HYPERMAP-BASED DISTRIBUTED SPATIAL DATA MODEL

Hui LIN* Haiyang HAN* Jianya GONG**
*Chinese University of Hong Kong, Hong Kong
Joint Lab for Geoinformation Science
Hhy@mail.jlgis.cuhk.edu.hk
**Wuhan Technology University of Surveying and Mapping, P.R.China
Lab for Information Engineering in Surveying, Mapping and Remote Sensing
jgong@rcgis.wtusm.edu.cn

KEY WORDS: distribution, GIS, hypermap, data model, interoperability

ABSTRACT

Distribution, interoperability, interactivity, component are the main four features of distributed geography information system. Based on the principle of hypermap, hypermedia and distributed database, the paper comes up with a kind of distributed spatial data model which is in accordance with those features of distributed geography information system. The model takes catalog service as the outline of spatial information globalization, and defines data structure of hypermap node in different level. Based on the model, it is feasible to manage and process distributed spatial information, and integrates multi-source, heterogeneous spatial data into a framework. Traditionally, to retrieve and access spatial data via Internet is only by theme or map name. With the concept of the model, it is possible to retrieve, load, and link spatial data by vector-based graphics on the Internet.

1 INTRODUCTION

Spatial distribution is the natural feature of geography information. In the past years, due to limitation of technology, communication and hardware, spatial data captured from different region has to be saved, managed and processed in the central place. As a result, spatial data can not meet the requirement of update and share [Kraak, 1996]. Recently, Hypertext-based browsers gradually become the doors of Internet. It is convenient and popular to access distributed information by hyperlink on the Internet and www. However, only hypertext protocol can not meet the requirement of GIS. Combining with the idea of hypertext and hypermedia, Laurini and Milleret-Raffort brought up the concept of hypermap and applied to GIS in 1990[Caporal, 1995][Huffmann, 1993]. It will largely push the development of DGIS (Distributed Geography Information System) which is one of important trends in the future.

Combined with some other definitions of hypermap such as Laurini and Milleret-Raffort [Nielsen, 1995], a descriptive definition is given as follow. Hypermap means that map, which is regard as a kind of media on the Internet, can interact with other hypermedia by hyperlink among different types of media. As far as map media is concerned, a feature such as point, line or polygon in the map can be treated as hyperlink node. The node may link to different type of media. Zooming, clicking or querying the specific hyperlink node will access sub-map or other media information. Such media may be text, voice, image, virtual reality, vector-based map, video, image and so on.

At present, hypermap-based DGIS is only applied to raster-based map (or image)[IconAuthor, 1995]. Prior to publish map on the Internet, a polygon area of an image is set as hot area. It represents the geographical range of hyperlink node, and hyperlink linked to target media is set [Jiang, 1995]. Such procedures can be finished by most of popular commercial software such as FrontPage. Most hypermap-based applications are like those. However, it is not suitable for vector-based map in this way.

The paper will try to point out limitation of such a way and potential trend of DGIS by analyzing the nature of distributed geography information system (DGIS), and bring up hypermap-based distributed spatial data model by combining with hypermap, distributed database and catalog service.

2 PROBLEMS OF DGIS

Distribution reflects the natural feature of spatial data for DGI. A descriptive definition is given as follow. DGI is Internet and TCP/IP-based GIS. All kinds of distributed spatial information, including vector-based map, image, data sets, analytical operation, text report, etc., can be accessed and further processed such as querying, mapping, analyzing spatial data on the Internet. Furthermore, products and services in relation to spatial information are provided for public and specialists via Internet in order to farthest share spatial data and its operations. DGI is made up of client-side, server-side and network communication. Interoperability, distribution, interactivity, component are its main features.

As far as DGI is concerned, distribution implies that data and corresponding operation are located in difference places; component represents the architecture of system construction; interoperability contributes to heterogeneous data and operations share [OGC, 1997]; interactivity represents bi-directional information flow by hyperlink among different types of hypermedia. Among those features, distribution is dominant. It reflects the natural features of DGIS, and will be the potential trend of DGIS. Furthermore, distributed construction make it more feasible to inter-operate spatial data than traditional way. DGIS represents its distributed architecture by component construction and message communication; Distribution consists of data distribution and operation distribution.

