MULTI-SCALE REPRESENTATION OF HYDROGRAPHIC NETWORK DATA FOR PROGRESSIVE TRANSMISSION OVER WEB

Tinghua Ai[1], Bo Ai[2], Yafeng Huang[1]
[1]School of Resource and Environment Sciences, Wuhan University, China
[2]Geomatics College, Shandong University of Science and Technology, China

Abstract The progressive transmission of map data over World Wide Web provides the user with a self-adaptive strategy to access remote data. The key technology in this transmission is the efficient multi-scale representation of spatial data and pre-organization on server at a scale associated order. This study aims at the progressive transmission of hydrographic data investigating some constraints of multi-scale representation and offering an hierarchical data model of river network at both object and geometric detail levels. Through watershed partitioning generates the channel order in significance increment and then by BLG tree structure divides and stores the different representation of individual river. For a given scale, the data model can derive which river segments to be transmitted and in what detail level to be represented. The experiment shows the application of this data model in progressive transmission is successful.

Keywords: progressive transmission, map generalization, hydrographic network.

1. Introduction

The appearance of Internet presents two challenges for cartography discipline. One is that we get a new space to be mapped, namely the cyberspace or the virtual world (Taylor 1997, Jiang & Ormeling 1997). Another is pushing mapping technology into web environment, including the delivery of map data through web, the remote access of map data, the on-line making map with data from different web sides, and so on. The previous challenge resulted from new visualization content leads to the cartography re-building basic map concepts which is quite different from those in a conventional map about physical space. The latter challenge resulted from mapping technology, on the one hand, provides new opportunities and methods to represent spatial phenomena, on the other hand, results in new challenges to live with web.

Once the map data can be downloaded from web side, the user will demand a high efficiency. There are two questions to be settled: (1) quickly finding the location of map data he needs with a search engine, and (2) quickly downloading data under an interactive control. The first question depends on the special map search engine to efficiently process metadata. For the latter question, the improvement of hardware and web infrastructure, such as broadband extending, is just partly a solution. The data organization on server and transmission approach across web play important roles. In this domain, the progressive transmission of map data from coarse to fine becomes a welcome transmission method. 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-adaptive transmission procedure in which the user and system can communicate interactively. As the complete data on server usually covers much 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(Ai *et al*, 2004).

Recently the progressive transmission of vector data becomes an active issue. Bertolotto & Egenhofer (2001, 1999) first presented the concept of progressive transmission of vector map data and provided a formalism model based on distributed architecture. Later Cecconi(2003) systematically studied the method of progressive transmission over web and examined the association with map generalization. Buttenfield (2002) investigated the requirements of progressive transmission based on the modified strip tree structure (Ballad, 1981) developing a model for line transmission. From the point of technology view, Han and Tao (2003) designed a server-client scheme for progressive transmission.

The progressive transmission of raster data and DEM/TIN has been successfully implemented in web transfer (Srinivas *et al*. 1999, Rauschenbach and Schumann 1999). But for vector map data, it still remains a challenge. The reason exists in that the multi-scale representation of vector data is much more difficult than that of raster or DEM data. It is hard to find a proper strategy to hierarchically compress vector data, like the quad tree to approximate raster data at different resolution. Some researches (Stell and Worboys 1989, Jones *et al* 1996, Sarjakoski 2007, Matthias 2008) examined the hierarchical structure of map data for the purpose of decomposition at multi-resolutions. But it is just at conceptual level not presenting concrete methods in algorithm realization. Some just aims at special features and the line feature attracts more interests(Zhou and Jones 2003, Buttenfield 2002).

The progressive transmission is the application of multi-scale representation of spatial data in web transfer environment, associated with map generalization. It can be regarded as the inverse process of map generalization with high granularity change. Sester and Brenner (2004).built a method of multi-scale representation of building data by point continuously change and this model can output data stream for progressive transmission. For the same feature building data, Ai *et al*(2004) developed a model, namely the changes accumulation model, to carry out the progressive transmission by pre-decomposing the building into series of rectangles and constructing a LOD composition to get different approximation representation. The ideal solution for progressive transmission is through dynamic continuous generalization to derive different representations at high resolution based on the single detailed data. Unfortunately this kind of technology currently is still a large challenge requiring prompt response generalization algorithm. Instead, we can decompose the data by off-line method and pre-organize the generalized data on server in a linear order with details increment. In latter progressive transmission let the data model output proper data stream.

In this study we try to settle the progressive transmission by the data re-organization method. We take the hydrographic data as an example building a multi-scale representation model to organize the river data in a linear order to refine the representation at both object and geometric detail levels. The rest of paper is organized as follows. Section 2 investigates the constraints of progressive transmission. Section 3 studies the establishment of multi-scale representation model including the watershed partitioning of catchment to organize the channels to a sequence, the segment decomposing by BLG tree structure and, data model integrating and the application of multi-scale representation model in progressive transmission. Section 4 concludes with some discussions on the established model.

2. Constraints of Progressive Transmission

2.1 Transmission Granularity

The granularity of progressive transmission refers to the minimum data unit in one step of transmission. From the viewpoint of visual cognition, only when the change between two transmitted data is small enough can manual eyes acquire the effects of continue or gradual animation. For the purpose of progressive transmission, on server site the degree that one decomposes the vector data into details has to meet the granularity requirements.

The component of a vector map can be represented as 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 hydrographic feature. The object is the independent entity with complete geographic meaning under one feature class, such as the line river, the polygon lake and the point fountain. The geometric detail is the component parts to compose one object, such as the bend contained in river curve. In 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 transmission granularities. From feature class to object and further to geometric detail, the transmission granularity increases and the changes between two consecutive transmissions also reduce correspondently.

The transmission granularity at feature class level is mainly decided by the semantic hierarchy rather than the spatial scale and obviously is too coarse. In the theory of scale measurement(Steven,1946), the feature class belongs to the nominal variable which is not comparable in significance grade. To download map data, generally the users on client requires not only the theme selection but also the spatial scale selection. It implies that the transmission granularity should reach to the level of object or geometric detail.

