Journal of Visual Languages and Computing (1998) 9, 267–286
Article No. vl980083

Bridging the Gap: The Use of Pathfinder Networks in Visual Navigation

CHAOMEI CHEN

Department of Information Systems and Computing, Brunel University, Uxbridge UB8 3PH, U.K., email: chaomei.chen@brunel.ac.uk

Received 24 October 1997

This paper describes a generic approach to the design of a 3D virtual environment for visual navigation. The basic design principle is that users need to understand how an environment is organised and how they can utilise the knowledge to find their way through the environment. In this paper, we focus on the relationship between implicit semantic structures associated with a collection of domain-specific documents and users' cognitive needs in the context of visual navigation. We extend the notion of Pathfinder networks in order to represent implicit semantic structures in a 3D virtual environment. Users' cognitive needs in visual navigation are conceptualised based on the concept of cognitive maps.

Keywords: pathfinder network, visualisation, virtual reality, cognitive map, generalised similarity analysis, user-interface design

1. Introduction

DISORIENTATION AND COGNITIVE OVERHEAD are two of the most significant problems associated with navigating in a complex information space [1]. The problem of balancing local detail and global context is well-known to user-interface designers of a large information space [2–4]. These problems highlight the role of underlying semantic structures in matching users' cognitive needs for visual navigation.

In this paper, we focus on the relationship between implicit semantic structures associated with a collection of domain-specific documents and users' cognitive needs in the context of visual navigation. There are several crucial design decisions that one must make. First of all, what is the source of implicit semantic structures? What aspects of these structures should be represented to the user? How do we visualise implicit semantic structures to meet the needs of users in visual navigation? And finally, do users indeed find it useful for their visual navigation tasks? In this paper, we argue that an efficient visual navigation must address the interplay between users' cognitive model and the visual environment. The missing link between implicit semantic structures and users' cognitive models is likely to increase disorientation and cognitive overhead in visual navigation.

Pathfinder networks were originally developed to analyse and model patterns in proximity data [5]. We have been exploring the applications of Pathfinder networks in various areas [6–9]. In this paper, we extend the notion of Pathfinder networks in order

1045-926X/98/030267+20 $30.00/0

to represent implicit semantic structures in a 3D virtual environment. Users' cognitive needs in visual navigation can be conceptualised based on the concept of cognitive maps [10, 11]. Drawing upon user navigation strategies in the physical world, we have designed a virtual environment in which users may use similar heuristics in their visual navigation through an abstract world of documents on specific topics. The virtual environment is modelled and presented to users in Virtual Reality Modeling Language (VRML) for its wide ranging potential as well as its increasing accessibility through the Internet.

This paper is organised as follows. First, we introduce the concept of cognitive maps and its implications for visual navigation interface design. Then, we describe how implicit semantic structures are extracted and represented as extended Pathfinder networks. We also explain how these implicit semantic structures are visualised with reference to the role of various types of knowledge required in visual navigation. Finally, we discuss usability issues concerning the novel environment for visual navigation.

2. Cognitive Map

A cognitive map is the internalised analogy in the human mind to the physical layout of the environment [10, 11]. The concept of cognitive maps plays an influential role in the study of navigation strategies, such as browsing in hyperspace and wayfinding in virtual environments [12, 13–15]. It has been regarded that the acquisition of navigational knowledge proceeds through several developmental stages from the initial identification of landmarks in the environment to a fully formed mental map [13]. In this section, we use the concept of cognitive map as the basis of our user interface design, in which a virtual world is structured so as to utilise the cognitive mapping between users' understanding of the environment and the abstract information space.

2.1. Levels of Knowledge

Landmark knowledge is the basis for building our cognitive map. According to Anderson [16], as summarised in Dillon *et al.* [13], the development of visual navigation knowledge would start with highly salient visual *landmarks* in the environment such as unique and magnificent buildings or natural landscape. One associates their location in the environment with reference to these landmarks.