Present DGIS is only in the primary stage and still has a long distance to go. Especially features in relation to distribution are not reflected and represented. Although data are captured in distributed regions, they are saved and managed in a central database. On the one hand, Heterogeneous raw data need to be transferred in advance. It is a time-consuming work. On the other hand, it is not convenient for data updating, especially against the requirement of information age, and all processing have to be finished in a central place. As far as GIS server is concerned, it may be cracked down when too many requests are received, furthermore, it is hard to process data in the most suitable Internet node in order to decrease the capacity of data flow and improve system efficiency.

Due to lack of catalog service and project, it is difficult to organize and manage distributed spatial information on the Internet. At present, most DGIS can only process isolated workplace map, and the relationship among different workplaces maps are not under consideration. It goes without saying spatial analysis based on distributed spatial data. To publish spatial data on the Internet will be the trend in the future; however, it is impossible to meet the requirement of constructing distributed national and province spatial data infrastructure in traditional way.

In addition, with the development of Internet, hypermedia become the mainstream mode of publishing information on the Internet, and Internet protocol, supporting vector-based map, will be implemented by w-3c consortium. In the near future, it is feasible to directly browse vector-based map by browser without the support of any other DGIS software. Hypermap will be as popular as hypertext or hypermedia.

What's the reason of those problems? It is the lack of distributed spatial data model. In the past years, many researches have focused on the construction of traditional spatial data model and data structure such as object-oriented model, object-relation model and so on. However, DGI become popular only in recent years and no corresponding theory is suitable for storage, organization and management of distributed spatial data. Traditional spatial data model can not meet the requirement of constructing distributed spatial data model.

3 HYPERMAP-BASED DISTRIBUTED SPATIAL DATA MODEL

The prime problem need to be solved is to look for an appropriate data structure and data model in order to develop DGI. Internet brings mountains of raw data, at the same time, it also brings up the problem of how

to organize and manage information. GIS data is also not an exception. How to access distributed, heterogeneous spatial information, and process, manage those spatial data as if they are in local machine? It is necessary to develop distributed spatial data model and data structure according to features of distributed information in order to solve those problems.

The objective of hypermap is to link different object, especially vector-based map and spatial data by hyperlink and hypermap node (hypernode). In this way, it is possible to access distributed spatial information. However, as the same as hypertext, too much hyperlinks will form a complicated network if without any organization. It goes without saying users; even designer will lose his way under such a hyperlink-based network. Furthermore, although browser supports display and link of raster-based image, it does not support vector-based data format. How to define the data structure of vector-based hypernode and display a map in a browser is also a problem need to be further discussed.

Following solution for distributed spatial data can be generalized in combination with hypermap, distributed database, and the idea of organizing and managing spatial data by project in traditional GIS.

Generally, traditional GIS data model is to integrate multi-workplace into a project and process analytical functionality in a project. Project index, catalog and workplace information, and data is saved in central database. Under Internet environment, spatial data is located in different places. Catalog service, combined with the concepts of project and workplace in traditional GIS can implement organization and management of Distributed spatial data. In addition, the hypermap-based relationship among distributed spatial information can be constructed by defining hypernode object. Therefore, a complicated network model, which is based on hierarchical structure and hypermap, is formed. With the support of distributed multi-server to process distributed spatial data, database and corresponding geoprocessing are both distributed in server-side, while the operations in client-side is centralized. User may access and process distributed spatial information only by a browser.

3.1 Concepts

The concepts of distributed spatial data model include hyperroot node database, hypernode database, hyperproject, hyperworkplace, hypernode, hyperlink, catalog service database, metadata database. Definition in detail is as below.

[hypernode] hypermap node. The node may be point, line, polygon or area in a vector-based map. It can link to other vector-based map or hypermedia by hypermedia. The relation of different workplaces and media is built up by hypernode. As a specific object, it reflects the link relation towards specific region or spatial object in the map. Its attributes include features of link object and relationship of hyperlink (accordance with TCP/IP protocol).

[hyperworkplace]hypermap workplace. The set of multi-layer in a vector-based map. A hypernode, which is defined as specific region or spatial object in a hyperworkplace, can be linked to other hyperworkplace or hypermedia by hyperlink. Certainly, the hyperworkplace can be linked by other hypernode too.

[hyperproject]hypermap project. The set of distributed multi-hyperworkplace. With the concept of hyperproject, the spatial data from distributed hyperworkplace is integrated and processed in a unified interface. It is the extension of concept of local project under network circumstance.