The GAP-tree structure with the linear 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 linear order, especially for those objects across different themes. The progressive transmission model proposed by Bertolotto & Egenhofer (2001) belongs to the object level that each step transmits one object. On client monitor, the transmission with the object granularity 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 or convex parts of polygon and so on (Ai *et al* 2002, Muller 1992). The gradual add of geometric details refines the object representation and let user get the image of dynamic evolution just like the pixel add to refine the image. 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. 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 object.

2.2 Data Volume

Reducing the data volume as much as possible is another requirement for data transmission over web (Buttenfield, 2002). Generally the data volume resulted from the progressive transmission is much larger than that of complete representation with full details due to the add 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. To settle the contradict between the progressive transmission and the large data volume, the data organization can be improved by compression way. We can apply three strategies: (1) only recording change parts rather than complete representation states, (2) distinguishing the key representation and removing unimportant ones, and (3) deriving the new representation state by transformation function.

In the compression of multimedia data, such as video data, we try to detect change parts between consecutive frames and record it in the compressed file. Considering the case that for vector representations over spatial scale two consecutive states 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 can use LOD technology to get different representation by details accumulation(Ai 2004). Unfortunately, many generalization algorithms can just output independent representations corresponding to one scale without providing connections among the series of representations over scale change. A post-process is required to extract such changes through the comparison between two consecutive output results.

Among the series of data transmission from coarse state to fine state, the contribution of each state in gradual evolution is not equal. Over the representation scale space, one object needs different operations to convert the representation and we can distinguish two change stages: the key stages and the 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

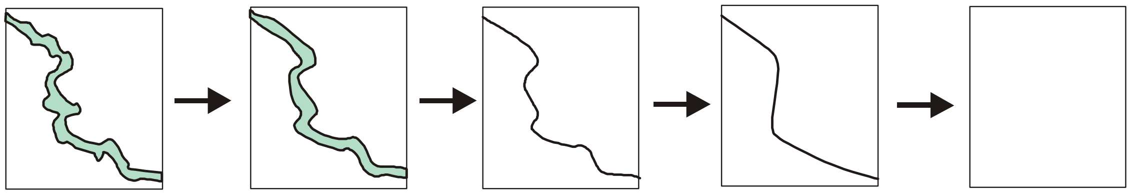


Fig. 1. The representation lifespan of river feature over scale space, including polygon simplification, collapse, line simplification, and elimination.

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 with the basic properties preserved in quality, such as the simplification of curve or polygon, local displacement, exaggeration, rectification of building and so on. The key stage happens at one point while non-key stage occurs within a 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 1 shows the representation lifespan of river object 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 scale transformation 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 conducted by off-line generalization requiring complicate algorithms and much running time while the non-key stage by on-line generalization. To reduce the data volume for progressive transmission, we can examine the representation lifespan over scale space to distinguish the key stage and the non-key stages and remove parts of non-key stages just recording versions of key-stag.

The third strategy of reducing data volume is to derive the new representation by on-line transformation function, say morphing(Cecconi, 2002). We can just store the limited key stage representation and let the transformation function later to derive the middle non-key stages. Under the control of two terminal key stages, the morphing transformation reflects as the interpolation of representation.

3. **Multi-scale Representation of Hydrographic Data**

After investigating the progressive transmission constraints on high granularity and reducing data volume as much as possible, this section takes the example of hydrographic network data to study the question of multi-scale representation for progressive transmission. We will present a model under the constraints above to support the progressive transmission.

As an important data in GIS representation, the hydrographic data has special characteristics with network distribution and hierarchical pattern(Wu 1997). Its data structure in GIS database can be illustrated as figure 2 including four levels: catchment, channel, segment and nodes. The channel is the element with complete geographic meaning under the high level catchment. It corresponds to the element in Horton order, which is composed of series of segments from outlet to joint node. The component of channel is the segment element which corresponds to the element in Strahler order with two terminal node, namely start node and end node along water flow direction.

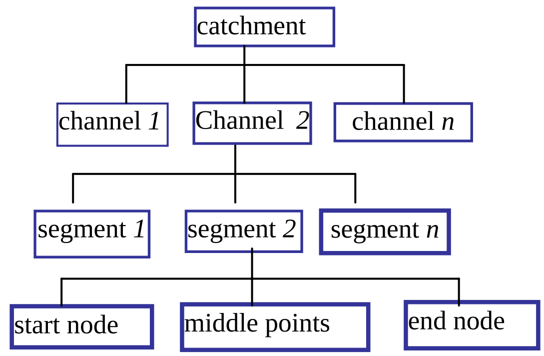


Fig. 2. The hierarchical tree of river network

For the purpose of progressive transmission of hydrographic network with high granularity and small data volume, on the one hand the channel elements have to be organized in a significance increment order to dynamically derive data step by step, on the other hand the output channel segment representation should be refined with detail add. The data model of multi-scale representation has to be established to answer the following three questions for a given scale: (1) How many branches to be selected? (2) Which channels are important to select? (3) How to simplify the selected channel?

The Töpfer law (Töpfer and Pillewizer, 1966) has answered the question one during the catchment data transformation from large scale to small scale by the computation of scale rate. For question three, there are lots of algorithms to conduct the line simplification under the consideration of special properties of river feature. But for question two, it is a decision question based on the analysis of river geographic properties and the context. At least three aspects has to be considered at different levels: the spatial distribution pattern at macro level, the distribution density and proximity relationship at meso level and the individual geometric properties at micro level. In traditional map generalization, some studies investigate the question 2. Richardson (1993) presented a method to select river based on Horton order and river length. Thomson and Brooks (2000) applied the Gestalt recognition principles in river network generalization judging the main channel and removing unimportant channels. Since the distribution of river network associates with the terrain surface, Wolf(1988) built a weighted network data structure integrating the drainage, ridge, peak and pit point. This data structure supports to determine the significance of river channel.