The acquisition of *route* knowledge is the next stage of developing a cognitive map. Route knowledge is characterised by the ability to navigate from point A to point B, using landmark knowledge acquired to make decisions about when to take turns without association with surrounding areas. If one with route knowledge wanders off the route, it would be very difficult for them to backtrack to the route on their own. They may easily get lost, even though they can start their navigation all over again based on their landmark knowledge. Route knowledge does not provide enough information about the contextual structure to enable the person to optimise his route for navigation.

The cognitive map is not fully developed until *survey* knowledge is acquired. According to Tolman [11], the physical layout of the environment must be internalised in the human mind to form a cognitive map. Dillon et al. [13] have noted that when users navigate through an abstract structure such as a deep menu tree, if they select wrong options at a deep level they tend to return to the top of the tree altogether rather than just take one step back. This strategy suggests the absence of survey knowledge about the structure of the environment and a strong reliance on landmarks to guide navigation. It is not clear if each individual develops through all stages in such a logical sequence. Nevertheless, studies suggested that intensive use of maps tends to increase survey knowledge in a relatively short period of time.

It is clear that visual navigation will rely on the cognitive map and the extent to which users can easily connect the structure of their cognitive maps with the visual representations of the underlying information space. On the one hand, the concept of cognitive map suggests that users need information about the structure of a complex, richly interconnected information space. On the other hand, it is clear that if all the information about connectivity is displayed, it would be unlikely for users to be able to navigate effectively in spaghetti-like visual representations.

In the following section, we will address criteria that might be useful in determining what information is important for visual navigation and therefore should be extracted and preserved in the visual representations, whereas redundant information may be filtered out to increase the clarity and simplicity of the visual environment.

2.2. Explicit and Implicit Structures

When users navigate through a virtual environment, the strategies and behavioural patterns of their visual navigation will be affected inevitably by a perceived structure, or metaphor, in association with the environment. There are many options in choosing the underlying structuring metaphor. For example, one can use an existing hierarchical logical structure to organise and represent the environment. The structuring metaphor may associate a variety of information with a geographical backdrop as seen in a geographical information system (GIS). On the other hand, an explicit organising structure may not always come naturally along with a given data set, or the existing structure may be simply inappropriate to specific tasks at hand. What methods are available for us to derive an appropriate structure? How can we connect such derived structures with the cognitive map in the user's mind?

Recent studies have shown that browsing through a table of contents is a preferred method over more analytical methods such as query formulation. Chimera and Shneiderman [17] examined three generally used interface methods for browsing hierarchically organised on-line information, including stable, expandıcontract and multipane tables of contents.

The major drawback to a stable interface is that users have difficulty in perceiving the entire hierarchical organisation of the text at a glimpse. Users must perform a considerable amount of scrolling with a stable interface and can get lost in a large document. The expandıcontract and multipane interfaces are used to overcome this problem by displaying the high-level information contiguously and giving users the choice of viewing specific section and subsection levels on demand. The motivation is to provide a balance of local detail and global context [18]. Not surprisingly, Chimera and Shneiderman's experiments confirmed the superiority of dynamic visual representations over static ones. Their findings also highlighted the role of structures in guiding people in visual navigation in a large database or information space.

Information on explicit, logical structures may not be readily available. In this paper, we focus on the situation where an explicit logical structure of a large collection of documents is not available or not appropriate for visual navigation. In the following section, we will address how salient structures can be extracted and transformed into a natural structure for visual navigation. We will also emphasize our generic, automated and integrated approach to the design and realisation of an extensible and re-configurable virtual environment.