[hypernode database]hypermap node database. Any computer connected to Internet is named after a network node on the Internet. Such a node provides and accesses information on the Internet. If network node acts as database server and is permitted to access database, then the database in accordance with the network node is named after node database. As far as GIS is concerned, spatial information is saved in spatial database, accordingly, database server is called hypernode databse. Hypernode database is a set of thematic spatial information in a specific region or themes.

[hyperroot database] hypermap root node database. The all and the one of hypernode databases.

[catalog service database] it aims to provide catalog service for nodes of distributed spatial model. As a result, all nodes are treated as an interactive integration. Catalog service of hypernode database provides hierarchical structure for distributed spatial model and IP address of hypernode database server. Catalog services provide index of all hyperworkplace and hypernode.

[metadata database] it provides essential information of hypernode database, hyperproject and hyperworkplace. The end-user can query spatial information necessary and its location on the Internet, and Administrator in server-side may organize hyperlinks of hypernode database and hypernode.

There are two different levels for distributed spatial data model. Hierarchical structure, which consist of hyperroot database and hypernode database, acts as the skeleton of network model; hyperproject, hyperworkplace and hypernode, which is related by hypermap-based link in different hypernode database, acts as the extension of network model. A network-based model is constructed under the basis of combination as above. Catalog service database and metadata database save relation information of model.

More than one hypernode database can be constructed in different regions for distributed geographic information. If global spatial information is taken as a whole and hypernode databases are constructed according to specific region, there is one hyperroot database and many hypernode databases. In addition, different hyperworkplaces are linked by hypermap links. Therefore, any organization or person can define his own spatial database as a hypernode database, and add it into the network model as a hypernode according to pre-defined regulations. The distributed spatial data model, constructed under such an architecture as above, is the combination of hierarchical structure and hypermap-based link. Catalog service database will be responsible for the update and management of hyperworkplace, hyperproject and hypernode database.

3.2 Distributed Spatial Sata Model

According to concept introduction and model basis as above, a distributed spatial data model as fig. 1 can be constructed. On the one hand, from the whole view, the model is based on hierarchical structure, so it benefits organization and management of spatial data of specific region; on the other hand, from a local view, hypernode and hyperlink construct relationship of spatial data located in different places. As a result, hypermap-based link extends original hierarchical structure, and relationship among distributed spatial information can be reflected.
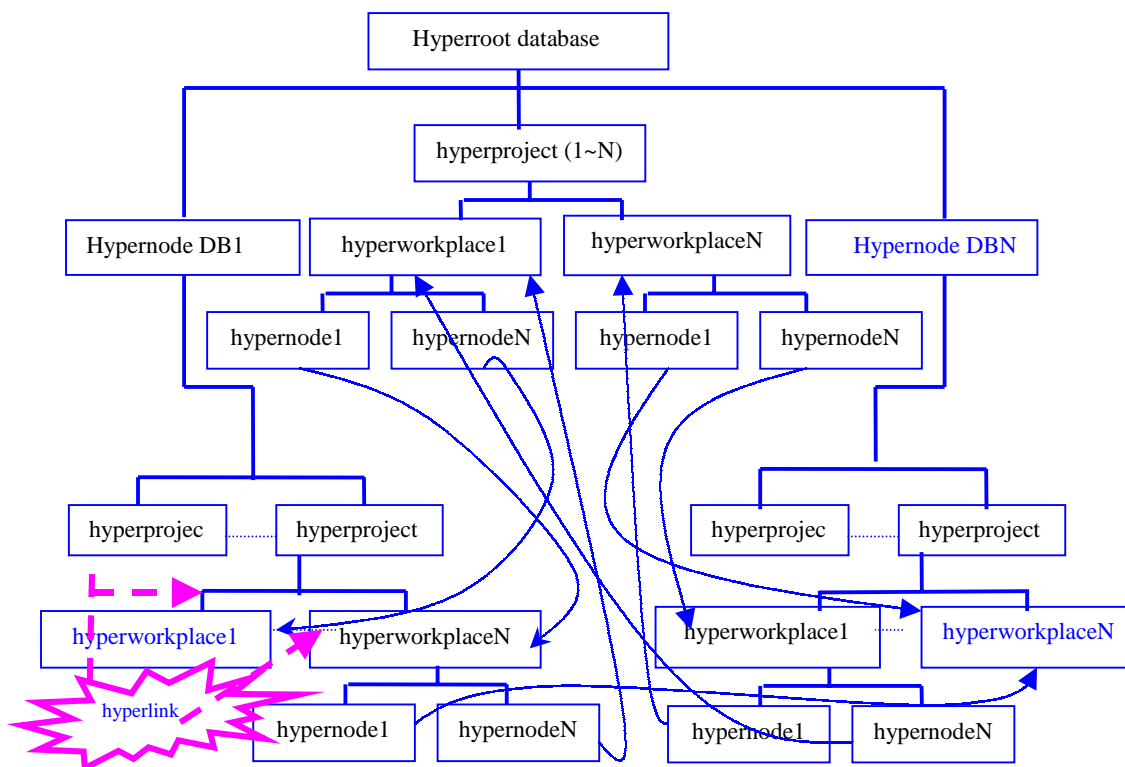
Figure1: DGI spatial data model