Aiming at progressive transmission, the multi-scale representation of hydrographic network needs to combine the channel selection at object level and the segment simplification at geometric detail level. Both are controlled by scale parameter. This study tries to build such a data model

3.1 Partitioning the Watershed by Delaunay Triangulation

3.1.1 Motivation

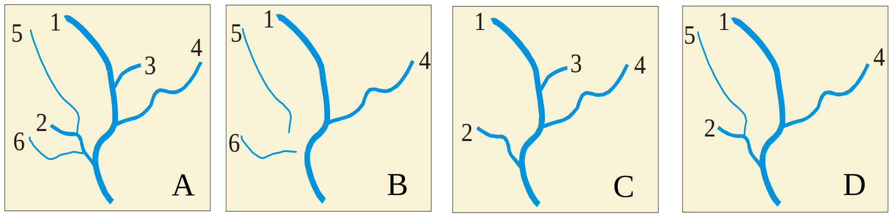


Fig.3. The river selection from catchment based on different parameter tolerances respectively. (A Original river network. B. Selection by length . C. Selection by Horton order. D. Correct selection)

The channel selection in river network generalization has to consider the order, the length, the distribution pattern and other parameters. The selection only based on one parameter condition can not get ideal result. In figure 3, the original river network is represented as *A* with river 1 the main channel, river *2,3,4* the second order and river *5,6* the third order. If we just consider the length tolerance, the selection result is illustrated as *B* in which short channel *2* is removed but the child channel *5, 6* remained as the dangled branches. If we just consider the Horton order, the selection result is as *C* in which the channels of third order namely *5,6* are removed. But channel *5* is very long although the order is low. The correct selection should be *C* in which the integration of length and order is taken into consideration.

How to find a simple parameter which integrates the length, the order and the distance between proximity rivers as the importance decision condition in the selection of river network? In this study we try to let the watershed area playing this role. The reason exists in that the watershed area describes the integration of three aspects.

First the watershed area strongly depends on the river distribution density(Tribe 1992). If the river distributes in a dense way, the channel can just compete to obtain a small watershed area. The watershed area considers the context impact that the same river in a high density area is less important than that distributes in a sparse region.

Secondly the watershed area has considered the order impact. The hierarchical structure shows the watershed area of high order channel is not smaller than the integration of that of all its child channels. This principle guarantees the parent channel has preference to its child channels in selection, not generating the case such as in figure 3 *B* that the short parent channel removed but long child channel remained.

Thirdly the watershed area has considered the channel length. Obviously the long channel extends in a large range obtaining a large watershed area. For the watershed-area-based selection conducting on the river network in Figure 3 A, we may see channel 3 has smaller watershed than that of channel 5 with higher order than channel 3. So channel 3 is removed but channel 5 remained. The generalization result is the same as manual operation.

3.1.2 Watershed Extraction

The rainfall down on the earth collects into basin or river. The watershed area of one river channel reflects the ability of its corresponding catchment to compete for the rainfall. The watershed line is the competition result between neighbor rivers. In geographic analysis we usually use Voronoi diagram or Delaunay triangulation (Preparata and Shamos 1985) to study such spatial competition question. In map generalization, they are also good solutions in such as spatial conflict judgment, neighbor object aggregation and local displacement (Ai and Oosterom 2002, Jones 1995, Ware and Jones 1995, Poorten and Jones 1999, Bader and Weibel 1997). Next we first generate the watershed line between proximity channels and then link the sequent lines to build the inclusion of watershed polygons. The whole process contains four steps.

____ **Triangulation construction.** Get all vertex points of catchment to make a point set *S*. For those segments containing too long direct line between two points, we interpolate serial middle points between terminal points to generate additional points to avoid the intersection between triangle edge and river segment in later triangulation construction. Add three outside points making a triangle to envelope the point set *S* and construct the Delaunay triangulation of S. Remove those triangles which are related to outside three points and the remained triangles compose the coverage of the river catchment as shown in figure 4. If the distance between two neighbor points is short enough, the triangulation does not result in the intersection between triangle edges and the channel segments. Otherwise replace the normal DT with the constrained DT.

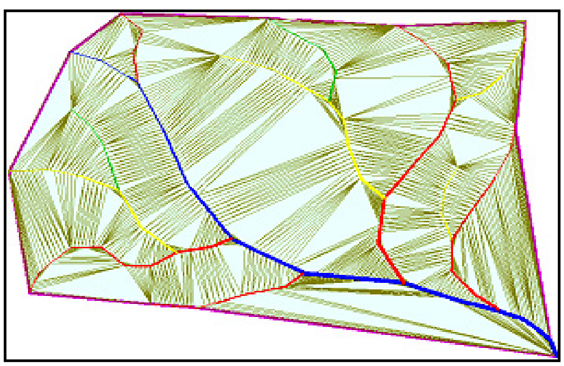


Fig. 4. Construct the Delaunay triangulation in the coverage of river catchment.

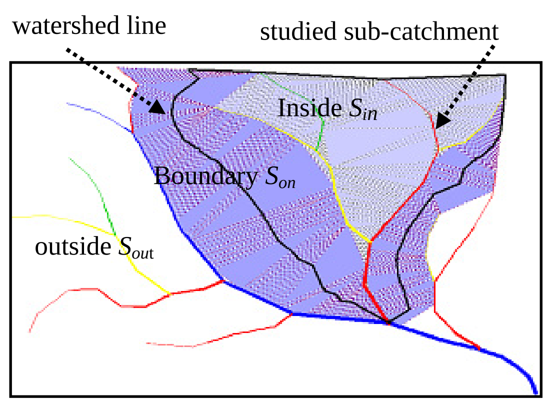


Fig.5. For one sub-catchment, identify three sorts of triangles: inside, outside and boundary triangle regions.

_____**Triangle Classification.** One channel *a* together with its descent channels makes a sub-catchment as part of its parent sub-catchment. According to the relationship between a triangle and the sub-catchment *a* , we can divide the triangles into three classes. Take one segment *b* of sub-catchment *a* into account. The triangle with at least one vertex locating on the segment *b* is assigned to be segment-related triangle. All segment-related triangles of sub-catchment *a* (including current channel *a* and its descent channels) are assigned to be sub-catchment-related triangles which makes the coverage region of sub-catchment *a* .

These triangles are further able to be classified as two types. One is the completely related triangle with all three vertexes locating on the segments of sub-catchment *a*, and partially related triangle with one or two vertexes on the segments of sub-catchment *a* . The other vertex of partially related triangle locates either on the segment of parent channel of sub-catchment *a* or the segments of brother sub-catchments of sub-catchment *a* . Finally for one sub-catchment *a*, the triangles are classified as three types, namely (1) the outside triangles S_{out} without relation to the sub-catchment *a*; (2) the inside triangles S_{in} being completely related to *a*, and (3) the boundary triangles S_{on} partially related to *a*. Three types of triangle are illustrated in figure 5. The core red-colored channel is currently studied channel. The light blue shaded triangle region is the inside S_{in} and the deep blue shaded triangle region the boundary S_{on}. The other white region belongs to the outside S_{out} .