3. From Cognitive Map to Virtual Environment

In this section, we focus on the relationship between a cognitive map in the user's mind and a visualised semantic structure for visual navigation. A fully developed cognitive map accommodates three types of knowledge of the surrounding environment, namely landmark, route and survey knowledge. We will show how semantic structures can be represented via information visualisation according to similar organisation strategies. Figure 1 shows the relationship between a cognitive map and information visualisation. The relationship is characterised as explicitıinherited or implicitıderived in nature. An inherited visualisation is usually straightforward and easy to understand. For example, many types of data can be organised geographically. An implicit or derived visualisation, on the other hand, is metaphorical in nature. For example, there may be a number of different ways to visualise the same latent semantic structures. Each visualisation may simply emphasize a particular aspect of the underlying semantic characteristics. Visualising implicit structures therefore is usually harder than visualising explicit or inherited structures.

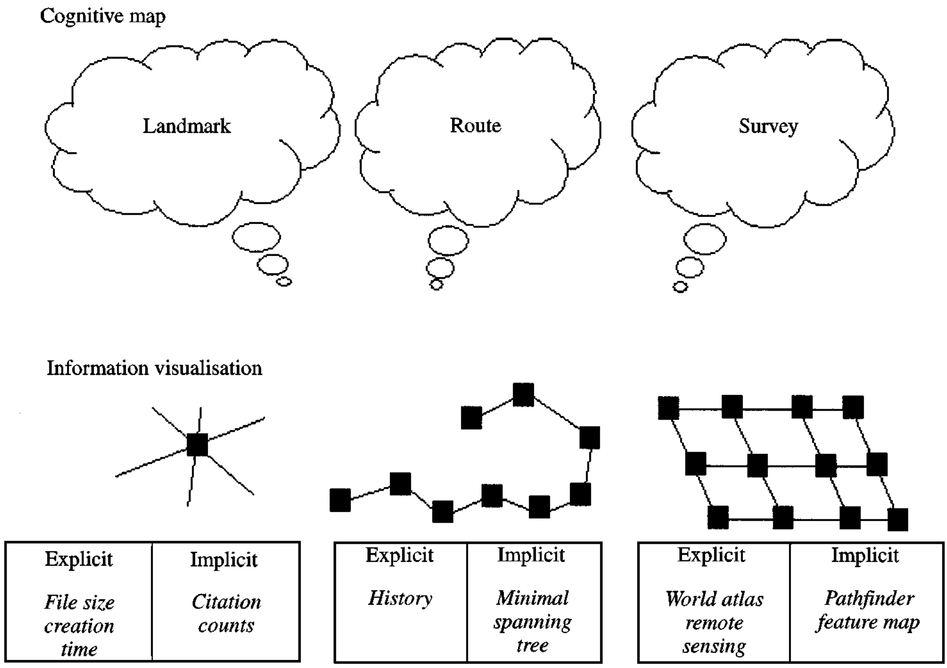

Figure 1. The relationship between a cognitive map and information visualisation

In our approach, latent semantic structures are implicit. They are derived from the proximity data according to some scaling or other transformation algorithms. For example, users' preferences can be visualised as a type of landmarks in a virtual environment such that relevant documents will be grouped in surrounding areas. In this way, landmarks provide distinct guidance for visual navigation.

3.1. Dimensionality

The issue of dimensionality is concerned with the overall aesthetics, complexity, clarity and accountability of the appearance of the visual world. Whether a high-dimensional data set is more appropriate to be embedded into a 2D or 3D space? What implications it might have if it becomes necessary to deal with a data set containing hundreds of thousands of distinct items? Is a 3D visual representation always superior to a 2D version? Can human users rely on their experience in visual navigation through the physical world to cope with the complexity of navigating in a 3D cyberspace? Does individuals' spatial ability count?

One needs to consider both technical and visual factors when choosing the dimensionality of the space in which the data are to be embedded. For visualisation purposes, the choice of dimensionality is normally whether to use a 2D or a 3D-space. On the one hand, node placement in a 3D space may turn out to converge faster than in a 2D space because a cluster of objects may have to negotiate for a minimal stress in a limited area of a plane. On the other hand, the development of some early information visualisation systems suggested that a full-fledged 3D configuration could be too complex for users to handle efficiently. Fractal dimensionality, e.g. a 2.1D or 2.5D space, has been suggested as a compromised solution. In these 'flat' 3D spaces, coordinates along the third dimension are restrained within specified-regions during the node placement process. There is notable lack of empirical studies in related areas, such as the study of effects of spatial ability of individual users on visual navigation (but see Chen and Czerwinski [8]).

