



# Using 3D interfaces to facilitate the spatial knowledge retrieval: a geo-referenced knowledge repository system

Bin Zhu<sup>a,\*</sup>, Hsinchun Chen<sup>b</sup>

<sup>a</sup>Department of Information Systems, Boston University, 595 Commonwealth Ave., Rm. 649, Boston, MA 02215, USA

<sup>b</sup>Department of Management Information Systems, University of Arizona, Tucson, AZ 85721, USA

Received 24 September 2002; received in revised form 20 January 2004; accepted 26 January 2004

Available online 19 March 2004

## Abstract

Retrieving knowledge from a knowledge repository includes both the process of finding information of interest and the process of converting incoming information to a person's own knowledge. This paper explores the application of 3D interfaces in supporting the retrieval of spatial knowledge by presenting the development and the evaluation of a geo-referenced knowledge repository system. As computer screen is crowded with high volume of information available, 3D interface becomes a promising candidate to better use the screen space. A 3D interface is also more similar to the 3D terrain surface it represents than its 2D counterpart. However, almost all previous empirical studies did not find any supportive evidence for the application of 3D interface. Realizing that those studies required users to observe the 3D object from a given perspective by providing one static interface, we developed 3D interfaces with interactive animation, which allows users to control how a visual object should be displayed. The empirical study demonstrated that this is a promising approach to facilitate the spatial knowledge retrieval. © 2004 Elsevier B.V. All rights reserved.

*Keywords:* Knowledge retrieval; Information visualization; Geographical information systems

## 1. Introduction

While knowledge management enables organizations to capture, organize, and access their intellectual assets effectively [14], its codification strategy [18] makes more and more knowledge repositories available. Knowledge is captured, indexed, and stored into repositories. Such knowledge repositories play an essential role in the sharing, the creation, and the application of knowledge [1,41,46]. It usually takes

two processes for users to retrieve knowledge from repositories: finding and internalizing related information. The interface design of a knowledge repository thus may have significant impact on its communication efficiency with its users [22]. A good design can facilitate the retrieval of knowledge stored and a not so good design may hamper it. While the interface requirement varies with knowledge domain and the media type, supporting cross-media knowledge retrieval is even more challenging when a knowledge repository stores knowledge into more than one media type [35].

Presenting the development and the evaluation of a geo-referenced knowledge repository system, this pa-

\* Corresponding author. Tel.: +1-617-353-4553; fax: +1-617-353-5003.

E-mail address: [bzhu@bu.edu](mailto:bzhu@bu.edu) (B. Zhu).

per explores the application of 3D (three-dimensional) interface in supporting the retrieval of spatial knowledge. While geographical information has become an important resource to support decision-making [10,32,38], the diversity of media in which information is stored also makes geo-referenced knowledge repository an interesting candidate to study the cross-media knowledge retrieval. This paper presents a system that combines realistic and abstract graphics to deliver spatial knowledge. Such combination has been shown valuable for linking “where” and “what”, the two most important concepts in the Geographical Information Systems (GIS) research [32]. The advent of multimedia leads most GIS systems [26,40,48] to use one window for the display of information in one media type. Usually there is a window displaying a map for the “where” type of information. The map can be color-coded to represent the distribution of a geographical attribute such as surface elevation, temperature, or population. More attributes can be presented on the same map by adding icons whose shape, size, or color may represent certain attributes. Except place names, attributes displayed on such a map are either numerical or taxonomic. While the window with a map indicates “where,” other windows present “what” in different media types. A window of “what” might be a statistic diagram, an image, or a text document. Users jump among windows to link “where” and “what” or to identify patterns and trends from “what.” As the volume of information increases, the windows on a computer screen can become unmanageable. The ability to pack more information into one window makes 3D interface a promising candidate for addressing this “small screen” problem [37]. A 3D interface becomes even more attractive in the context of spatial knowledge retrieval due to its similarity to the terrain surface that itself is three-dimensional. Three-dimensional figures have been used for data exploration and hypothesis generation by earth and atmospheric scientist for years. And previous GIS study also found that the 3D depiction of a terrain surface enhanced the understanding of 3D structure of the real world in pilot training [30].

Despite its advantage in visualizing geo-referenced information, most previous 3D–2D comparison studies have not yet provided supportive evidence for the use of a 3D interface [24,34,43]. However, those studies used only static interface, forcing human

subjects to observe a 3D object from a given perspective, inevitably giving rise to the “hidden object” problem. This paper, on the other hand, proposes and evaluates 3D interfaces that enable users to choose his/her own view angle to observe 3D objects through interactive animation.

## 2. Related work and research issues

### 2.1. Knowledge repository

According to Ref. [27], “A knowledge repository is an online, computer-based storehouse of expertise, knowledge, experience, and documentation about a particular domain of expertise. In creating a knowledge repository, knowledge is collected, summarized, and integrated across sources.” Based on source, a knowledge repository is either internal or external. A knowledge repository may also be structured or unstructured.

However, the usefulness of a knowledge repository may largely depend on the extent to which its interfaces help its users find and internalize related information. While most knowledge repositories have stored knowledge in textual documents, multimedia knowledge repositories pose more challenge to the retrieval of knowledge.

### 2.2. Information retrieval vs. knowledge retrieval

Among several perspectives to distinguish knowledge from information [2,45], this paper adopts the one of [2], who believe that knowledge and information to be interchangeable. Information becomes an individual’s knowledge once it is processed in the mind of that person. And knowledge becomes information when it is described and represented in artifacts. Therefore, although knowledge is coded and stored into knowledge repositories, but a user retrieves only information from a repository system. The information retrieved becomes the user’s own knowledge after he/she processes it in his/her own mind. Knowledge retrieval thus refers to both finding and processing related information to generate an individual’s own knowledge, similar to the internalization process, one of the four knowledge generation processes described in Ref. [33].

Applying visualization technologies can facilitate such internalization process. Because human eyes can process visual cues in a parallel manner, a visual interface can help users perceive patterns that might be invisible if information is presented in numbers and tables. In addition, visualization makes solutions perceivable [48], reducing the cognitive load of mental reasoning and mental image construction, necessary for internal knowledge generation.

### 2.3. Existing systems

As indicated in Ref. [11], graphical interfaces usually belong to one of two categories: the realistic or the abstract. The spatial dimensions in a realistic graphic correspond with those of the object depicted, while spatial dimensions in an abstract graphic represent non-spatial data. For instance, a geographical elevation map or a diagram of molecule is a realistic graphic, whereas various statistic diagrams such as pie chart and bar chart are examples of abstract graphic. A GIS system usually utilizes both types to link “where” and “what.” While research in cartography studies various techniques of presenting a geographical map, a variety of GIS systems focus on how to help users specify their information need. For instance, the GKRS system described in Ref. [9] and Ref. [50] supports concept-based cross media searching. The Alexandra Digital Library (ADL) project develops a system that enables users to specify their queries over a map. All those systems use multiple windows to display information in different media types and users jump among windows to obtain information. Such processes may become unmanageable to users as the volume of information increases.