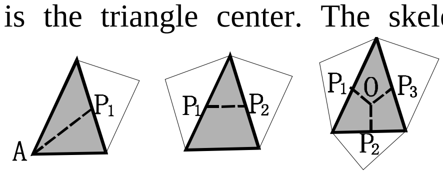
Suppose the terrain slope, the soil filter and the vegetation abstraction distributes in a uniform way. The watershed area will be determined by the spatial relation of river channels. There exists the close principle that the rainfall finds the nearest path to collect into basin or river. Take the sub-catchment *a* into account, the rainfall on the S_{in} region will completely flow into *a* no matter the flow directly down to channel *a* or through its descent channels. The rainfall on S_{out} region will completely down flow to other sub-catchments having nothing to do with sub-catchment *a*. But for S_{on}, the rainfall faces the competition since it locates as a bridge between sub-catchment *a* and context neighbors. The buffer area as shown in deep shaded in figure 4 needs to be divided into two parts by some way.

______**Watershed Line Extraction**. Based on the analysis above, the watershed extraction should be conducted in the area of boundary triangles S_{on} . We use the skeleton of DT method (Ai and Oosterom, 2002)to extract the watershed line. Just consider the triangles in S_{on} and distinguish them three types according to the number of neighbor triangles, namely type I with only one neighbor, type II with two and type III with three. The skeleton connection way for three types of triangle is described in figure 6, where P_1, P_2, P_3 is the midpoint of corresponding triangle edge, and O is the triangle center. The skeleton segments are created by means of the next paths:

Type I : $A \rightarrow P_1$;

Type II: $P_1 \rightarrow P_2$;

Type III: $O \rightarrow P_i$, i=1,2,3



Type I Type II Type III

Fig. 6. Center-line connection ways for three types of triangle.

The skeletonization result of S_{on} is illustrated as black line, namely the watershed line, in figure 4. As the skeleton line is closed, the watershed area automatically generates.

Trace the channels of river tree one by one to repeat the method above. After all sub-catchments have been processed, we finally get the watershed line distribution result as shown in figure 7.

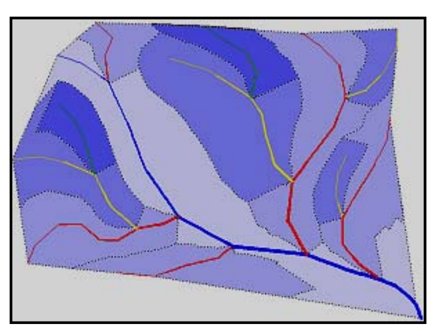


Fig. 7. The result of hierarchical partitioning of the river catchment

____ **Inclusion relation generation.** The hierarchical structure of river channels acts as the minor channel joins into the main channel. This hierarchical relationship is mapped as the inclusion between watershed polygons. It means the sub-catchment a with child sub-catchments b1, b2,…bn corresponds to the watershed polygon of a enveloping that of b1, b2,…bn. Based on this association, it is easy to build the inclusion relationship of watershed polygons. The child watershed polygon must be included within the parent watershed polygon. But the integration of all child watershed does not equal to the parent watershed. Some regions belong to the parent channel rather than any child channel. At same level, the watershed of brother channels does not overlap to each other. Note that the boundary line between neighbor watershed be exactly the same, since it is extracted from the same sub-triangulation by the same skeletonization method. Figure 8 right illustrates the inclusion among the hierarchical watershed by the different level of shaded color.

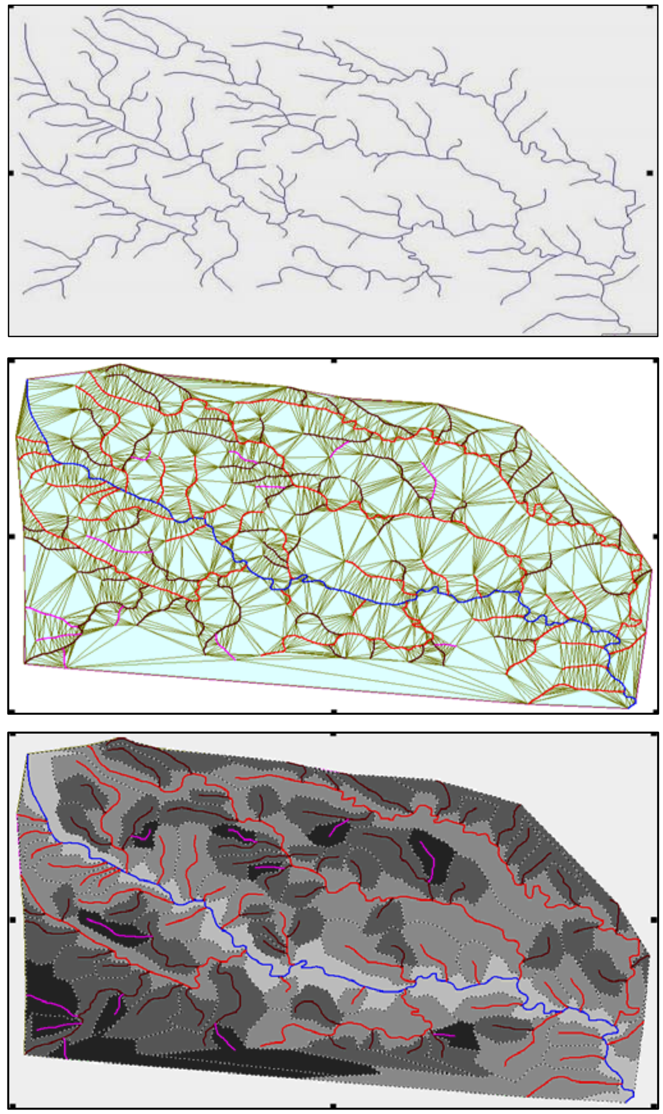


Fig. 8. The construction of Delaunay triangulation within a catchment in middle and the watershed extraction which is illustrated as hierarchical shaded area at bottom

After four steps of partitioning, every channel with a sub-catchment obtains a watershed polygon. According to the spatial competition analysis above, the polygon size describes the significance of corresponding channel. Finally we sort the channels on the watershed area to a linear sequence at decrement order as a preparing work for the latter multi-scale representation model.

