size grid and this representation is stored as connected raster features using Freeman Chain Code. There is also a grid table that keeps track of all features located in each grid mesh. CFMDBF was used by Bidoshi (1995) to compare, update, and conflate road data.

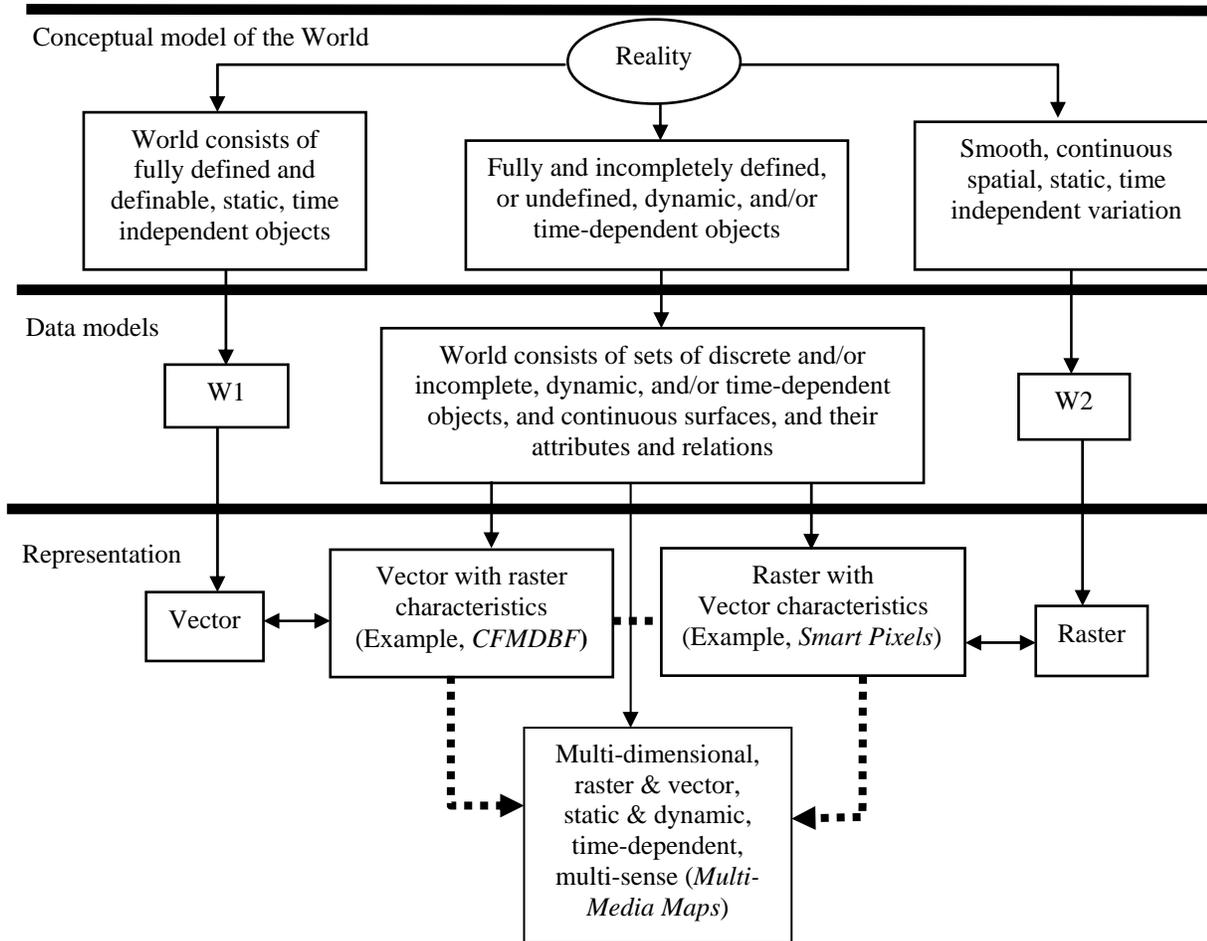


Figure 3. Vision of an Extended Model

The Smart Pixel Format is a raster format where features are expressed as attributed continuous structures using three-dimensional Freeman Codes developed at the Center for Mapping. We are currently researching the geometry of the grid space to implement rotation and scaling of raster features (all the pixels describing a feature). McCullan (1997) presents preliminary results.

We recognize that the model shown by Figure 2 is just the first step toward the development of a more complete and useful model to represent geographic data. Geographic data today are multi-dimensional, multi-spectral, static and/or dynamic, and a representational model must be able to accommodate all these data types as well as to incorporate sound, and natural and special effects. A vision of this model is shown in Figure 3.