Using a 3D interface is one way to pack more information into one window. Advanced hardware technology and specialized 3D software fortunately have made it faster and faster to accomplish 3D transformations, hidden-surface removal, and surface models. Researchers and software developers have invented several 3D designs, including the Cone Tree [19,36], Hyperbolic Tree [25], the WebBook [5], Information Cube [13], and information landscape [6]. In addition, a 3D interfaces is closer to the 3D terrain surface than its 2D counterpart. But as more and more 3D prototype systems have been developed to visualize large-scale information, little work has

been done to incorporate this technique into the GIS research.

### 2.4. Spatial knowledge

The goal of a GIS is to deliver spatial knowledge. MacEachren [28] cited the work of Golledge and Stimson [17] and listed three types of spatial knowledge that users may acquire from a geographical information visualization system: declarative knowledge, procedural knowledge, and configurational knowledge. Declarative knowledge, which calls for the least cognitive development, denotes to the knowledge about places and their attributes (i.e., place name and location). One example of such knowledge is as “the population of Boston is five hundred eighty nine thousand.” This type of spatial knowledge does not require contextual information but most of the time does need both realistic and abstract graphics. Procedural knowledge, considered to be at a higher cognitive level, is characterized by the awareness of how to get from one place to another and is also called routing knowledge. One example of routing knowledge is “what is the shortest path to drive from Boston to New York?” A user usually obtains routing knowledge through a realistic map depicting land surface. At the highest level of cognitive processing is configurational knowledge, which refers to the spatial relationships among places and knowledge of geographical patterns. For instance, “Arizona is in USA” is an example of configurational knowledge. Graphical interfaces provide both local and contextual information to deliver configurational knowledge, which, again, may take both realistic and abstract interfaces. How information is organized by the means of realistic and abstract interface may have great impact on users’ perception. While the effectiveness of a particular interface varies with the type of spatial knowledge, there is little current research comparing the effectiveness and efficiency of 2D and 3D approaches in conveying all three types of spatial knowledge.

### 2.5. Previous 3D–2D comparison studies

Most geographical information visualization (GIV) systems involve describing terrain surfaces, which requires a realistic map to provide information about

both location and elevation. Even though a 3D terrain description is more similar to the earth surface than a 2D description, its biggest disadvantage is the problem of hidden objects. Kumler and Groop [24] found that a 2D terrain-depiction method, continuous tone shading, performed better than a 3D depiction in facilitating static map reading. Cognitive psychology studies have also obtained similar results when abstract graphics were used. For instance, Pilon and Friedman [34] indicated that search among 2D objects is more efficient than search among 3D objects because 3D display is more complex than 2D display. In addition, Swan and Allan [43] conducted an experiment to evaluate the user interface of their textual document retrieval system. They compared the usefulness of the 2D and 3D interfaces in facilitating textual information retrieval. Again, no support for 3D interface was found.

One limitation of the cited studies is that only static 3D interfaces were provided and there was no interaction between subjects and the interface. Providing a single point of view of a 3D object inevitably causes the hidden-object effect. As indicated by Moellering [30], a 3D terrain-description can increase map understanding when multiple perspectives are provided. However, Goldberg et al. [16] argued that providing several perspectives of 3D depiction might increase a user's cognitive load because he/she has to rotate the map mentally to synthesize the incoming information, which can be a difficult task. In addition, more than one type of mental rotation is involved when human subjects read a series of 3D depictions [16]. On the other hand, information visualization research indicated that interactive animation enables users to have sole control over the rotation of 3D objects displayed, which can effectively shift the cognitive processing to the perceptual process [37]. One important example that applies this theory is the using of 3D Cone Tree system [19,36] in the visualization of structures of a large collection of text documents. However, little user study has been conducted regarding the effectiveness and efficiency of a 3D interface with interactive animation.

Another limitation of previous evaluation studies is that they focused only on the delivery of declarative and configurational knowledge, using either realistic graphics [16,24] or abstract graphics [34,43]. Most studies have indicated that a static 2D interface

(realistic or abstract) was more efficient than a static 3D interface (realistic or abstract) in conveying these two types of knowledge. For the routing knowledge, Elvins et al. [12] found that using 3D landmarks enhanced people's ability to find their way. Again, few studies have compared the effectiveness of 3D and 2D interfaces in delivering routing knowledge.

In summary, there has been very little research comparing the relative effectiveness and efficiency of 3D and 2D interfaces in both realistic and abstract types. In addition, most comparisons provided a single perspective of a 3D display, introducing the possibility of a hidden-objects effect. Some other studies provided a series of multiple perspectives of a 3D object. However, since human subjects had no control on the perspective of display, the mental rotation required for this approach still may have increased the cognitive load. Furthermore, their research attention focused only on the declarative and configurational knowledge. Little comparison with routing knowledge has been done.

## 2.6. Research formulation

The emerging 3D techniques provide a chance to increase information density and thus to decrease the number of windows. In addition, 3D interfaces have shown to enhance users' understanding of land surface [31] because of their similarity to the real world.

*Research question 1: How to apply 3D interface technique to facilitate cross-media information browsing?*

The system developed provides interactive animation that enables users to choose how 3D objects should be displayed. This provides an opportunity to further explore the differences between 3D and 2D interfaces, given the fact that most previous comparison studies used static interface and found no supportive evidence for 3D interfaces.

*Research question 2: Can a 3D interface with interactive animation have at least comparable performance as its 2D counterpart?*

## 3. System development

The data used by the 3D system includes aerial photos, Digital Elevation Model (DEM) data, Geo-

graphical Name Information Systems (GNIS), and geo-referenced textual documents. The testbed covers three media types including image, number, and text. Data from south Santa Barbara County in California State was selected. The system developed delivers all three types of spatial knowledge. The declarative knowledge includes the elevation of a location and its land surface type (i.e., urban area, orchard, beach, or mountain). With the access to textual documents, a user may also acquire declarative knowledge such as the “ground water information of Santa Barbara city.” He/she can estimate the driving distance between two locations based on the elevation change over the land surface (the routing knowledge). Moreover, configurational spatial knowledge can also be obtained by comparing attributes of a location with those of other locations nearby. Whereas all information was geo-referenced, the geographical coordinates were used as glue to combine information in different media types. Section 3.1 describes the development of the system in detail followed by the description of interface functionalities in Section 3.2.

### 3.1. Technology description

OpenGL is the programming language utilized to render interfaces. Compared with other 3D software

such as Virtual Reality Modeling Language (VRML) and Java3D, OpenGL does not provide high level command to describe 3D objects and developers need to draw pixel by pixel on the screen, which in turn gives developers more power in developing realistic and interactive 3D objects.

Fig. 1 is the 3D aerial photo that visualizes the aerial photos, Digital Elevation Model (DEM) data, and Geographical Name Information Systems (GNIS) data by applying the *textual mapping* technique. *Texture mapping* is an approach in computer graphics that glues an image over an object drawn on a computer screen. The texture mapping ensures the visual correctness as the object is transformed and rendered. For instance, an image of brick wall can be put on a rectangular to depict a brick wall. The bricks will look smaller as the wall gets farther away from the viewpoint. The 3D system drew a 3D object based on the DEM data and maps several air photos as the texture on the surface of the 3D object. The mapping was based on geographical coordinates.