3.2 Decomposing individual channel by BLG tree

The sort on watershed area has organized the channels into a sequence at significance decrement. Next let the hydrographic data to further decompose the individual channel segment into different details. The data model needs to provide not only different number of significant channels but also the channel representation with different details. The latter decompose is related to line simplification. We apply the Douglas-Peucker (1973) algorithm and by BLG tree(Oosterom 1994) to enhance the multi-scale representation model.

BLG tree is a binary tree structure to store the series of simplified result of Douglas-Peucker algorithm(Oosterom 1994). The tree node with a record of offset distance represents the role of corresponding point in line representation. The root corresponds to the first point separating the initial line when the algorithm applying. Extracting parts of nodes by BLG tree trace obtains the line representation with different details. So by off-line BLG tree construction we can get a data structure to output the line representation with different resolution, as shown in figure 10.

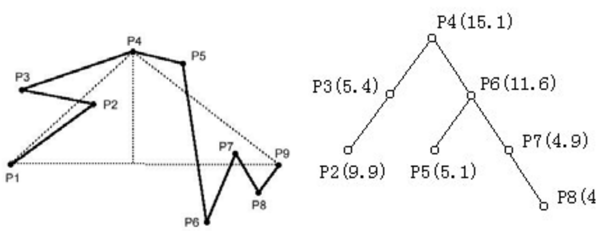


Fig 9. A polyline example and corresponding BLG tree

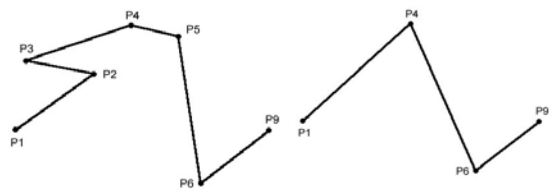


Fig 10. Selecting nodes by offset threshold 5 (left) and 10(right) obtains two representations

One question in original BLG tree is that the offset distance of parent node may be smaller than that of its child nodes not guaranteeing the top level node has larger offset than that of low level nodes. For example in figure 9 the offset 5.4 of node P3 is smaller than that of its child node P2, namely 9.9. It implies that the node selection based on the offset size will take child node but missing parent node. In order to avoid this situation, we modify the offset of parent node to be same as its child node. In figure X, the offset of P3 is changed to 9.9 same as its child node P2.

Trace the BLG tree and record the nodes in the order of offset decrement and let the start and terminal point of original line be recorded firstly by pre-setting its offset a large value. We can get a linear data structure as shown in figure 11. From left to right selecting some nodes whose offset is larger than the predetermined threshold is able to derive different representations of channel segment. Farer away from the begin position the obtained representation is closer to the complete representation. The selected points should be adjusted the sequence to make the line representation as series of consecutive points.

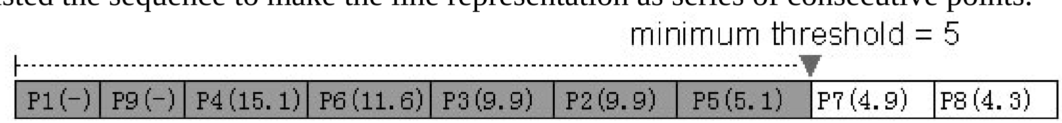


Fig. 11 Select nodes by offset threshold 5 from the linear BLG-Tree

3.3 Building the multi-scale representation matrix

Based on two linear structures, namely the channel order at object level and the linear BLG tree at geometric detail level, we can now construct the multi-scale data model to represent the hydrographic network. The data model reflects as a matrix structure. The row from top to bottom represents the channels in the sequence of watershed area decrement. The column represents one channel (river) by series of linear BLG tree structure to describe the line representation with detail refined. As one channel is composed of several segments, one column has several linear BLG tree structures. The matrix of multi-scale representation of hydrographic data is shown in figure 12.

In the multi-scale representation matrix, the row order and the column order are both controlled by variable scale. It means for given scale, the model can automatically compute which row channels will be output and in what detail refined by column selection. For the quick response in latter progressive transmission, the parameter scale needs to be inserted into the row and column record in this matrix.

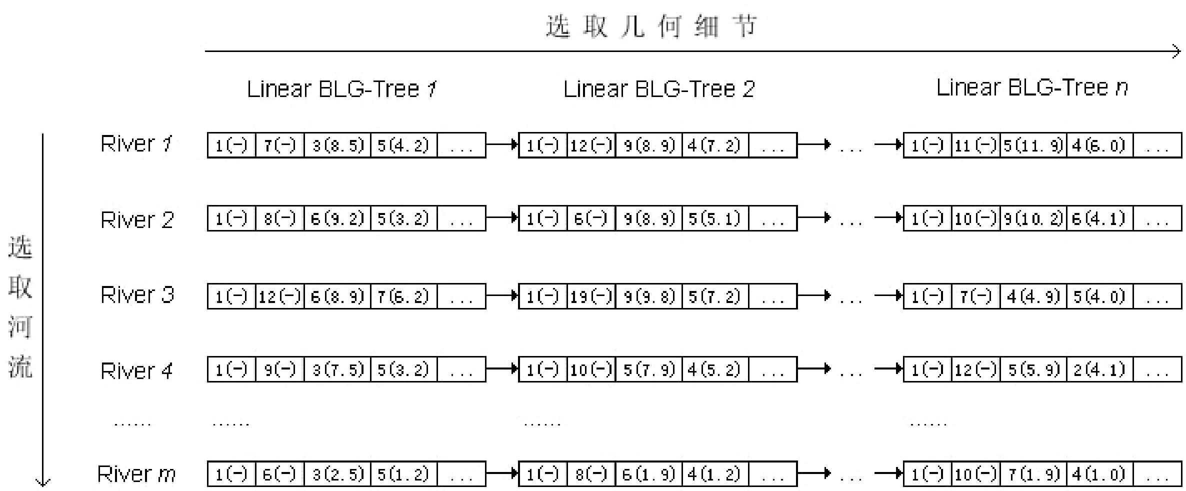


Fig 12. An example of matrix of multi-scale representation of hydrographic network.

