Multiple representation of spatial data, which means obtaining different detailed representation of geographic phenomena based on the same spatial database or different data versions, is one of the key technologies for digital earth. It plays an important role in such applications as seamless data navigation, progressive web transfer and self-adaptable visualization. Multiple representation technology meets the contradiction between data volume and high granularity representation. This paper aims at the objective by small size of data volume to realize high granularity representation, presenting a concept, namely representation lifespan over scale space to describe that the spatial data has different representation scene in scale space. The scale transformation is represented as $t_{ij}$: $f_i \rightarrow \{[s_i, s_j], \{g_{ij}\}\}$, where $f_i$ is the transformation function, $[s_i, s_j]$ the scale range controlled by two key scale points $s_i$ and $s_j$ and $\{g_{ij}\}$ the base representation status associated with the key scale point of $s_i$ or $s_j$. The study summaries that there are four transformations, namely generalization, LOD accumulation, morphing and extraction.

1. INTRODUCTION

Digital Earth concept requires to represent the geographic phenomena on the earth covering a large scale change. Goodchild points out that the range of scales implied is over at least four orders of magnitude, from a resolution of 10km that would be appropriate for rendering of the entire globe, to the 1m resolution needed to render a local neighbourhood (Goodchild, 1999). It presents a strong requirement for multi-scale representation of GIS and generalization technology. Current GIS can not support the dynamic zoom-in function, which means when zoom-in the geographic data we are able to get more detailed information rather than just the graphic symbol exaggeration. This dynamic zoom-in process is also called on-the fly generalization (Oosterom, 1995). Multiple representation of spatial data is to obtain different detailed representation of geographic phenomena based on the same spatial database or different data versions. It plays an important role in such applications as seamless data navigation, progressive web transfer and self-adaptable visualization.

Two kinds of generalization methods can be used to conduct multi-scale representation. One is to store different scale version map in the database and then to extract different layer data to display according to the scale change. The data layers describe the same geographic area but with varied resolution. The important feature which will be displayed at a determinate scale can be pre-decided in the database through a visualization related field definition in the feature table. The value in what scale it will be displayed can be assigned into the field. This method requires huge memory space to store redundant map data and additional preparation works to organize the various scale map layers. Furthermore, the representation can only carry out limited discrete layers in scale change, so it is not a really continuous scale change. The advantage of this method exists in quick running because it just read data to visualize without additional algorithm model computation. Another method is based on special generalization model which automatically compute and determine whether objects are displayed or not and how detailed graphic to be displayed in the process of zoom-in. The latter automatic method needs hierarchical data structure and abstraction algorithm to support. The research associated with this method is active in the field of map generalization.

One of applications of multi-scale representation is the progressive transfer of spatial data over web from coarse to fine. In the sequence of significance, the map data is transferred and visualized on the client step by step with increasing details. Once the user finds the accumulated data meets his requirements, he can interrupt the transmission at any time. It is a self-adaptable transmission procedure in which the user and system can communicate interactively. As the complete data on server usually covers many details over the requirements of users, the interruption can save much time for some users. The progressive transmission not only speeds up the web transfer but also respects the principle from coarse to fine details in the cognition of spatial information. From the point of view of information acquisition, the progressive process behaves as an efficient navigation guide. Recently the progressive transmission of vector data becomes an active issue. Bertolotto & Egenhofer (2001, 1999) first present the concept of progressive transmission of vector map data and provided a formalism model based on distributed architecture. Buttenfield (2002) investigates the requirements of progressive transmission and based on the modified strip tree (Ballad, 1981) developed a model for line transmission. From the point of technology view, Han and Tao (2003) design a server-client scheme for progressive transmission.

This technology can be regarded as the inverse process of continuous map generalization. The key solution is to pre-organize generalized data on server site in a lineal order with details increment. This study tries to offer such a model to...
describe the data representation over scale space and apply in the progressive transfer. The model is representation lifespan over scale space, which is to depict spatial object in multi representation driven by different transformation.

The rest of paper is organized as follows. Section 2 investigates the general questions of multiple representation focusing on the granularity partitioning, data volume reducing and consistency maintenance. Section 3 presents the model of representation lifespan over scale space. Four transformations are discussed in section 4 to realize the representation from one scale range to other scale range. Section 5 provides the conclusion with the future works.