Fig. 2 presents the semantic map that supports users' browsing, an effective way to access knowledge stored in textual documents, especially when a user does not have a specific goal. Providing subject categories can facilitate such browsing behavior. The subject categories can be generated automatically by

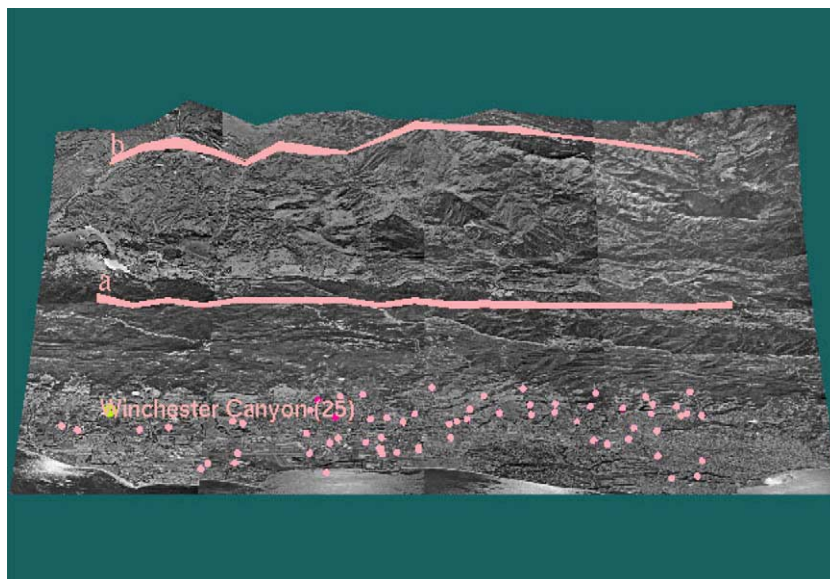


Fig. 1. A 3D aerial photo.

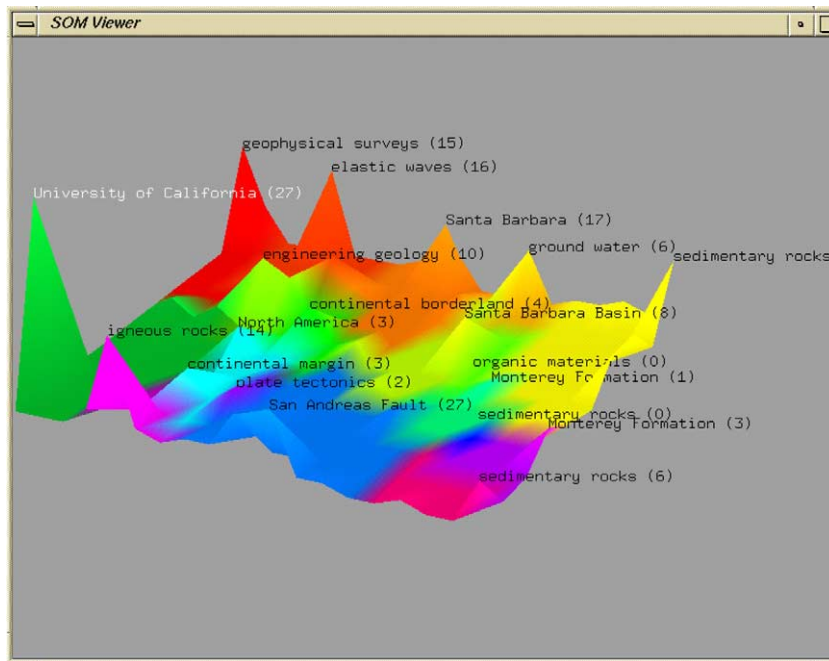


Fig. 2. A 3D semantic map.

applying various clustering algorithm when manual approaches become impossible due to the large number of documents available. Compared with other clustering algorithms, self-organizing map (SOM) appears to be a better candidate in the generation and presentation of subject categories [8]. The 3D system thus utilized indexing and categorization tools to generate the self-organizing map, called semantic map in this paper.

- *Indexing* is the process of representing a document automatically with a vector of terms [39]. The indexing tool used in this paper consists of two parts. The first operation is to use a noun-phrasing tool, Arizona Noun Phraser, to identify relevant noun phrases. The natural language processing noun-phrasing technique has been used in information retrieval to capture a rich linguistic representation of document content [3]. Allowing multi-word (or multi-phrase) matching [15], such approach has potential to improve precision over other document indexing techniques. Arizona Noun Phraser was found to have higher precision than the other noun phrasing tools in extracting

phrases [44]. The second operation is to select a subset of the phrases extracted to represent a document. The indexing tool selected phrases according to the phrase frequency (number of times a phrase occurs in a document) and document frequency (number of documents in which a phrase occurs). In addition, the place names extracted from a textual document were used to associate this document with geographical coordinates. Because the coordinates of a place can be found in the Geographical Name Information Systems (GNIS) data, a document was thus assigned coordinates based on the place names it contained.

- *Categorization* is the assignment of items to groups based on semantic association. As an information categorization and visualization tool, SOM was first proposed by Kohonen [23], who based his neural network on the associative neural properties of the brain. Table 1 provides a detailed description of the SOM algorithm. SOM is defined as a mapping from a high-dimensional input space into a two-dimensional array of output nodes, where spatial proximity represents the semantic proxim-

Table 1  
Description of SOM algorithm

- 
- Present each document in order: represent each document by a vector of  $N$  key terms and present to the system.
  - Compute distances to all nodes: compute distance  $d_j$  between an input document and each output node  $j$ .
  - Select winning node  $j^*$  and update weights to node  $j^*$  and neighbors: select winning node  $j^*$  with minimum  $d_j$ . Update weights for node  $j^*$  and its neighbors to reduce their distances to the input document.
  - Assign documents: after the network is trained through repeated presentation of all input documents, assign a document to a node that has a shortest distance from the document. Similar documents are assigned to the same or similar nodes.
  - Label regions in map: submit unit input vectors of single terms to the trained network and assign a node the name of the winning term that has the shortest distance to this node. Neighboring nodes with the same name then form a concept or topic region. The resulting map thus represents regions of important concept patterns.
- 

ity. In addition, its two-dimensional output makes SOM an ideal candidate for information visualization. Several recent studies adopted the SOM approach to textual analysis. Examples are the Distributed Script processing and Episodic memory Network (DISCERN) developed by Mikkulainen [29] as a natural language processing system, the WEBSOM system developed by Kohonen's group [21] for newsgroup classification, and the multi-layered SOM system developed by the Arizona Artificial Intelligence Group for Internet web page categorization [7]. Their work suggests a high applicability of the SOM approach to large-scale classification.

### 3.2. Interface of the 3D system

The interface of the 3D system had two parts, a 3D aerial photo (Fig. 1) and a 3D semantic map (Fig. 2). The 3D aerial photo depicts the elevation change of the landscape according to the Digital Elevation Model (DEM) data. At the same time, a user can see the picture of a place because the aerial photos have been placed on top of the land surface. In Fig. 1, the  $X$  and  $Y$  coordinates of each point on the map show its spatial location on the land surface and the  $Z$  coordinate suggests its elevation. Red dots on the 3D aerial photo indicate the locations of places named in the Geographical Name Information Systems (GNIS).