For given row which corresponds to one channel element, it has a qualification scale to express when to be selected. Down this scale it will be removed and above this scale be selected. We use the The Töpfer law (Töpfer and Pillewizer, 1966) to determine the qualification scale for each row. Suppose the initial large scale is M_0 (scale denominator) and at this scale all channels will be selected with the number N_0. According to Topfler law, for given scale M_x, the number of selected object should be

$$N_x = N_0(M_x/M_0)^{1/2}$$

Then the scale

$$M_x=M_0\ (N_x/N_0)^2$$

It means that the qualification scale M_x of row N_x is determined by the number N_x, because the accumulation number of selected object before

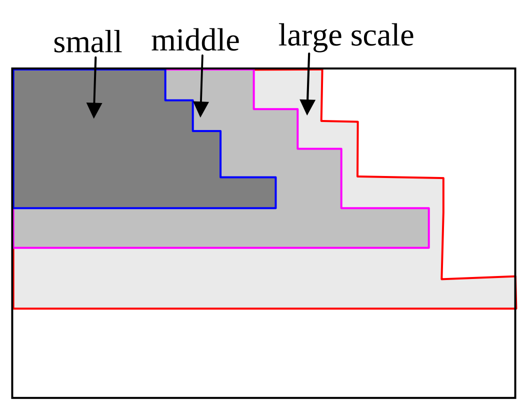


Fig 13, The selected area in matrix expands as scale increases

the current row is $N_x\text{-}1 \approx N_x$. So we can apply the equation above computing the qualification scale and attaching it to the record of each row.
For given column of matrix, each point in linear BLG tree structure has a record of offset distance which is related to visual resolution. The large offset distance implies the point in curve line strongly impacts on the visual cognition and should be selected at even small scale. According to the minimum graphic element to be identified for common users, it is easy to map the offset distance to scale parameter. All scale parameters set for the point number in linear structure should be larger than the initial scale determined by the row, channel object selection.

After being added the parameter scale, the matrix of multi-scale representation can derive proper representation for given scale by first gradually scanning rows downward and then scanning one row from left to right column until accessing to the scale threshold. As the scale increases gradually, the selected element data in the matrix acts as the gradual expanding area around the top left corner as illustrated in figure 13. And the corresponding representation is illustrated in figure 14

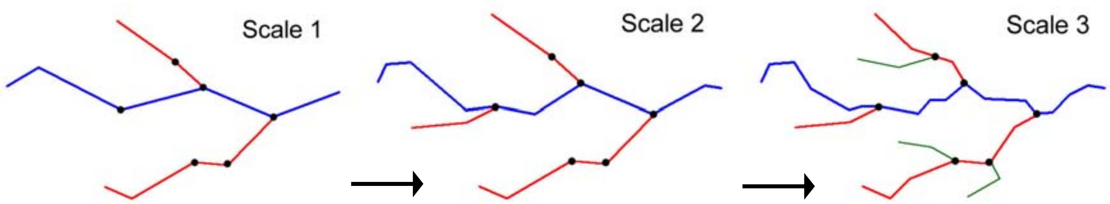


Fig 14. The derivation of different representations at different scales based on the matrix above

3.4 Applying in progressive transmission

The multi-scale representation matrix model considers the scale impact on data representation able to derive different representations with different resolutions. It is able to support the progressive transmission over web well. We develop an experiment system to conduct the progressive transmission of hydrographic network. Some snaps during the data transmission are extracted as shown in figure 15. Moving the scale button, the dynamic displayed hydrographic data results in an animation not only adding the channel segment gradually but also converting the line graphic more and more complex.

To apply in the progressive transmission, the multi-scale representation matrix is stored on server site and let scale parameter control data output. After receiving the data application and the parameter scale denominator M_n, the server first generates a sequence set of scale $\{ M_0, M_1, M_2,\ldots,M_n\}$ with $M_0 < M_1 < M_2 <,\ldots, < M_n$, then for each scale $M_i(0 \leq i \leq n)$ scans the multi-scale representation matrix to extract the data composition for scale M_i . Finally a series of data version is bound and transmitted to client over web. More number the scale separates, higher granularity of data version outputs. It means the multi-scale representation matrix model has prepared high granularity data but the data output is dynamic able to adjust the resolution. The number of scale separation depends on the resolution requirement from user, the data volume and the web speed. For a given web environment the system can check the web speed and determine how many scales to be separated making the progressive transmission a self-adaptable system.