This model incorporates static, dynamic, multi-dimensional, multi-spectra, and time-based phenomena, as well as sound and natural and special effects. To develop the extended model, we integrated our previous research (Ramirez, 1997) on the nature of raster and vector data with the model presented in Figure 2 to come up with the extended model shown in Figure 3. In the 1997 research, we investigated some fundamental questions about data

representations, such as: are raster and vector different types of datasets? If not, is there a common framework, which expresses both datasets as a unique data type? We found that raster and vector datasets belong to the same type of data structure, some kind of super-structure, and that it is possible to express them by a single set of common expressions (their position and attributes). In the next sections, the most important elements of the extended model are discussed.

3. The “Super-Structure” for Raster and Vector Data

Ramirez (1997) found that raster and vector data could be expressed in an orthogonal three-dimensional coordinate system (X, Y, Z) by a single set of equations of the type:

$$X = BU x, \quad Y = BU y, \quad Z = BU z, \quad (1)$$

Where,

x = number of basic units in the first primary direction (for example the X-axis)
y = number of basic units in the second primary direction (for example the Y-axis)
z = number of basic units in the third primary direction (for example, the Z-axis)
BU = basic unit

$$BU = \text{Distance. } 1, \quad (2)$$

And,

Distance = 1 (for vector model)
Distance > 1 (for raster model).

The Cartesian distance between two space points **A** and **B** is given by,

$$d_{AB} = [(X_A - X_B)^2 + (Y_A - Y_B)^2 + (Z_A - Z_B)^2]^{1/2}$$

or,

$$d_{AB} = [(BU x_A - BU x_B)^2 + (BU y_A - BU y_B)^2 + (BU z_A - BU z_B)^2]^{1/2}$$

$$d_{AB} = BU [(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2]^{1/2} \quad (3)$$

And the fundamental transformations are expressed as follows.

Scaling of the line **AB** by a factor **S** is given by,

$$d_S = S d_{AB} \quad (4)$$

Replacing, d_{AB} by (3), we obtain:

$$d_S = S BU [(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2]^{1/2} \quad (5)$$

And, the coordinates of the scaled line are given by:

$$\begin{aligned} X_A &= S BU x_A \\ Y_A &= S BU y_A \\ Z_A &= S BU z_A \\ X_B &= S BU x_B \\ Y_B &= S BU y_B \\ Z_B &= S BU z_B \end{aligned} \quad (6)$$

Translating a line **AB** is given by the expressions:

$$\begin{aligned} X_A &= BU x_A + BU dx \\ Y_A &= BU y_A + BU dy \\ Z_A &= BU z_A + BU dz \\ X_B &= BU x_B + BU dx \\ Y_B &= BU y_B + BU dy \end{aligned} \quad (7)$$

$$Z_B = BU z_B + BU dz$$

Finally, rotation of a line AB by the angles φ , ω , and γ is given by,

$$\begin{aligned} X_{Ar} &= \\ Y_{Ar} &= \\ Z_{Ar} &= \\ X_{Br} &= \\ Y_{Br} &= \\ Z_{Br} &= \end{aligned} \quad (8)$$

4. Describing Features in the Global Model

The global model encompasses the geometric aspect of the traditional raster and vector model. The description provided by this model is equivalent to the skeletal representation developed by Ramirez (1992) for vector data. In order to provide a complete global model, three additional aspects need to be discussed next: description of features in the raster model, graphic variables (or graphic attributes), and non-graphic attributes. We present next a summary of raster feature description. The reader interested in the other two aspects is referred to Ramirez, 1997, for a complete discussion of these topics.

4.1 Description of Raster Features

Traditionally, raster images are composed of “dumb” pixels. “Dumb” pixels have no connectivity, geometric, or feature related information. They carry only positional (I, J) information and an attribute. A typical example of “dumb” pixel images is orthophotos. Orthophotos show the surface of a particular area of the earth in an orthogonal projection (distances and angles are equivalent to the ones on the ground). They carry a large amount of implicit information, but little explicit information. Ideally, we would like to have images with “smart” pixels. “Smart” pixels of images carry a large amount of explicit information (similar in some fashion to the information in the vector model).

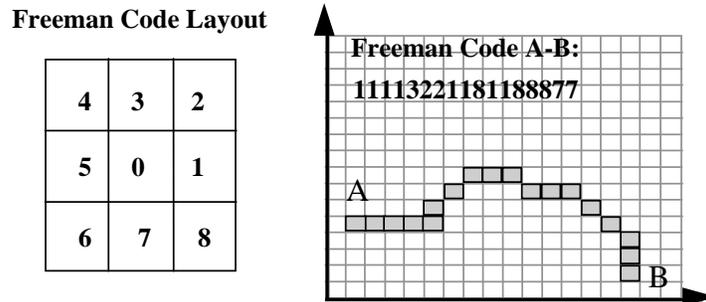


Figure 4. Freeman Structure

A simple way to define features in the raster model is by using the Freeman code. The Freeman code carries connectivity and geometric information. Skeletal representation of features in the raster model can be expressed by the Freeman code. Figure 4 illustrates the basic structure of Freeman Codes.