A user can not only see the elevation and the land surface type of a place, but also know the place name by putting the mouse cursor over a red dot. In addition, the 3D aerial photo also delivers the routing knowledge when a user tries to figure out what is the best path between two locations. Configurational spatial knowledge such as elevation comparison between two places can also be acquired over the 3D aerial photo.

The semantic map displayed in Fig. 2 is an example of abstract interface on which the  $X$  and  $Y$  coordinates are not associated with any place in the world. Instead, the semantic map presents subject categories of textual documents describing the geographical attributes (i.e., ground water) of places displayed on the 3D aerial photo (Fig. 1). The subject categories were generated automatically by applying the SOM. The layout of the semantic map indicates the semantic relationship among categories. Categories are semantically close if they are spatially close on the map. Furthermore, the volume of each category represents the number of textual documents associated to that category. Each category has its own color and the label of each category includes the category name and the number of documents within the category. The semantic map provides a content summary of documents about places on the 3D aerial photo. For instance, a user may know that there are six documents about "ground water" and 16 documents about "elastic wave" (declarative knowledge) in the collection by just one glance. He/she may also know that "elastic wave" has more documents than "ground water" does (configurational knowledge).

Clicking on the category of "ground water" on the semantic map will have all places related to documents of "ground water" highlighted on the 3D aerial photo. A user may also know the geographical attributes of a place by clicking on it on the 3D aerial photo, which will have all categories with documents about that place highlighted on the semantic map. In another word, the 3D aerial photo depicts locations of documents, whereas the semantic map presents content summary of the collection of documents.

The 3D system enables users to rotate and to zoom in/out both the aerial photo and the semantic map. Users thus can manipulate the object within a window by dragging the mouse or using the keyboard. For instance, dragging the mouse from left to right rotates

the object from west to east along the  $Y$ -axis and dragging the mouse from up to down leads to the rotation from up to down along the  $X$ -axis. A user can also move the object up, down, left, and right by striking designated keys on the keyboard. By interacting with the interface, a user can obtain multi-perspectives of a visual object. He/she may also avoid the hidden objects effect by rotating the visual object to choose how a visual object should be displayed.

## 4. System evaluation

### 4.1. Benchmark system

A 2D system was implemented as a benchmark system to be compared with the 3D interfaces developed. The 2D system was similar to the 3D system except that each map was two-dimensional. In order to deliver as much information as a 3D aerial photo, the 2D aerial photo (Fig. 3) needed to be supplemented by a 2D elevation map (Fig. 4). The aerial photo suggests the land surface type and the elevation map indicates elevation changes within the same area. The elevation map utilized color shades to represent the elevation of each point on the map according to a

color scheme indicator displayed by the map. Moreover, because the 2D semantic map (Fig. 5) does not have  $Z$  coordinates, the number of documents associated with each category was indicated only by its label. The 2D system also enables users to move among maps and to manipulate the view angles.

### 4.2. Hypothesis development

Most studies have found a static 2D interface to be more effective and efficient in presenting declarative and configurational knowledge than a static 3D interface. However, providing only one perspective of a 3D object inevitably causes the “hidden object” effect, which in turn affects the effectiveness of a 3D interface. Displaying multiple views of the same 3D object can solve this problem. When a series of discrete views of a visual object is provided, information about each view passes the eye and visual cortex to the sensory register. The sensory register serves as a temporal buffer to hold the information. At the same time, information about other views is also acquired and stored. Such perceptual recentering process [4] allows a user to link different views and to capture more complete features, but the mental rotation required might increase the cognitive load [16].

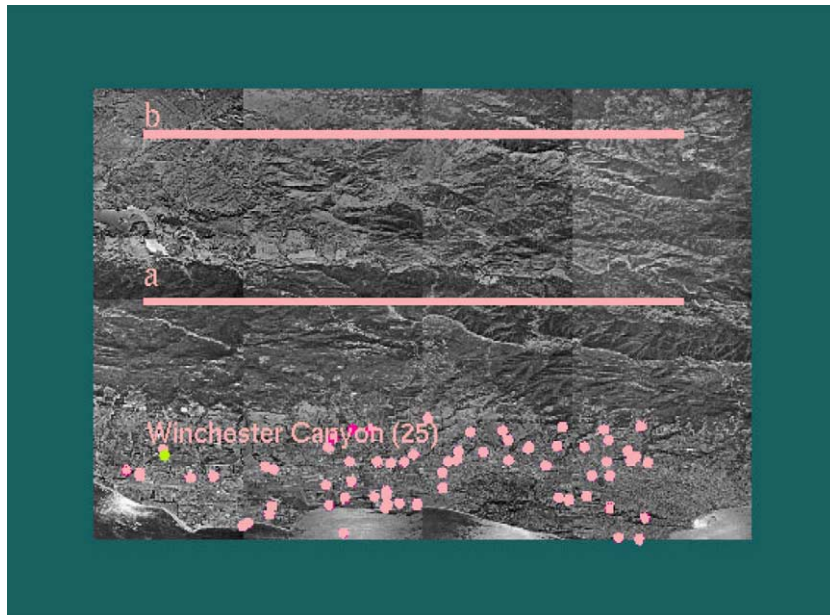


Fig. 3. A 2D aerial photo.

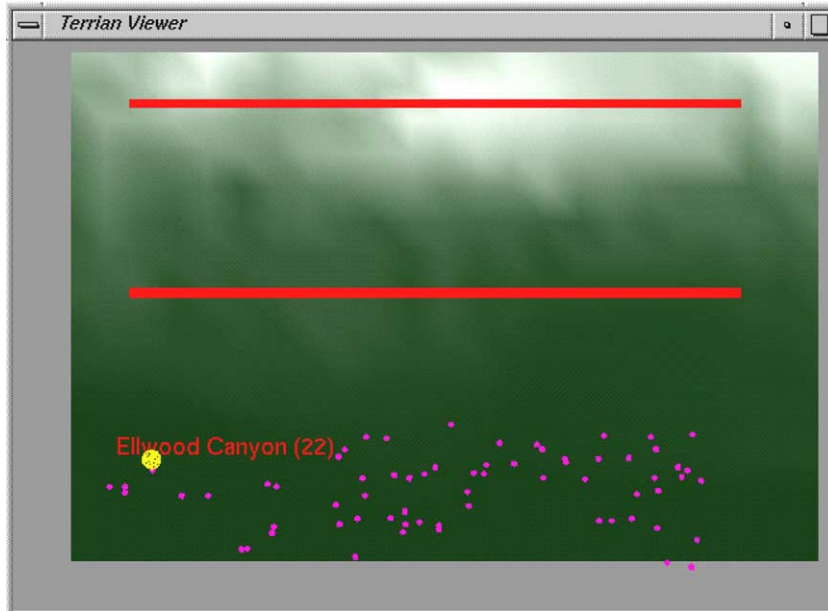


Fig. 4. A 2D elevation map.