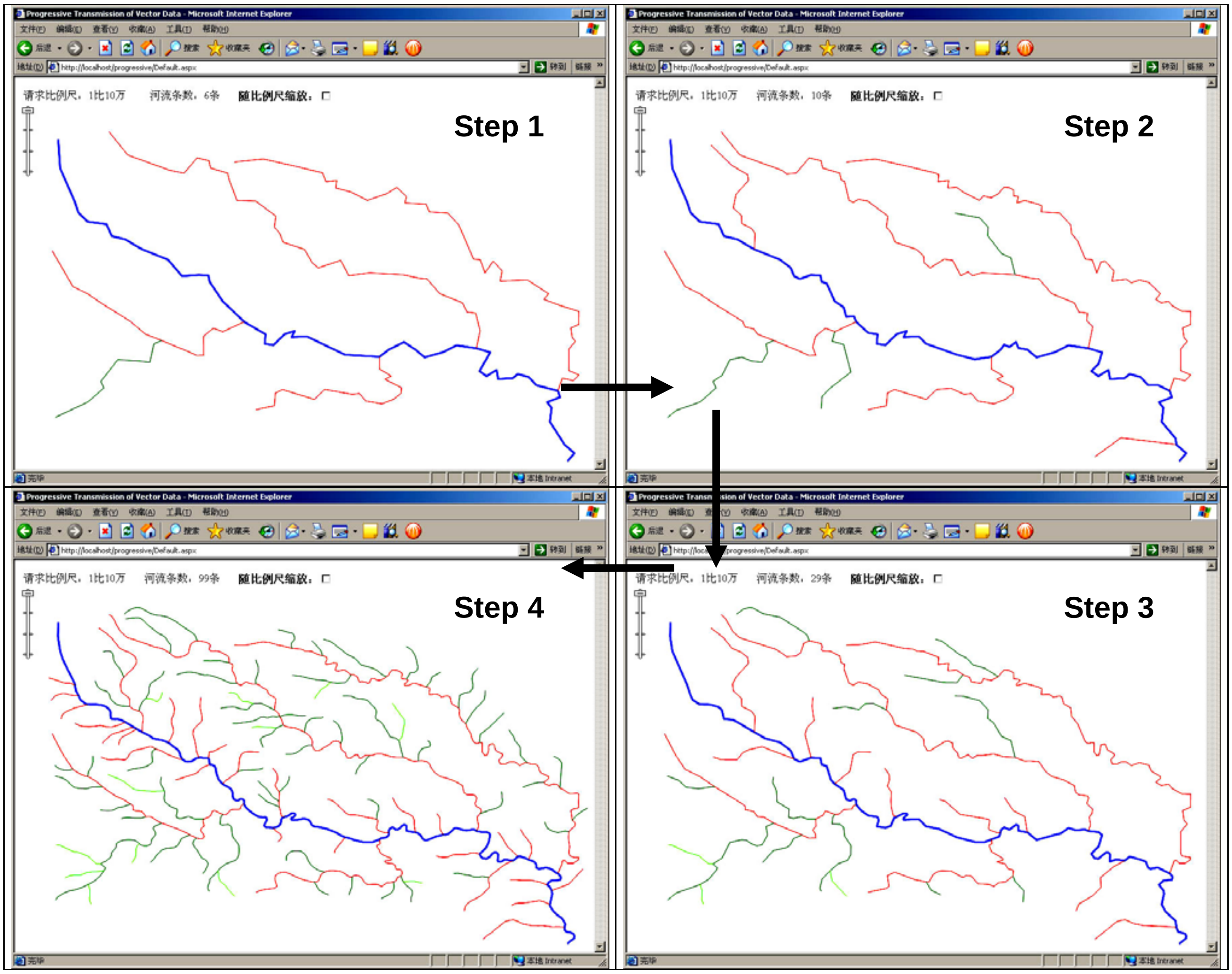


Fig. 15. The application of multi-scale representation matrix in progressive transmission over web

The application of multi-scale representation matrix in progressive transmission has the following characteristics:

______ **Quick response**. The dynamic data stream is just produced from scanning an existed multi-scale representation matrix which has been built off-line without complex operations to derive the versions of data representation. The Douglas-Peucker algorithm is used to simplify the channel segment. But in progressive transmission just retrieves the simplified result rather than actual executing the algorithm. The computation complexity of scanning the matrix is O(n). So the multi-scale representation matrix can quickly respond to the data application from client.

_____ **High granularity change**. Among three levels of data granularity for progressive transmission, namely the feature class, object and geometric detail level, the matrix model represents the channel segment with point add accessing to geometric detail level. In some GIS software, such as ArcMap, the system provides the multi-scale visualization strategy by attaching a qualification scale to every object to be displayed. The object can be displayed only when the visualization scale reaches to some value. But after visualization, the object will keep unchanged even when the scale changes. Our model can make the visualization an animation with high granularity.

______ **Small data volume**. Compared with the traditional GIS data representation of line and network, this model does not attach additional information except for the scale description for every object and every point. The matrix model just adjusts the sequence of point series in the channel segment to make a new point stream, namely the linear BLG tree. So the data volume increases not significantly. Also the data consistency such as the topological relationship between channel segments, is kept well across different representations. Instead to store multi-version of complete data for different scale is a simple method to realize multi-scale representation. But the data volume increases greatly and the consistency maintenance of different representations is hard.

4. Conclusion

The progressive transmission provides users with a self-adaptive method to access data over web. 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 transmission method. But the progressive transmission of vector data is still an open question. One challenge is to build the efficient multi-scale representation data model. The multi-scale representation and hierarchical organization of vector data is a key technology for progressive transmission. Indeed, the progressive transmission can be regarded as a mapping process of data representation from spatial scale to temporal scale. The data details separated on the basis of spatial scale is then transmitted in time range domain. Each snapshot in time domain corresponds to one representation at certain spatial scale. Taking a longer time to wait will get the representation closer to complete data.

In this paper we discuss the characteristics of progressive transmission of vector data, investigating the constraints in transmission granularity separation, data volume compression and response at real time. Considering these constraints we present a data model, namely the multi-scale representation matrix, to realize the progressive transmission of hydrographic network data. The channels under catchment are sorted on the watershed at increment order and each segment is re-organized the point sequence. The previous organization is based on the Delaunay triangulation to partition the watershed and the latter the BLG tree separation. Two ordered organizations are integrated in a matrix with row describing channel object and column the individual segment in geometric detail level. The study investigates the watershed partitioning in detail with experiment illustrations.

Due to the difference of geographic feature representation, we can not build a common multi-scale model for all GIS data. The progressive transmission should apply different strategies for different features. A basic idea is integrating the existed models and algorithms in map generalization domain and similarly considering the constraints in data volume compression, quick response and others to realize the progressive transmission. The multi-scale representation matrix in this study is an example to combine Delaunay triangulation method and BLG tree structure together to progressively transmit the hydrographic data which has a network pattern with hierarchical structure. Two methods are good solution in map generalization. This idea can be extended to other features to settle the question of progressive transmission.

References

1) Ai T, Liu Y, and Chen J. 2006. The Hierarchical Watershed Partitioning and Data Simplification of River Network[A]. In proceedings of the 12th Advances in Spatial Data Handling, Vienna: Springer-Verlag, pp617-632.

2) Ai, T., Li, Z. and Liu, Y. 2004. Progressive transmission of vector data based on changes accumulation model, Proceedings of the 11th International Symposium on Spatial Data Handling, SDH 2004. Springer-Verlag, Berlin, pp85-96.

3) Ai, T. and Oosterom, P. van. 2002. GAP-tree Extensions Based on Skeletons. In: Richardson D and Oosterom P van(eds) Advances in Spatial Data Handling, Springer-Verlag, Berlin, pp501-514.

4) Bader, M. and Weibel, R. 1997. Detecting and Resolving Size and Proximity Conflicts in the Generalization of Polygonal Maps. In: Proceedings of the 18th ICA/ACI International Cartographic Conference, Stockholm. pp1525-1534.