2. CHALLENGES FROM MULTI-SCALE REPRESENTATION

2.1 Granularity Partitioning

The GIS data has different resolutions represented by term granularity which refers to the minimum data unit in data representation. In multi-scale representation, the series of representations over scale space can acts as change accumulation at different levels. The component of a vector map can be divided into a hierarchical structure with three levels: feature class, object and geometric detail. The feature class refers to the object set with the similar theme, such as the hydrological feature. The object is the independent entity with complete geographic meaning under one layer feature, such as the line river, the polygon lake. The geometric detail is the component parts to compose one object, such as the bend contained in the river curve. From one scale to another scale, the representation will jump from coarse to fine with some information addition or reduction. In progressive web transmission, the data element that each step transmits can correspond to different levels in vector map component structure. We define the transmitted data in one step the transmission granularity. Then there are three kinds of granularities. From layer feature, object to geometric detail, the representation granularity decreases and the changes between two adjacency transmissions also reduce correspondently.

The granularity at feature class level means the change between consecutive representations over scale is one feature class. Obviously this kind of granularity is too coarse. In Steven’s theory of scale measurement (Steven, 1946), the concept theme belongs to the nominal variable not comparable in significance order. On server side the pre-organization of data can not predict the later demands from client users according to feature importance. The organization sequence of feature layers usually is not what users want. To download map data, generally the users on client not only have requirements in theme selection but also in representation scale.

Under a feature layer, the object elements can be sorted on significance grade in spatial representation. For example, in catchment representation the river branches can be organized to a lineal sequence based on Horton code and river length. In this multi-scale representation, the change between consecutive representation states is one object. The transmission of river branches in this order will result in the catchment representation with increasing details. The GAP-tree structure with the lineal sequence of polygon organization is able to support the progressive transmission of categorical area features (Oosterom, 1995; Ai and Oosterom, 2002). Unfortunately not all objects can be structured in such a lineal order, especially for those objects across different themes. However, the object sequence in progressive transmission is not strict, usually behaving as the order among object groups. Objects within group A have to be transmitted before those within group B. But the transmission sequence of objects within same group, i.e. A or B is not of importance. The transmission granularity in object level is enough for the applications whose users only care the representation resolution in object level. The progressive transmission model proposed by Bertolotto & Egenhofer (2001) belongs to this level that each step transmits one object not involving the geometric details. On client screen, the transmission with the granularity of object level reflects as either appearance or disappearance of one complete object. Once an object appears it remains the same scene without details add. It is still a coarse transmission as far as the granularity is concerned.

The granularity of geometric detail usually reflects as the segment of line, bend of curve, concave/convex parts of polygon and so on (Ai, Guo etc. 2000; Muller, 1992). The gradual add of geometric details refines the object representation and let user get the image of dynamic evolution. Compared with the pixel to compose image, the vector data is more complicated in both element structures and component methods. The decomposition of object into series of details is a difficult question when considering scale impacts. Thus the transmission of vector data under the granularity of geometric detail becomes a bottleneck (Under the other two levels is relatively easy.) The LOD technology in the field of computer graphics can be introduced to resolve vector data decomposition. But most algorithms on LOD are based on grid or mesh structure and aim at three dimensional objects.

How to determinate the transmission granularity among three levels? The determination of transmission granularity depends on what spatial concept that users interest and how broad the cognition ranges in scale space. If the user is interested in the catchment representation, the representation granularity at river branch is suitable. The further separation curve into bends is not necessary. But if users are interested in the representation of Line River, the representation granularity has to reach to the level of curve bends. So there is the principle that the representation granularity is down one level compared with the level of spatial concept that users interest in vector map data component structure.

Once the representation granularity is determined, we need to apply certain operations to decompose the representation into details adaptive to this granularity. From the point of view of map generalization, it is the scale transformation that separates the transmission granularity. The progressive transmission can be regarded as the mapping from the representation over spatial scale to the representation over temporal scale. Each time transmits (or displays) one representation suitable for a certain spatial scale. Every generalization operation must yield representation changes but the change degree is different. In generalization operation, there are three hierarchical strategies, namely operators, algorithms and parameters (Shea and McMaster, 1991). We can design a generalization method with different hierarchical strategies to get the transmission granularity we want according to the change degree.
The operators selection/elimination and typification result in the appearance or disappearance of one object completely (change at object level). The operators simplification, collapse, aggregation, amalgamation generate changes in geometric details of one or many objects. To tune the tolerant parameters in generalization algorithm can further adjust the changes in different granularity levels. Figure 1 shows the morphing transformation from a line with two end-points to a complex curve. The transformation steps and speed will determine the detail changes and the transmission granularity when the series data is sent over web. Not all existing generalization algorithms are adaptive to decompose transmission granularity. If one generalization algorithm is sensitive in scale, which means the algorithm can output a new representation state once the scale occurs little change; this generalization algorithm is suitable for the progressive transmission.