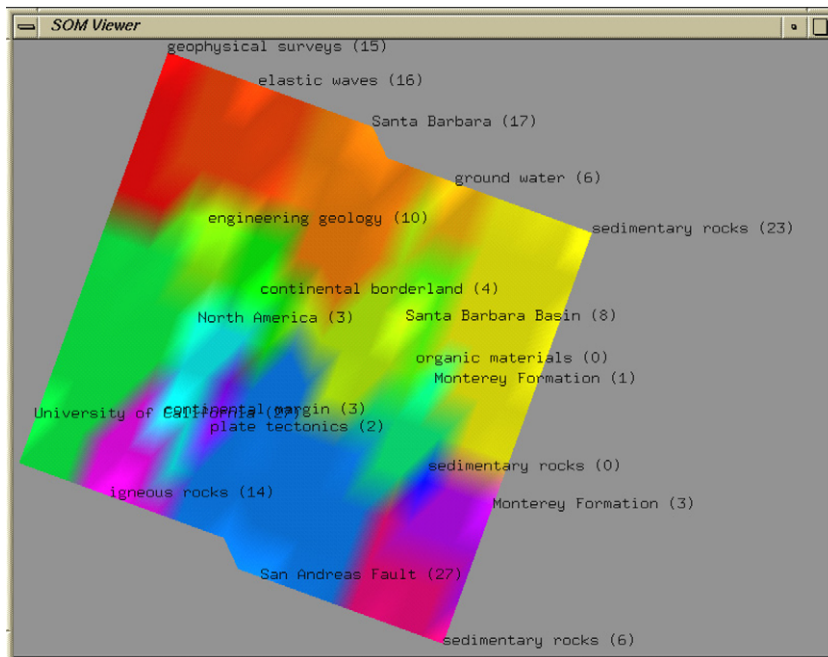


Fig. 5. A 2D semantic map.

This, however, may be relieved by using interactive animation. The interactive animation allows users to have sole control of the rotation of 3D objects. They can rotate a visual object smoothly from one perspective to another, which removes the need for mental rotation and converts a user's cognitive process to a perspective process [37]. Therefore, a 3D interface may have higher effectiveness when interactive animation is provided. Given that both our 3D and 2D systems contained both realistic and abstract interfaces, the empirical study developed its hypotheses as follows:

- H1: With interactive animation, the 3D aerial photo (*realistic graphic*) is *at least as effective and efficient* as the 2D aerial photo in conveying *declarative knowledge*.
- H2: With interactive animation, the 3D aerial photo (*realistic graphic*) is *at least as effective and efficient* as the 2D aerial photo in conveying *configurational knowledge*.
- H3: With interactive animation, the 3D aerial photo (*realistic graphic*) is *more effective and efficient* in conveying *procedural (or routing) knowledge* than the 2D aerial photo.
- H4: With interactive animation, the 3D semantic map (*abstract graphic*) is *at least as effective and efficient* as the 2D semantic map in delivering *declarative knowledge*.
- H5: With interactive animation, the 3D semantic map (*abstract graphic*) is *at least as effective and efficient* as the 2D semantic map in delivering *configurational knowledge*.
- H6: With interactive animation, the 3D system (*realistic + abstract*) is *as effective and efficient* in conveying *declarative knowledge* as the 2D system.
- H7: With interactive animation, the 3D system (*realistic + abstract*) is *as effective and efficient* in conveying *configurational knowledge* as the 2D system.

#### 4.3. Experiment design

We designed seven tasks to evaluate how well an aerial photo, a semantic map, and combined use of an aerial photo and a semantic map presented various types of spatial knowledge (Table 2). The task types

Table 2

Task assignment

	Description	Interface Type	Spatial Knowledge Type
Task 1	Please use the aerial photo to locate a place named University of California at Santa Barbara (UCSB)	realistic	declarative
Task 2	There are two roads displayed on the aerial photo. According to your visual judgment, which road will take shorter time for you to drive? <ul style="list-style-type: none"> <li>• Road a</li> <li>• Road b</li> <li>• Either road</li> </ul> Please explain your answer	realistic	routing
Task 3	Looking at the semantic map, which category or categories do not have any documents? Please find their locations on the semantic map	abstract	declarative
Task 4	Looking at the semantic map, which category, the “geophysical surveys” or the “Santa Barbara”, has more documents?	abstract	configurational
Task 5	Does the place named UCSB have documents about ground water?	combination of realistic and abstract	declarative
Task 6	How many places on the interface have documents about ground water?	combination of realistic and abstract	configurational
Task 7	Among the places that have documents about ground water, how many are in the urban area?	realistic	configurational

were selected from those used in previous studies. The identification task has been proposed to be one of the elementary task types in evaluating a graphical interface [47,49]. This type of task requires subjects to find certain visual objects based on certain attributes. We therefore chose the identification task to evaluate the delivery of declarative spatial knowledge (tasks 1, 3, and 5). While locating surface extremes is one of the map-reading tasks frequently used in previous studies [24], task 3 involves finding extreme categories that

have no document. Another widespread map reading task, estimating relative surface values, has also been used to evaluate the delivery of configurational spatial knowledge (tasks 4, 6, and 7). Such a task type requires subjects to identify not only an attribute of a given object but also to understand its contextual information. In addition, we used drive-distance estimation to measure the delivery of routing knowledge (task 2). While routing knowledge is about how to go from one place to another, estimating actual driving distance is an important factor to be considered. It has also been demonstrated in previous cognitive psychology research to be an appropriate measure for routing knowledge [20]. In addition to selecting task types based on previous studies, we also designed tasks according to the available geographical information, which includes land surface information, surface elevation information, text documents about places, and place locations and names.

Since each task required human subjects to find an answer (right or wrong), we used *task completion* as the measure of effectiveness. The system was to be considered effective if users could accomplish the tasks. To measure the efficiency, we employed *time on task* as the measure.

## 5. Experiment procedure and results

Sixty graduate and undergraduate students participated in this experiment. According to previous studies, subjects' individual map-reading skill, spatial competence and familiarity with computer may have had an impact on their performance [42]. To minimize such impact, we divided subjects randomly into two groups, with one group working on the 3D system and the other on the 2D system. The entire procedure was as follows. We first provided a written introduction to the experiment and showed the subjects how to use the system. The subjects then took as long as they wished to familiarize themselves with the system. We required each subject to accomplish the seven tasks one by one and looking ahead was not allowed. We recorded the time that a subject took to finish each task. Subjects were also encouraged to think aloud and their responses were documented.

We also took notes on how subjects interacted with the interface. The system provides several ways to

manipulate the visual objects displayed, including moving up/down, moving left/right, zooming in/out, rotating along the *X*-axis, and rotating along the *Z*-axis. One interaction score was added to a task for a subject when he/she used one type of interaction to manipulate a visual object on the interface. For instance, in order to accomplish task 7, a subject zoomed in on the aerial photo to get a close look and moved the visual object from left to right to find the desired location. We therefore assigned an interaction score of 2 to task 7 for this subject. At the end, we found that subjects using 3D interfaces had significantly higher overall interaction scores than subjects using 2D interfaces ( $p=0.067$ ), suggesting that 3D users were more inclined to interact with the interfaces provided. The hidden-objects effect posed by 3D display may have forced 3D subjects to change their view angles constantly. 2D users appeared to have more difficulties than 3D users when they had to interact with interfaces in order to acquire correct answers.