Hypernode database organizes and manages hyperprojects and hyperworkplaces by hierarchical structure. Its catalog service, based on hyperlink, can manage distributed spatial data efficiently. The organization structure is as below.

```
┌──────────────┐     ┌──────────────┐     ┌──────────────┐
│Catalog service│     │ Thematic map1 │     │  hypernode1  │
├──────────────┤     ├──────────────┤     ├──────────────┤
│ Metadata DB  │     │ Thematic map2 │     │  hypernode2  │
├──────────────┤     ├──────────────┤     ├──────────────┤
│ hyperproject1│     │ Thematic mapi │     │  hypernodeN  │
├──────────────┤     ├──────────────┤     ├──────────────┤
│ hyperprojectN│     │Thematic map N-1│    │  hypernodeN  │
└──────────────┘     ├──────────────┤     └──────────────┘
   │ hyperworkp │     │ Thematic mapN │
   │ hyperworkp │
```
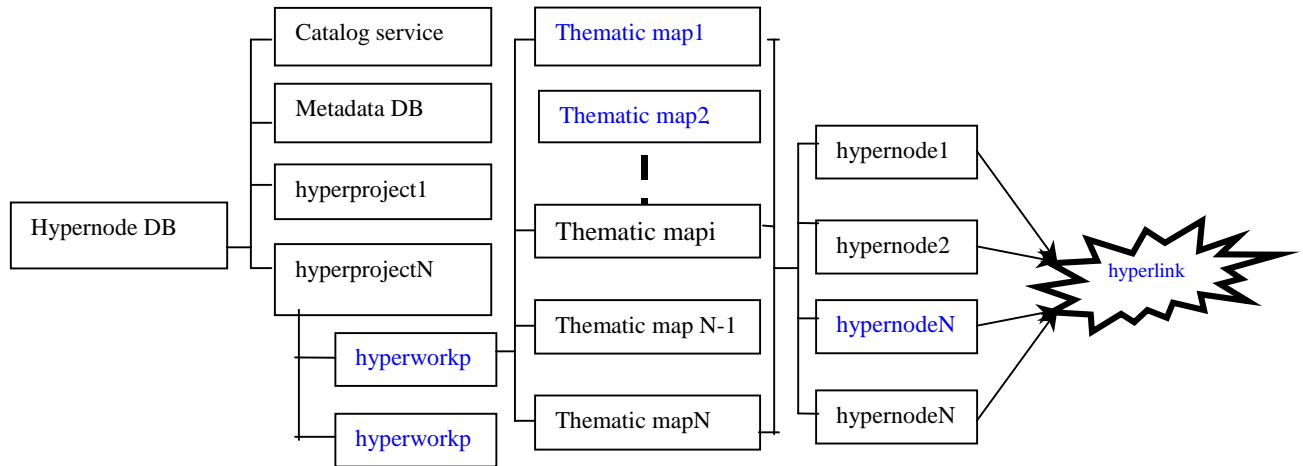
Figure2:Hypernode database construction

Hypernode database consists of hyperproject, catalog service database, metadata database. In addition, each hyperproject consists of multi- hyperworkplace. Hyperproject is used to organize distributed multi-workplace and builds up the relationship of different hyperworkplace. Different hyperworkplaces in the same hyperproject should have the same coordinate reference and map scale. It depends on user's requirement. Compared to traditional GIS, the core of distributed spatial data model is to apply the idea of project in traditional GIS and organize distributed spatial data by hypermap link. From the view of end-user, it seems to access and process distributed data in local machine, actually, user takes advantage of data distribution.

Catalog service of hypernode database is the index of the whole model. Its function is to reflect essential information of hypernode database such as IP address, name, and save other information in relation to hyperproject,hyperworkplace such as hypernode attribute, hyperlink structure.