2.2 Data Volume

Reducing the data volume as much as possible is another requirement for multi-scale representation of spatial data. Generally the data volume that aims at the multi scale representation is much larger than that of complete representation with full details due to the adding of middle gradual representations. If the user wants to download the whole data, he will suffer from the progressive transmission taking more time than that of direct transmission. However its advantages in other aspects show it a valuable process. The solution of the contradiction between multi-scale representation and large data volume exists in the compression of vector data. We may settle this question by three strategies: (1) only recording change parts rather than complete representation states, (2) distinguishing key representation and removing unimportant ones, and (3) deriving new representation state through the transformation function.

In the compression of multimedia data, such as audio data and video data, we try to detect change parts and record it in the file. In vector representations over spatial scale two consecutive states must have much overlap parts. We can also extract the change parts to express the vector representation. Based on this idea and for the purpose of data volume decrement, we will present our method, the changes accumulation model in section 3 to record details in different levels. Unfortunately, many generalization algorithms can just output independent representations corresponding to one scale without providing connections among the series of representations over scale range. A post-process is required to extract changes through the comparison between two consecutive output results.

From coarse state to fine state, among the series of representations the contribution of each state in gradual evolution is not equal to each other. When we put the representations in a line layout, we may find some of them are key stages but others not. The removal of unimportant stages will not affect the progressive transmission, but the miss of important stages will destroy the refining process. So distinguishing and removing some of unimportant representations is a useful step to reduce data volume. If we can find a suitable transformation function, say morphing (Cecconi, 2002), to automatically change the representation from one key stage to the next key stage and also can output middle representation stags, we can just store the key stage representation and let the function later to derive the non-key stages. Apparently the data volume is greatly reduced. Under the control of two terminal key stages, the morphing transformation reflects as the interpolation of representation. This on-line transformation is able to output generalized results in real time. Maintaining the spatial relationships, such as the handling of conflicts when too many objects simultaneously appear, can be accomplished by the mathematic transformations of translate, rotation and scale of some object. For instance, we can record an offset variable with object A and in later transmission if neighbour B appears let A move the offset distance through translate operation. The data volume is smaller compared with the storage of two states. Another kind of transformation is the combination of details based on “change parts” which have been decomposed in pre-organization process. The combination operation acts as the simple addition and subtraction of changed details.

To reduce the data volume in multi-scale representation, the above three strategies have something to do with map generalization but it concerns more transformation procedures. The traditional generalization technology which focuses on state representation at target scale needs to be improved in some degree.

2.3 Consistency Maintenance

If the transmitted data is just for visualization, the decomposed details arriving at the client need no post-process and the procedure of gradual adding details has reflected the progressive effects. It is just like the raster data transmission and visualization from coarse to fine by LOD technology. However, if the transmitted vector data is downloaded for the purpose of spatial analysis or imported into other application systems, the decomposed details need to be composed and restored as the original form, just like the decoding process in signal processing (Bertolotto and Egenhofer, 2001). In GIS applications, the spatial object has the same representation in logical level among different systems, although the concrete data structures and realization methods may differ. For example, the polygon object is represented as a closed coordinate string with possible one or more inside loops. But for the progressive transmission the resulted data may be quite different from this...
For the multi-scale representation at object level, the element remains the independent geometry and the component operation is able to simply through the union operation. For that in the level of geometric details, the set operations may be complex depending on the decomposing methods. The details are usually organized in a hierarchical structure. To restore to the original representation, it needs to determine which detail “nodes” in the hierarchical tree should be selected and what relations they have to each other. Some details make the positive contributions in object component, but others make the negative contributions.

In data restore, another question is the maintenance of topological relationship for the final sub-set data when the user interrupts the transmission after the data has enough details to meet his requirements. The retrieved sub-set data is dynamic under the control of user, leading to the difficulties in the maintenance of topological relationship. First detect where the topological relations have been destroyed among the sub-set data through the comparison with original relation based on the neighbourhood analysis. Usually the relation destruction results from either the reduction of the number of objects or the change of geometric representation compared with the original representation. Part of relation destructions reflect as the spatial conflicts in map generalization. Then use the consistency operation to restore the original relation. In Figure 2, when the user interrupts the transmission, the middle road is represented as line rather than narrow polygon and then two neighbour land-use parcels will no longer have “touch relation” to each other. Due to the collapse of road object, the gap between two land-use parcels needs to fill by polygon extension toward to keep the original “touch relation”. In Figure 3, two buildings with full details representation have “touch relation” to each other. During the progressive transmission, retrieving the approximate representation at a coarse detail, such as the bounding rectangle, leads to the overlap conflict, against the original relation. It needs displacement operation to correct the destroyed relation.