Table 3 provides a summary of results for each task, in which it could be seen that using the 3D aerial photo was significantly more effective and efficient than using 2D aerial photo in task 2, while the system

Table 3  
Experiment results

	Interface type	Spatial knowledge type	Results	
			Effectiveness	Efficiency
Task 1	aerial photo (realistic)	declarative	identical	no difference ( $p=0.817$ )
Task 2	aerial photo (realistic)	routing	3D is better ( $p=0.000$ )	3D is better ( $p=0.002$ )
Task 3	semantic map (abstract)	declarative	no difference ( $p=0.679$ )	no difference ( $p=0.402$ )
Task 4	semantic map (abstract)	configurational	no difference ( $p=0.321$ )	no difference ( $p=0.240$ )
Task 5	realistic + abstract	declarative	no difference ( $p=0.155$ )	no difference ( $p=0.218$ )
Task 6	realistic + abstract	configurational	3D is better ( $p=0.025$ )	3D is better ( $p=0.024$ )
Task 7	aerial photo (realistic)	configurational	no difference ( $p=0.389$ )	no difference ( $p=0.186$ )

combining a 3D aerial photo+3D semantic map had significantly higher effectiveness and efficiency than its 2D counterpart in task 6. The rest of this section discusses results in detail.

### 5.1. Aerial photo, declarative knowledge (task 1)

All human subjects were able to find the location of the University of California at Santa Barbara (UCSB). The 3D and 2D aerial photos had identical effectiveness. There was also no significant difference in efficiency between the 2D and 3D aerial photo ( $p=0.817$ ).

### 5.2. Aerial photo, routing knowledge (task 2)

We assigned a score of 1 when a subject gave the right answer and a score of 0 when a subject's answer was wrong. Every 3D subject had no problem with estimating actual driving distance. But even with the 2D elevation map, most 2D subjects still failed to take the elevation change into consideration. In addition, since the two roads looked identical on the interface, most subjects took a longer time to finish because they tried to figure out whether this was a trick question. The results indicated that the 3D aerial photo was significantly more effective ( $p=0.000$ ) and more efficient ( $p=0.002$ ).

### 5.3. Semantic map, declarative knowledge (task 3)

This task required subjects not only to find those categories with no document, but also to highlight them on the semantic map. It is important to know the actual location of a category on the semantic map because both its location and size have semantic meanings. The spatial proximity among categories indicates semantic proximity and the volume of a category is associated with the number of documents in it. The interface was designed to require a subject to place the mouse cursor on the color part of the category in order to highlight the category. In other words, if a subject were to point the mouse cursor on the label of a category, the category would not be highlighted. The interface was designed in this way to examine whether users could associate a text label with the category itself. The semantic map contained two categories with zero documents. The score of

each subject was calculated as the percentage of categories correctly selected. No significant difference in effectiveness ( $p=0.679$ ) and efficiency ( $p=0.402$ ) was detected. We observed that hidden objects to be the main cause of 3D human subjects' failing to recognize all categories. But the problem of 2D subjects stemmed from the label overlapping on the 2D semantic map and the difficulty in associating text labels with their color parts. It seemed that the same label layout algorithm resulted in less label overlapping on the 3D interface than on its 2D counterpart. The overlapping can be avoided by rotating the map, but 2D subjects were less effective in interacting with the interface when they had to.

### 5.4. Semantic map, configurational knowledge (task 4)

Again, for this task we recorded a score of 1 if the subject gave the right answer and a score of 0 otherwise. We required subjects not only to find the category having more documents than the other but also to highlight which two categories they needed to compare. As in task 3, label overlapping caused difficulty for 2D subjects, but there was no significant difference in effectiveness ( $p=0.321$ ) or efficiency ( $p=0.240$ ) between 3D and 2D semantic map use.

### 5.5. Aerial photo+semantic map, declarative knowledge (task 5)

This task required subjects to use both the aerial photo and the semantic map. They could start with the aerial photo, click on UCSB, and then check the semantic map to see if the category of "ground water" was highlighted. Alternatively, they could start with the semantic map, click on the category of "ground water," and then find out whether UCSB was highlighted on the aerial photo. We observed that most subjects started with the aerial photo, probably because they had located UCSB on the aerial photo in the task 1. Subjects working with 2D interface had difficulty associating text labels with their corresponding color parts if they chose to start with the semantic map. They may click on the wrong category and obtained incorrect answers. Overall, there was no significant difference in effectiveness ( $p=0.155$ ) or efficiency ( $p=0.218$ ).

### 5.6. Aerial photo+semantic map, configurational knowledge (task 6)

As in task 5, a subject needed to use both the semantic map and the aerial photo to accomplish this task. We recorded the score for each subject by calculating the percentage of correct locations the subject identified. Still, hidden objects and label overlapping appeared to be main problems to 3D and 2D users, respectively. We found the 3D system to be significantly more effective ( $p=0.025$ ) and efficient ( $p=0.024$ ) in this task. Compared with task 5, the text overlapping seemed to have more impact on this task because it requires subjects to start with the semantic map. Such problem was avoided in task 5 when most subjects started with the aerial photo. As a result, 2D users were more likely to click on the wrong category and then make a mistake in this task.

### 5.7. Aerial photo, configurational knowledge (task 7)

We recorded the results by calculating what percentage of locations selected by subjects as urban locations actually are in the urban area. In this task, a subject had to use interactions such as zooming in/out, rotation, and moving in order to obtain the correct answer. We also found that 3D users appeared to have less difficulty in manipulating the objects in the interface, probably because more user interaction with the 3D system already had been observed in previous tasks. Again, results indicated that there was no significant difference in effectiveness ( $p=0.389$ ) and efficiency ( $p=0.186$ ) between the use of 3D and of the 2D aerial photos.

## 6. Summary and conclusion

The system described in this paper provides an example of applying 3D interfaces to support cross-media knowledge retrieval. Combining realistic and abstract graphics, the system has been shown to facilitate users' spatial knowledge retrieval. In addition, the system utilizes SOM to categorize textual documents and to generate a semantic map. Integrated with 3D visualization technique, the third dimension of the SOM map indicates the number of documents

in each node. The 3D interface also has been used to incorporate information in numerical and imagery media types. Such incorporation decreases the number of windows required on a computer screen. However, all information displayed is geo-referenced, which makes it possible to integrate information based on coordinates. To apply the same approach to other domains will require creating a knowledge structure and meta-data that will allow mapping information in different media types. In addition, this approach was limited to three media types: number, text, and image. More investigation is needed when more media types, such as sound and video, are involved.

The paper also presents an experiment that compared the effectiveness and efficiency of 3D and 2D interfaces in facilitating spatial knowledge retrieval. That experiment demonstrated that with interactive animation, a 3D aerial photo was more effective and efficient than a 2D aerial photo in conveying routing knowledge, whereas a 3D aerial photo + 3D semantic map system (realistic + abstract interfaces) was more effective and efficient in presenting configurational knowledge. We list other results from this experiment as follows:

- With interactive animation, the 3D aerial photo (realistic) was at least as effective and efficient as the 2D aerial photo in conveying declarative and configurational knowledge.
- With interactive animation, the 3D semantic map (abstract) was at least as effective and efficient as the 2D semantic map in delivering declarative and configurational knowledge.
- With interactive animation, the 3D system (realistic+abstract) was as effective and efficient in conveying declarative knowledge as the 2D system.