Catalog service of hyperroot database includes IP location of all hypernode databases and their relationships.

3.3 Object-oriented Hypernode Object

The main feature of distributed spatial data model is to build up interactive relationship among different map workplace and other media by hypernode and hyperlink, and organizes hypernode database and hypernode by catalog service. Therefore, the data structure of distributed spatial data model focuses on the definition of data structure of hypernode and catalog service.

According to the idea of hypermap, map acts as graphical reference and index. Generally, a hot area of hypernode, which is defined as a specific region in the map, can link to any other media such as text document, image, map, video, etc, How to define the bound, target link of a hot area depends on the data structure of hypernode. However, vector-based map is different from image or other media at all. If hyperlink relation is built only by setting specific polygon region, it is hard to present the essential feature of vector-based map such as point, line, and polygon. In addition, the range of a polygon-based hot area is random, without a rule to obey.

Combined hypermedia with object-orient theory, it is more suitable to pack hypernode as an object. In addition, such an object will be one of attributes of point, line, polygon, workplace or project. Thus, vector-based map with other media can build up hypermap-based link relationship based on essential spatial features: point, line, polygon, etc. It is in accordance with the features of spatial data. Furthermore, it is very convenient to modify attributes of hypernode object. For example, adding, editing, or deleting a hyperlink in an object-oriented hypernode.

According to object-oriented theory, hypernode object consists of methods and attributes. Attributes focus on defining metadata information of hyperlink; methods aim to add, delete, and modify the attribute of hypernode object.

Especially, various hyperlinks of hypernode objects form a network structure. Operations, such as adding, deleting, can be done by calling methods of hypernode object. In addition, all hypermap-based links from the same hypernode object form a sub-network, so it is easy to access any linked media or spatial information by querying a hypernode object, and directly skip to the specific target media by querying result IP address represented by hyperlink.

Therefore, the design of hypernode objects not only needs to take medadata and catalog information of next sub-hyperlink into account, but also dynamic and multi-direction.

There are two kinds of meaning for dynamic. On the one hand, the frequency of clicking hyperlink can be adjusted dynamically according to clicked count. The highest count will be the default hyperlink of hypernode object. On the other hand, if SQL querying is taken towards hypernode object, all of hyperlinks in accordance with SQL condition can be queried. Although the present hypernode object does not direct to those hyperlinks, dynamic adjusting of hypernode object makes it possible to access metadata and IP address of those hyperlinks.

The meaning of Multi-direction reflects that one hypernode object can link to many other media, not a single one. Due to different background of users, they have different requirement for spatial data, so multi-hyperlink is more flexible than only single one.
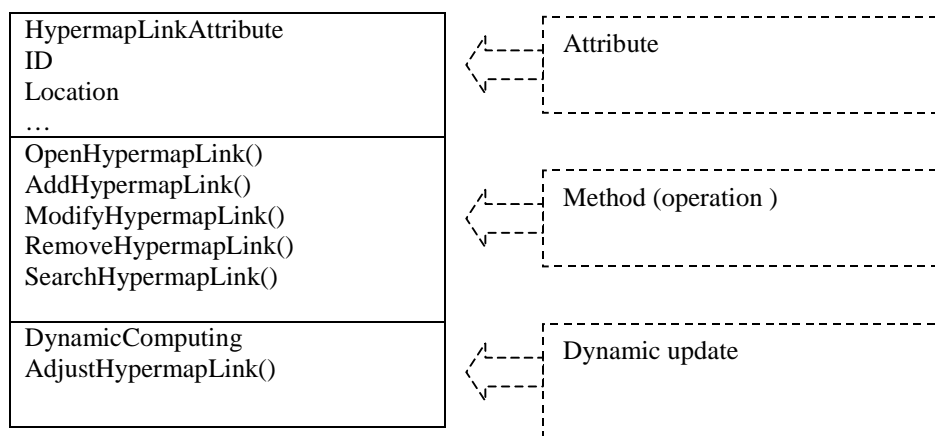

Figure3: the structure of hypernode object

Hypernode object is linked to corresponding spatial feature with a global unique identify string(GOID), so it is unique for a hypernode object on the Internet. The attributes of hypernode object aim to save linked metadata information. the method is used to update and modify the attributes.

3. 4 Attribute Structure of Object-oriented Hypernode Object

The link target of hypernode object includes text, image, video, vector-based map, VRML and other media. The paper focuses on data structure of vector-based hypernode object in relation to spatial information.