4.2 Expanding the Freeman Code

The traditional representations of features in raster images by Freeman code are still short, by several factors, of the equivalent vector representations. Some of them (graphic variables or graphic attributes, and non-graphic attributes) were mentioned earlier. A factor not mentioned yet, is the geometric dimension of conventional raster data (2-D) as compared to spatial vector data (3-D). A simple solution to this problem is to extend Freeman from two-dimensions to three. This can be accomplished as follows (see Figure 5).

The planar representation of Freeman code can be extended to a volumetric representation by considering the cube (instead of the square) as the fundamental representational unit. In this particular case, a three-dimensional pixel (or voxel) will have twenty-six (and only twenty-six) adjacent pixels. These pixels are located at a level (Level 1) below

the pixel of interest, at the same level (Level 2) as the pixel of interest or at a level (Level 3) above the pixel of interest. Figure 5 shows the basic unit, the pixel of interest and the twenty-six adjacent pixels, and the three levels with and identification number for each pixel. Pixel 10 is the pixel of interest.

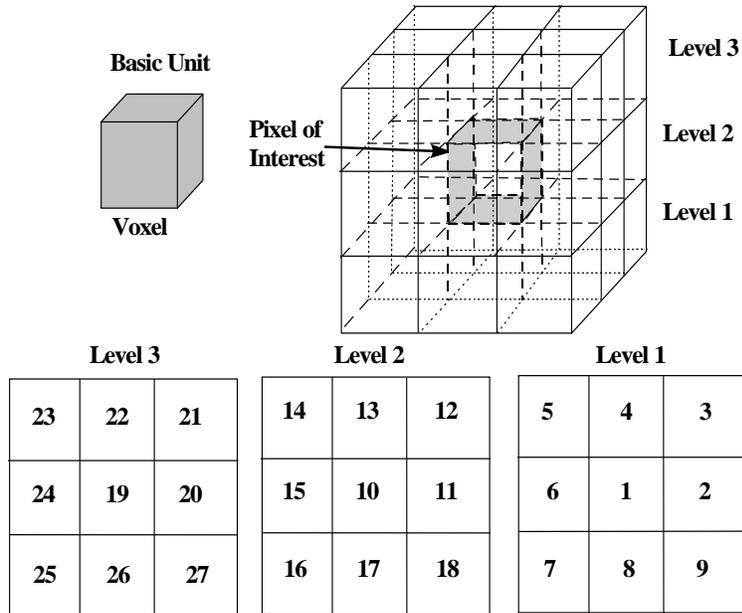


Figure 5. Three-D Freeman Code

4.3 Use of the Three-D Freeman Code

Spatial features can be expressed by Three-D Freeman code. Figure 6 illustrates one of such features. The darker blocks indicate the skeletal representation of the feature A-B. The 3-D Freeman code describing this feature is:

25 19 15 15 15 15 15 24 19 22 15 15 24 24

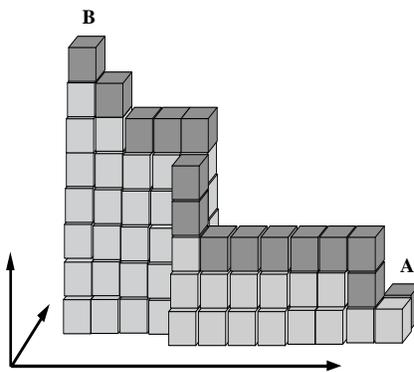


Figure 6. A Three-D Feature

5. Some Components of the Model

The following are some of the new components of the model:

5.1 Static and/or Dynamic Components

To the naked eye, the environment we live in contains static features (such as, roads, buildings, etc.) and dynamic features (such as streams, clouds, etc.). However, current representational models deal only with static features by representing the environment at a unique instant of time. For that particular instant of time, the environment looks static to the naked eye.

Most of the geographic problems we deal with cover an interval of time, not a single instant, and, therefore, include static and dynamic features and events. Examples of these problems are traffic flow at rush hour, flooding, erosion, tracking a missile, etc. Therefore, a realistic representational model must incorporate dynamic objects and time.

Incorporation of dynamic objects as part of the representational model can be viewed from two different perspectives. The first perspective is to consider multiple instants of time. For example, a sequence on time instances such as 10:00:01, 10:00:02, 10:00:03, 10:00:04, and so forth. Each instant of time is static in nature but the multiple instants of time taken together provide the idea of a dynamic environment. In the second perspective, we assume that the environment is dynamic in nature, that every thing is changing, and that static features are dynamic features with changes equal to zero.