Bertolotto & Egenhofer (2001) pointed out the consistency is an essential property for the usability of data. The consistency question is associated with the horizontal contexts and by now few generalization methods have settled the context consistency well. Decomposing the representation at different detail levels and selecting part of objects or details must result in inconsistency. For progressive transmission, on server site we can not forecast what objects will appear together and in what detailed representation. It is nearly impossible to find a way to decompose details to guarantee all possible consistencies among different components of later transmitted data. So the post-process on client side becomes an inevitable step.

3. REPRESENTATION LIFESPAN OVER SCALE SPACE

In map representation space, every spatial entity has a limited range of representation scale (Cecconi, 2002). For instance, the building can be represented as polygon under large scale (1:5,000), a rectangle under middle scale (down to 1:20,000), and a point under small scale (down to 1:50,000). When further down scale such as 1:50,000, the building will disappear if just considered the impacts from spatial scale without special semantic purpose. We define the scale range for one object representation from birth to disappearance the “representation lifespan over scale”.

Over the lifespan, one object faces different operations to abstract the representation and we can distinguish two change stages: the key stages and non-key stages. The key stages are those associated with steep change in geometric or semantic aspects, such as the disappearance of one object (elimination), the decrement of spatial dimension from three to two or from three to one (collapse), the amalgamation of various objects within a region to get a new concept object, and so on. The non-key stages are those related to smooth change in quantity with the basic properties preserved in quality, such as the simplification of curve or polygon, local displacement, exaggeration, rectification of building. The key stage happens at one point while non-key stage occurs within duration over scale range. The key stage and non-key stage happens in turn, which means a key-stage is followed by a non-key stage and vice versa.

Figure 4 shows the representation lifespan of the river representation from detailed to simplified states, the inverse of refining transmission, in the order: polygon simplification (non-key stage), collapse (key stage), line simplification (non-key stage), elimination (key stage). The generalization related to key stage is usually more difficult than that to non-key stages due to the consideration of more constraints and more complexities to maintain the relationships after steep change. The key stage transformation is usually finished in off-line generalization requiring complicate algorithms and much running time while the non-key stage is finished in on-line generalization. To reduce the data volume for progressive transmission, we can examine the generalization lifecycle to distinguish the key stage and non-key stages and remove part of non-key stages.

Figure 4. The representation lifespan of river feature, including polygon simplification, collapse, line simplification, and elimination.
The representation lifespan thinks the process of spatial representation from fine to coarse as the process of evolution from birth to death. Just like the lifespan of a natural life, from birth to baby to youth to midlife and finally to death, the geographic object also has different cartographic representation patterns while the scale continuously changes from the visualization perspective. The representation period of one geographic object in scale space can be denoted as

$$\sum_{s_{1}}^{s_{n}} R_{s} = \int_{s_{1}}^{s_{n}} dR/ds$$

The concept of the representation lifespan has the following basic properties:

1. The length of representation lifespan is limited. Each geographic object can only be represented in a certain scale range \([s_{0}, s_{n}]\). Once the representation scale out of the range the object will disappear from visualization scene. For example, a long and narrow river in 10km may have a representation period from 1:1 to 1:1 000 000, and in the spatial scale 1:1 000 000 the river may be represented as a line of 10mm length. If the scale continues to decrease less than 1:1 000 000 the river will disappear.

2. The representation state is multiple. The same reality has different representations in different spatial scale ranges. This property implies: \(\exists j \in [s_{1}, s_{2}], \forall i \in [s_{1}, s_{2}], \text{ makes } R_{i} \neq R_{j} \).

3. The evolution rate is different with slight changes and steep changes occurring alternatively. The change rate can be represented as \(dR/ds\).

4. The granularity of \(dR/ds\) is hierarchical. The steep change over scale space usually includes several slight changes.