5) Ballard, D. 1981. Strip Trees: A Hierarchical Representation for Curves. Communication of the Association for Computing Machinery, vol. 14: 310-321.

6) Bertolotto, M. and Egenhofer, M. 2001. Progressive Transmission of Vector Map Data over the World Wide Web. GeoInformatica, 5 (4): 345-373.

7) Bertolotto, M. and Egenhofer, M. 1999. Progressive Vector Transmission. Proceedings, the 7th International Symposium on Advances in Geographic Information Systems, Kansas City, MO: pp152-157.

8) Buttenfield, B. P. 2002. Transmitting Vector Geospatial Data across the Internet, Egenhofer, M.J. and Mark, D.M (eds.) Proceedings GIScience 2002. Berlin: Springer Verlag, Lecture Notes in Computer Science No 2478: 51-64.

9) Buttenfield, B. P. 1999. Sharing Vector Geospatial Data on the Internet. Proceedings, 18th Conference of the International Cartographic Association, August 1999, Ottawa, Canada, Section 5: 35-44.

10) Cecconi, A. 2003. Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping. PhD Thesis, University of Zurich, Zurich.

11) Douglas, D. H. and Peucker, T. K. 1973. Algorithms for the Reduction of the Number of Points Required to Represent a Line or Its Caricature. The Canadian Cartographer, vol. 10(2): 112–122.

12) Han, H., Tao, V. and Wu, H. 2003. Progressive Vector Data Transmission, Proceedings of 6th

AGILE, Lyon, France.

13) Jiang, B. and Ormeling, F.J. 1997. Cybermap: the Map for Cyberspace. Cartographic Journal, 34 (2):111-116.

14) Jones C B, Kidner D R, et al. 1996, Database Design for a MultiScale Spatial Information System. International Journal of Geographical Information Systems, 10(8):901-920.

15) Jones, C. B., Bundy, G. L., and Ware, J. M. 1995. Map Generalization with a Triangulated Data Structure. Cartography and Geographic Information System. 22(4): 317-331.

16) Matthias Bobzien, Dirk Burghardt, et al, 2008. Multi-Representation Databases with Explicitly Modelled Intra-Resolution, Inter-Resolution and Update Relations. Cartography and Geographic Information Science, 35(1):3-16.

17) Muller, J. C. and Wang, Z. 1992. Area-path generalization: A competitive approach, The Cartographic Journal, 29(2), 137-144.

18) Oosterom, P. Van .1995. The GAP-tree, An Approach to On-the-Fly Map Generalization of An Area Partitioning. In: Muller J C, Lagrange J P, Weibel R (eds) GIS and Generalization: Methodology and Practice. London: Taylor & Francis. pp 120-132.

19) Oosterom, P. van, 1994. Reactive Data Structure for Geographic Information Systems. Oxford University Press, Oxford.

20) Poorten, P. and Jones, C. B. Customisable Line Generalization Using Delaunay Triangulation, CD-Rom Proceedings of the 19th ICC, Ottawa, Canada, Section 8,1999.

21) Preparata, F. P. and Shamos, M. I. 1985. Computational Geometry An Introduction, Springers-Verlag.

22) Rauschenbach, U., Schumann, H. 1999. Demand-driven Image Transmission with Levels of Detail and Regions of Interest. Computers & Graphics, 23(6): 857-866 .

23) Richardson, D. E. 1993. Automatic Spatial and Thematic Generalization using a Context Transformation Model. PhD Thesis, Wageningen Agricultural University

24) Sarjakoski, T., 2007, Conceptual Models of Generalisation and Mulitple Representation. In Generalisation of Geographic Information: Cartographic Modelling and Application (Elsevire Ltd), p11-35.

25) Sester M. and Brenner M. 2004.Continuous Generalization for Visualization on Small Mobile Devices. Proceedings of the 11th International Symposium on Spatial Data Handling, SDH 2004. Springer-Verlag, Berlin pp355-368.

26) Srinivas, B. S. R., Ladner, M., Azizoglu, and Eve A. Riskin. 1999. Progressive Transmission of Images using MAP Detection over Channels with Memory. IEEE Transactions on Image Processing, 8(4): 462-475.

27) Stell, J.G. and M.F. Worboys. 1998. Stratified Map Spaces: A Formal Basis for Multi-resolution Spatial Databases. Proceedings of the 8th International Symposium on Spatial Data Handling, SDH 1998. Columbia (British): Taylor and Francis pp180-189.

28) Steven S. S. 1946. On the theory of scales of measurement. Science . vol 103: 677-680.

29) Taylor, D. 1997. Maps and Mapping in the Information Era. In: Ottoson L.(ed) Proceedings of the

18th ICA/ACI International Cartographic Conference, Stockholm, Sweden, Gävle. pp 23-27

30) Thomson, R. C. and Brooks, R. 2000. Efficient Generalization and Abstraction of Network Data Using Perceptual Grouping. Proceedings of the 5th International Conference on GeoComputation,

31) Töpfer, F. and Pillewizer, W. 1966. The principles of selection: a means of cartographic generalization, The Cartographic Journal, 3(1):10-16

32) Tribe, A. 1992. Automated Recognition of Valley Lines and Drainage Networks from Grid Digital Elevation Models : A Review and a New Method, Journal of Hydrology, vol. 139 : 263-293,

33) Ware, J. M. and Jones, C. B. 1995. A Triangulated Spatial Model for Cartographic Generalization of Areal Objects, In: Proceedings COSIT'95, Semmering, Austria, pp 173-192.

34) Wolf, G. W. 1988. Weighted Surface Networks and Their Application to Cartographic Generalization, Visualization Technology and Algorithm, W. Barth (ed.), Berlin : Springer-Verlag : pp199-212.

35) Wu, H. 1997. Structured Approach to Implementing Automatic Cartographic Generalization. Proceedings of the 18th ICC, Stockholm. Sweden□Vol.1: 349-356.

36) Zhou, S. and Jones, C. B.2003. A multi-representation spatial data model. Advances in Spatial and Temporal Databases, Proceedings of the 8th International Symposium on Spatial and Temporal Databases SSTD pp394-411.