The experiment indicated that 3D users had more interactions with the interface than did 2D users because 3D users were forced to rotate visual objects to find an appropriate display angle. 2D users did not have to deal with this issue and consequently were less experienced when they had to interact with the interface. They considered it difficult to rotate the visual objects to avoid text-label overlapping and also had trouble in associating text labels with analogous color parts on the 2D semantic map. Although a better

layout algorithm might help address the text-label overlapping problem, this study nevertheless demonstrated that it is easier to put text-labels over a 3D semantic map than over a 2D semantic map when the same layout algorithm is applied. But even with interactive animation, hidden subjects remained the main impediment for a 3D interface to deliver spatial knowledge, perhaps because manipulation of objects on the interface may still add extra cognitive load. Despite this, the results nevertheless support the conclusion that, with interactive animation, a 3D interface is a promising approach to visualizing spatial knowledge. However, the 3D interface may not have the same consistency of representation when applied to other knowledge domains. The choice of 2D or 3D interfaces still needs to take into consideration the knowledge domain, users' expectations, and users' task types.

### Acknowledgements

The authors would like to thank Dr. Daniel Zeng who kindly provided valuable help in recruiting human subjects. This research was supported by:

- Department of Defense, Advanced Research Projects Agency (DARPA), B. Schatz and H. Chen, "The Interspace Prototype: An Analysis Environment Based on Scalable Semantics."
- NSF Digital Library Initiative-2, H. Chen, "High-performance Digital Library Systems: from Information Retrieval to Knowledge Management."

### References

- [1] A. Abecker, A. Bernardi, K. Hinkelmann, O. Kühn, M. Sintek, Toward technology for organizational memories, *IEEE Intelligent Systems* 3 (3) (1998) 40–48.
- [2] M. Alavi, D.E. Leidner, Review: knowledge management and knowledge management systems: conceptual foundations and research issues, *Management Information Systems Quarterly* 25 (1) (2001) 107–136.
- [3] P.G. Anick, S. Vaithyanathan, Exploiting clustering and phrases for context-based information retrieval, 20th Annual International ACM SIGIR Conference on Research and Development, Philadelphia, PA, ACM Press, New York, NY, 1997, pp. 314–323.
- [4] R. Arnheim, *Visual Thinking*, University of California Press, Berkeley, 1969.
- [5] S.K. Card, G.G. Robertson, W. York, The WebBook and the Web Forager: an information workspace for the World Wide Web, *Conference Proceedings on Human Factors in Computing Systems (CHI '96, April, 1996)*, 1996.
- [6] M. Chalmers, R. Ingram, C. Pfranger, Adding imageability features to information displays, *Proceedings of ACM Symposium on User Interface Software and Technology (Seattle, WA, November 1996)*, ACM Press, New York, NY, 1996, pp. 33–39.
- [7] H. Chen, C. Schuffels, R. Owig, Internet categorization and search: a machine learning approach, *Journal of Visual Communication and Image Representation Science* 7 (1) (1996) 88–102.
- [8] H. Chen, A.L. Houston, R.R. Sewell, B.R. Schatz, Internet browsing and searching: user evaluation of category map and concept space techniques, *Journal of the American Society for Information Science, Special Issue on AI Techniques for Emerging Information Systems Applications* 49 (7) (1998) 582–603.
- [9] H. Chen, T.R. Smith, M.L. Larsgaard, L.L. Hill, M. Ramsey, A geographic knowledge representation system (GKRS) for multimedia geospatial retrieval and analysis, *International Journal on Digital Libraries* 1 (1998) 132–135.
- [10] A.R. Dennis, T.A. Carte, Using geographical information systems for decision making: extending cognitive fit theory to map-based presentations, *Information Systems Research* 9 (2) (1998) 194–203.
- [11] D. DiBiase, A.M. MacEachren, J.B. Krygier, C. Reeves, Animation and the role of map design in scientific visualization, *Cartography and Geographic Information Systems* 19 (4) (1992) 201–214.
- [12] T.T. Elvins, D.R. Nadeau, R. Schul, D. Kirsh, Worldlets: 3D Thumbnails for 3D Browsing, *Proceedings of ACM CHI '98 Conference (Los Angeles, CA, USA, April 1998)*, ACM Press, New York, NY, 1998, pp. 163–170.
- [13] S. Feiner, C. Beshers, Worlds within Worlds: metaphors for exploring n-dimensional virtual worlds, *Proceedings of ACM UIST '90*, ACM Press, New York, NY, 1990, pp. 76–83.
- [14] R. Garner, Knowledge management! *Computer World* 33 (1999) 50–52, GartnerGroup, Summer KM Forum (No. 32, <http://gartner6.gartnerweb.com/public/static/home/home.html>).
- [15] M.R. Girardi, B. Ibrahim, An approach to improving the effectiveness of software retrieval, in: D.J. Richardson, R.N. Taylor (Eds.), *Proceedings of the 3rd Annual Irvine Software Symposium*, University of California at Irvine, Irvine, CA, 1993, pp. 89–100.
- [16] J.H. Goldberg, A.M. MacEachren, X.P. Korval, Mental image transformations in the terrain map comparison, *Cartographic* 29 (2) (1992) 46–59.
- [17] R.G. Golledge, R.J. Stimson, *Analytical Behavioral Geography*, Croom Helm, New York, 1987.
- [18] M.T. Hansen, N. Nohria, T. Tierney, What's your strategy for managing knowledge, *Harvard Business Review* (1999 March–April) 106–116.