We propose to use the first perspective here. The basic reason for this decision is the fact that we have a lot more experience with this perspective. A movie, for example, is a set of static frames, that when projected sequentially, provides the illusion of movement and a dynamic environment. This approach will allow us the use of the above mathematical expressions (formulas 1 to 8) as the basic mathematics for incorporating dynamic components.

5.2 Incorporate Fuzzy Boundaries

Objects on the ground sometimes are well defined (with crisp boundaries), sometimes they have smooth continuous variations, and sometimes they are incompletely defined or indefinable (with fuzzy boundaries). As indicated earlier, vector representation assumes crisp boundaries for all the objects represented, and raster representation assumes smooth variations. However, current representational models do not support the definition of fuzzy boundaries.

Fuzzy boundaries are common in our environment, such as the edges of dirt roads, forests, snowfields, soil types, coastlines, etc. Fuzzy definitions and representations of ground objects require revising the traditional ideas of how ground objects are represented on maps and/or images. In this context, objects can be represented not only by crisp boundaries and/or smooth continuous variation, but also by fuzzy boundaries. Lines and points (in 2-D), or areas, lines, and points (in 3-D) represent features with crisp boundaries; regular grids (in 2-D) or regular polyhedrons (in 3-D) represent smooth continuous variations; and 2-D or 3-D buffers or uncertainty corridors represent incompletely defined or indefinable objects.

5.3 Incorporate Natural Events

The geographic environment we live in is affected by all kinds of natural events. Weather, for example, continuously affects our perception of the world. Other natural events of catastrophic natures, such as flooding, earthquakes, hurricanes, tornadoes, landslides, etc., could greatly change the geographic environment.

Today, there is a great deal of data about natural events. Weather information can be accessed over the Internet twenty-four hours a day; hurricane trajectories and flooding are forecast days in advance. It makes sense to incorporate these natural events into an extended representational model.

Incorporation of weather information results in incorporating illumination changes, and shadows, snow, fog, rain, etc. The result is a more realistic representation of the environment. Computer graphic experts have studied the problem of illumination for about forty years. Realistic images of the environment are generated for scientific and leisure purposes. But, the question in consideration today is: are these images correct? Would they accurately represent the existing environment? Greenberg and et al. consider that the answer to this question is no! "Yet the results are appealing because the resulting images are believable." Realistic and accurate images of the environment are our goal.

We propose to incorporate weather conditions in a range going from schematic to realistic. In all cases, accurate representation of the environment will be a must.

Catastrophic natural events are part of the extended model. Their practical implementation in the extended model is constrained by our understanding of their behavior. Experts in the appropriate fields have developed prediction models and they are working to improve them. We foresee the extended model as the interface to connect the description of the environment with the effect (or possible effect) of catastrophic natural events.

5.4 Incorporate “Augmented Reality”

One of the most important applications of computers and computer graphics is the capability to present and analyze “what-if” situations. Computer simulation is a well-established part of computer sciences. It includes concepts that range from the very profound (such as Artificial Intelligence, the simulation of human intelligence processes by computer systems) to very practical concepts (such as traffic flow in a city). As indicated earlier, current representational models are static structures representing an area of interest for a moment of time, and therefore, they are not designed for simulations. We would like to have “augmented” objects as part of the extended model.

“Augmented” objects are virtual (non-existing) objects merged with real ground objects in the model. The actual environmental representation would be automatically modified to reflect what would happen if such objects did exist. These objects would be removed and placed under user control.

“Modified” objects are a special case of “augmented” objects. They are a combination of objects existing in the environment and virtual objects. The combination can be additive or subtractive. Therefore, the resulting object may be equal to the original ground object augmented by some virtual component, or may be equal to part of the original object. These objects would be removed and placed under user control.

6. Conclusions

In the previous paragraphs we have introduced an extended representation model, its mathematics, and described some of the components of the extended model. This description is not exhaustive. It is only a first step toward the realization of a comprehensive model describing the environment by geographic data. Our goal is to visualize the environment in a realistic and accurate fashion.

Reference

- Bidoshi, K., A Database System as a Tool for Comparing, Updating, and Conflating Spatial Data from DLG and GPSVan™, *Unpublished Thesis, The Ohio State University*, Columbus, 1995.
- Burrough, P.A., “Natural Objects with Indeterminate Boundaries,” in *Geographic Objects with Indeterminate Boundaries*, Editors Burrough, P.A. and Frank, A.U., London: Taylor&Francis, 1996.
- McCullan, J.D., Exploration of the Geometry of Rotations in the Pixel Domain, *Unpublished Thesis, The Ohio State University*, Columbus, 1997.
- Ramirez, J.R., “Development of a Common Framework to Express Raster and Vector Datasets”, *Proceedings of Auto-Carto 13*, Seattle, 1997.
- “Development of a Language for Expert Cartographic Systems”, *Proceedings of the XVII ISPRS Congress*, 1992.