In a vector-based map, there are two ways to define hot area. One is to set specific spatial region as hot area by drawing a polygon or a rectangle; the other is to take point, line, polygon or other features as hot area. As the latter way is better to reflect the nature of spatial data, it is the emphasis of this paper.

After the hot area of a hypernode object is defined in a hyperworkplace, it is necessary to encode hypernode object. The code is globally unique (GOID). One way is to combine IP address of hypernode database with feature OID. Node name in accordance with hypernode object code should reflect the object-linked attributes.in order to link to sub-media, and the code and IP address of sub-media is also necessary. Due to different types of sub-media, it may be text, image, vector-based map or other types, and sub-media is in accordance with corresponding data format and processing method, so the features of sub-media has to be reflected in the data structure of hypernode object. For example, the format of an image may be gif or bmp. If the link target is vector-based map, it is more complicated. Now, there is various GIS software, so data format has to be pre-defined. When the hyperlink is clicked, the corresponding data processing component is called, and the result is sent to end-user.

Field Name	type	description
VNodeOID	long int	global OID
VNodeName	char	node name
VWorkplaceOID	long int	workplace OID
VObjectOID	long int	Object OID
VLinkRegion	polygon object	Polygon(x[]` y[])
VLinkScale	int	display scale
VfLinkDB	char	sub-DB OID
VfLinkWSOID	long int	sub-workplace OID

Table1: Part of Attributes of Vector-based hypernode

In tab.1, if polygon region is set as hot area, then VLinkRegion is used, or else it is null; if OID or featureID is taken as hot area, then VObjectOID is used.

By clicking hot area or zooming to defined scale, the vector-based map can be loaded by hyperlink.

It is simpler to present the hypermap-based link of an image. General way is to define hot area by a rectangle or polygon. In this way, the attributes of an image hypernode are as the same as vector-based maps. It is necessary to define linked media' type, IP address and data format. At present, most of browsers support general image data formats such as GIF or JPG. However, browser does not support vector-based map, so it deeply depends on specific DGIS software in order to display and process vector-based spatial data.

Hypernode data structure and hypermap-based link construct relationship among different media, however, it can not meet the requirement of dynamic organizing and updating distributed spatial data without catalog service. In addition, catalog service reflects essential information of all hyperprojects, hyperworkplaces and hypernode databases.

All essential information of hypernode databases, which act as a node on the Internet, have to be added into catalog service database, including their interactive relationship. If the index of a hypernode database is deleted or added from the catalog service database, then corresponding relationship in the model has to be modified dynamically.

If users enter into sub-hyperworkplace in a vector-based map, by clicking or zooming specific region or feature, model will capture the hypernode object and its attribute such as data format, type, IP address and so on. In the end, the data will be loaded and displayed in the client-side according to that meta information.

4 CONCLUSION

Distributed spatial data model focuses on distributed computing of spatial information on the Internet. Catalog service of hypernode database, which is based on Hypermap, is the core of model. It provides effective index and hierarchical structure. However, distributed spatial data model is only the basis of DGIS, it is necessary to construct multi-browser/multi-server-based architecture in order to access and process distributed, heterogeneous spatial data. The idea of model is applied in our own DGIS software---GeoSurfV3.0. The web site is http://www.jigis.cuhk.edu.hk/wwwgis

ACKNOWLEDGMENTS

REFERENCE

Caporal, J., 1995, HyperGeo: a geographical hypermedia system: Proc. Joint European Conference and Exhibition on Geographical Information, Den Haag, v.1.p.90-95

Huffmann, N.H., 1993, Hyperchina: adventures in hypermapping: Proc. 16th Intern. Cartographic Conference Intern. Cartographic Association, Cologne, v.1, p.26-45

IconAuthor, 1995, Reference manual for Windows and OS/2, version 7.0: AimTech Corporation, Nashua, variously paged.

Jiang, B., Kainz, W., and Ormeling, F., 1995,Hypermap techniques in fuzzy data exploration: Proc. 17Th International Cartographic Conference, Intern Cartographic Association, Barcelona, v.2, p.1923-1927

Kraak, M.J., and Ormeling, F. J., 1996, cartographic, visualization of spatial data: Addition-Wesley Longman, London, 222 p.

Nielsen, J., 1995, Multimedia and hypertext, the Internet and beyond: Academic Press, London, 480 p.

OGC. 1997, The OGC Technical Committee Technology Development process, Open GIS Consortium, Way-land, Massachusetts.