- [19] M.A. Hearst, C. Karadi, Cat-a-Cone: an interactive interface for specifying searches and viewing retrieval results using a large category hierarchy, *Proceedings of SIGIR '97*, ACM Press, New York, NY, 1997, pp. 246–255.
- [20] S.C. Hirtle, P.B. Heidorn, *Behavior and Environment: Psychological and Geographical Approaches*, Chapter 7, Elsevier Science Publishers, Amsterdam: North-Holland, 1993.
- [21] T. Honkela, S. Kaski, K. Lagus, T. Kohonen, Newsgroup exploration with WEBSOM method and browsing interface, Report A32, Helsinki University of Technology (1996).
- [22] P.J. Hu, P. Ma, P.Y.K. Chau, Evaluation of user interface designs for information retrieval systems: a computer-based experiment, *Decision Support Systems 27 (1999)* 125–143.
- [23] T. Kohonen, *Self-Organized Maps*, Springer-Verlag, Berlin, 1995, chap. 3.
- [24] M.P. Kumler, R.E. Groop, Continuous-tone mapping of smooth surfaces, *Cartography and Geographic Information Systems 17 (4) (1990)* 279–289.
- [25] J. Lamping, R. Rao, P. Pirolli, A focus+ context technique based on hyperbolic geometry for visualizing large hierarchies. In *Conference Proceedings on Human Factors in Computing Systems (CHI '95, Denver, CO, ACM Press/Addison-Wesley, New York, NY, May 1995)*, 401–408.
- [26] R.R. Larson, Geographical information retrieval and spatial browsing, in: L.C. Smith, M. Gluck (Eds.), *Geographic Information Systems and Libraries: Patrons, Maps, and Spatial Information*, The Publications Office, Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, 1996, pp. 81–123.
- [27] J. Liebowitz, T. Beckman, *Knowledge Organization: What Every Manager Should Know*, St. Lucie Press, Boca Raton, FL, 1998.
- [28] M. MacEachren, The role of maps in spatial knowledge acquisition, *Cartographic Journal 28 (1991)* 152–162.
- [29] R. Mikkulainen, *Subsymbolic Natural Language Processing: an Integrated Model of Scripts, Lexicon, and Memory*, The MIT Press, Cambridge, MA, 1993.
- [30] H. Moellering, The real time animation of three-dimensional maps, *American Cartographer 7 (1980)* 67–75.
- [31] M. Monmonier, Authoring graphic scripts: experiences and principles, *Cartography and Geographic Information Systems 19 (4) (1992)* 247–260.
- [32] L.D. Murphy, Geographical information systems: are they decision support systems? *Proceedings of 28th Annual Hawaiian International Conference on System Science*, IEEE Computer Society Press, Los Alamitos, CA, 1995, pp. 104–131.
- [33] I. Nonaka, A dynamic theory of organizational knowledge creation, *Organization Science 5 (1) (1994)* 14–37.
- [34] D.J. Pilon, A. Friedman, Grouping and detecting vertices in 2-D, 3-D, and quasi-3-D objects, *Canadian Journal of Experimental Psychology 52 (3) (1998)* 114–126.
- [35] S. Rathnam, M.V. Mannino, Tools for building the human-computer interface of a decision support system, *Decision Support Systems 13 (1995)* 35–59.
- [36] G.G. Robertson, J.D. Mackinlay, S.K. Card, Cone trees: animated 3D visualizations of hierarchical information, *Proceedings of ACM SIGCHI '91*, ACM Press, New York, NY, 1991, pp. 189–194.
- [37] G.G. Robertson, S.K. Card, J.D. Mackinlay, Information visualization using 3D interactive animation, *Communications of the ACM 36 (4) (1993)* 56–71.
- [38] D. Robey, S. Sahay, Transforming work through information technology: a comparative case study of geographic information systems in county government, *Information Systems Research 7 (1996)* 93–110.
- [39] G. Salton, *Automatic Text Processing*, Addison-Wesley Publishing, Reading, MA, 1989.
- [40] T.R. Smith, A digital library for geographically referenced materials, *IEEE Computer 19 (5) (1996)* 54–60.
- [41] E.W. Stein, V. Zwass, Actualizing organizational memory with information systems, *Information Systems Research 6 (2) (1995)* 85–117.
- [42] L.A. Streeter, D. Vitello, A profile of drivers' map-reading abilities, *Human Factors 28 (2) (1986)* 223–239.
- [43] R.C. Swan, J. Allan, Aspect windows, 3-D visualizations, and indirect comparisons of information retrieval systems, *Proceedings of the 21st Annual International ACM/SIGIR Conference (Melbourne, Australia, August 1998)*, ACM Press, New York, NY, 1998, pp. 173–181.
- [44] K.M. Tolle, H. Chen, Comparing noun phrasing techniques for use with medical digital library tools, *Journal of the American Society for Information Science 51 (4) (2000)* 352–370.
- [45] R. Van Der Spek, A. Spijkervet, Knowledge management: dealing intelligently with knowledge, in: A. Liebowitz, Wilcox (Eds.), *Knowledge Management and Its Elements*, CRC Press, 1997.
- [46] J.P. Walsh, G.R. Ungson, Organizational memory, *The Academy of Management Review 16 (1) (1991)* 57–91.
- [47] S. Wehrend, C. Lewis, A problem-oriented classification of visualization techniques, *Proceedings of IEEE Visualization '90*, IEEE Computer Society Press, Los Alamitos, CA, 1990, pp. 139–143.
- [48] J. Zhang, The nature of external representations in problem solving, *Cognitive Science 21 (2) (1997)* 179–217.
- [49] M.X. Zhou, S.K. Feiner, Visual task characterization for automated visual discourse synthesis, *Proceedings of ACM SIGCHI '98*, 1998, pp. 392–399.
- [50] B. Zhu, M. Ramsey, H. Chen, R.V. Hauck, T.D. Ng, B. Schatz, Support concept-based multimedia information retrieval, *Proceedings of ICIS '99, 20th Annual International Conference on Information System*, Association for Information Systems, Atlanta, GA, 1999, pp. 1–14.

**Bin Zhu** received her PhD in Management Information Systems from the University of Arizona. She is an assistant professor in the Information Systems Department at Boston University. Her current research interests include human-computer interaction, information visualization, computer-mediated communication, and knowledge management systems. She has been a lead author for papers that have appeared in *Journal of the American Society for Information Science*, *IEEE Transaction on Image Processing*, and *D-Lib Magazine*. Her research also received an IBM faculty award in 2003.

**Dr. Hsinchun Chen** is McClelland Professor of Management Information Systems at the University of Arizona, and was Andersen Consulting Professor of the Year in 1999. He received his BS degree from the National Chiao-Tung University in Taiwan in 1981 and PhD degree in Information Systems from the New York University in USA, in 1989. He authored four books and more than 150 articles covering various topics such as intelligence analysis, data/text/web mining, digital library, knowledge management, medical informatics, and Web computing. He currently serves on the editorial board of *Journal of the American Society for Information Science and Technology*, *ACM Transactions on Information Systems*, and *Decision Support Systems*. Dr. Chen is a Scientific Counselor of National Library of Medicine (NLM), USA, and has served as an advisor for major National Science Foundation (NSF), Department of Justice (DOJ), and other international research programs in digital library, digital government, medical informatics, and intelligence analysis. He is a founding director of Artificial Intelligence Lab and Hoffman E-Commerce Lab. The UA Artificial Intelligence Lab, which houses 40+ researchers, has received more than \$15M in research funding from NSF, NIH, NLM, DOJ, CIA, and other agencies over past 10 years. The Hoffman E-Commerce Lab, which has been funded mostly by major IT industry partners, features state-of-the-art e-commerce hardware and software in cutting-edge research and education environment. Dr. Chen has been a Principal Investigator (PI) of several major NSF Digital Library and Digital Government research programs. He is a conference co-chair of ACM/IEEE Joint Conference on Digital Libraries (JCDL), 2004 and has served as the conference general chair or international program committee chair for the past six International Conferences of Asian Digital Libraries (ICADL), 1998–2003. Dr. Chen was also conference co-chair of the NSF/NIJ Symposium on Intelligence and Security Informatics (ISI) 2003. His COPLINK research has been widely adopted in the law enforcement (CA, AZ, TX, MI, MA, WA, etc.) and intelligence communities (CIA, NSA, and Department of Homeland Security) in the USA, and had been featured in *New York Times*, *Newsweek*, *Los Angeles Times*, *Washington Post*, and *Boston Globe*. Dr. Chen has also received numerous industry awards in knowledge management education and research including: AT&T Foundation Award, SAP Award, and the Andersen Consulting Professor of the Year Award.