To construct the object-oriented model of representation lifespan, we consider that in the whole lifespan \([s_{0}, s_{n}]\), an object has some base representations \(g_{0}, g_{1}, \ldots, g_{m}\) which relates to some key point \(s_{0}, s_{1}, \ldots, s_{n}\) over scale space. An essential evolution of the representation arises at each key point, such as collapse from polygon to centreline. But between two adjacent key scale points \(s_{i}\) and \(s_{j}\), the evolutions of the representation are smooth. So the middle representations between \([s_{i}, s_{j}]\) can be derived by a transformation function \(f\). Based on this idea, the scale based transformation is defined as a tri-tuple:

\(T_{ij}: <f_{i}, [s_{i1}, s_{i2}], \{g_{ij}\}>\),

where \(f_{i}\) is the transformation function, \([s_{i1}, s_{i2}]\) the scale range controlled by two key scale points \(s_{i1}\) and \(s_{i2}\) and \(\{g_{ij}\}\) the base representation status associated with the key scale point of \(s_{i1}\) or \(s_{i2}\). Then the representation of the object at any scale \(s_{x}\) can be formulated as:

\[R_{x} = f(\{g_{0}\}, x), x \in [s_{i1}, s_{i2}]\]

4. FOUR TRANSFORMATIONS

From one representation state to another state, it is the scale transformation that drives the representation change. There are different scale transformations to conduct the representation change. Considering the transformation properties of slight change or steep change, we summarize four types of transformation functions.

1. Traditional map generalization \(R_{s}=G(g_{0}, s)\), from the point of view of map generalization, it is the scale transformation that results in the different representations. Using \(G\) transformation, we can get a relative coarse representation from it’s refine version. In Figure 5, the top one shows this kind of transformation, the left is a detailed building distribution, after map generalization the buildings are simplified into several blocks (the right).

2. Interpolation (morphing) transformation \(R_{s}=M(g_{0}, t, s)\), unlike the traditional transformation, the interpolation state is controlled by two input states: the initial state \(g_{0}\) and the end state \(g_{t}\). In figure 5, the second one show the Morphing transformation, in which the \(s_{0}\) and \(s_{1}\) are two key frames, the series representation between \(S0\) and \(S1\) are the inter-medias.

3. LOD accumulation \(R_{s}=L(g_{s}, s) = g_{s} + \sum \Delta R_{s}\). This transformation considers the spatial representation from one scale to another as an accumulation of the set of changes. The difference between two consecutive representations is recorded in a linear order and through gradually addition or subtraction of “change patches” the goal representation is achieved. By such a technique, the data volume can be reduced without losing the necessary details. In figure 5, the third one shows this kind of transformation, by the addition of LODs , a refined representation can be drive out.

4. Equivalence transformation \(R_{s}=E(g_{0}, s)=g_{e}\). It is just to extract the pre-organized representation version without real geometric transformation. It is applied in multi-scale representation by different data versions. In figure 5, the last one is the \(E\) transformation, in which the two representations are equal.

To build the model of representation lifespan, the key question is how to rationally divide the representation period \([s_{0}, s_{n}]\).
into several sub spatial scales \[s_{i1}, s_{i2}\]. In practice, the elements relates to some special operator for generalization. Experientially, collapse by which the geometry dimension descending, and aggregation by which many object becoming into one, and delete by which an object disappearing etc all cause the essentially representation changes. In conclusion, the scale point, at which the representation changes essentially, is the key scale point.

5. CONCLUSION

The multi-scale representation provides users with a self-adaptive method to access data over web at different resolution. Also it plays an important role in the data navigation for users to acquire spatial information from coarse to fine, consistent with the process of information cognition. The image data, raster data and DEM has been realized this multi-representation. The multi-scale representation of vector map data is still an open question. In this paper we discuss the characteristics of multi-scale representation of vector data, investigating the constraints in granularity separation, data volume compression, and data restore. The multi-scale representation and hierarchical organization of vector data is a key technology for progressive transmission. This paper builds a new model “representation lifespan over scale” considering the scale change from birth to disappearance with different representation states driven by four transformations.

In technology the multi-scale representation is associated with map generalization. If one generalization can output dynamic data within a wide scale range rather than at one scale point, the series of data is well suitable for multi-scale representation. Unfortunately, most of existing generalization algorithms can just derive new data at some scale point. Rather than the generalization, the other transformations such as LOD technology and morphing methods need to apply in multi-scale representation.

REFERENCES:


Ceconi, A., 2002, Improving Map Generalization for On-Demand Mapping by Multiscale Databases, Joint ISPRS/ICA workshop on Multi-Scale Representations of Spatial Data, Ottawa, Canada.