In fact, we found that it might be sensible if we simply reserve the third dimension for incorporating additional navigational information subsequently. In later sections, we will show an integrated environment in which users would have a wider range of options. They would be able to utilise visual representations for both search and navigation strategies so that it becomes easier to match the visual navigation to their specific cognitive knowledge.

3.2. Spatial Layout

Documents are represented as nodes. Interconnections between documents are denoted as links. A spatial layout in a 2D or 3D space is essentially determined by interrelationships among these nodes, especially the nature and strengths of these relationships. The major concern of node placement is how to optimise the configuration of nodes such that the resulting spatial relationships truly reflect the essence of an underlying semantic structure. The position of a node is not determined on its own; one must take into account at least its most closely related nodes in a wider context. Fortunately, force-directed placement algorithms, a special class of graph drawing algorithms, have been developed to deal with these type of problems.

In a visual representation of an information structure, salient interrelationships are highlighted with associative links. Ideally, the length of a link in the visualisation should reflect the semantic closeness between the two nodes it connects together. Essential links should be distinguished from redundant links. To improve the clarity, only essential links should be preserved in the visualisation.

3.2.1. SemNet

SemNet is one of the pioneering systems using graphical representations to visualise the structure of a knowledge base [18]. SemNet was designed to enable users to have an easy access to a large knowledge base. It addressed questions such as, how should large knowledge bases be presented to users? How should users explore, manipulate and modify these knowledge bases?

The fundamental design rationale of SemNet was to enable users to draw upon the skills that they have already developed in everyday life for recognising and manipulating visual patterns in three-dimensional space. SemNet embedded the knowledge base into a three-dimensional (3D) space as a directed graph. The designers of SemNet considered a few node placement methods. For example, SemNet used Multidimensional Scaling (MDS) and simulated annealing to position the nodes so that two nodes directly connected by a link were closer to each other than those not directly connected.

Our approach differs from SemNet in the following aspects:

- SemNet derived proximity measures from logical relations encoded into a particular knowledge base, whereas our approach extracts structural patterns directly from a collection of documents.
- Node positioning in SemNet takes all the available information into account concerning inter-document connections, whereas we investigate how a concise and representative spatial layout can be derived from a subset of the information to reduce the overall complexity.

3.2.2 Intermedia

Intermedia is among the first hypertext systems using local maps to support browsing [15]. More recent examples include Harmony's local map, which generates a visualisation of a document's hyperlink relationships in Hyper-G [19].

Intermedia explored various strategies to provide a compact spatial layout to the user. One example was 'minimal links placement', in which the distance between two documents in the information space was defined as the smallest number of links connecting the two documents [15].

Utting and Yankelovich noted that updating a spatial layout may destroy all the patterns and clusters that users are familiar with, for example by adding new nodes or deleting existing links. Users, therefore, may find it difficult to predict how their actions on the network will alter the spatial layout.

They suggested a semi-spatial approach to minimise radical changes in a spatial map. In this approach, users can specify a number of key documents. The positions of key documents are fixed in a spatial layout, while the positions of other documents are fully configurable. This approach has an appealing self-organising effect in that relevant documents tend to be clustered around key documents specified by users. Utting and Yankelovich pointed out that a user who is unfamiliar with the content of the space may still find it difficult to find a good place to begin browsing. Even a key document may not always provide a good starting point for browsing. They suggested an integration of search and browsing. One can use queries such as: 'Find 10 documents with the highest number of outgoing links' and 'Find all the documents created yesterday' to locate specific starting points. The following section describes the self-organising model used in this study.

3.3. GSA

Generalised similarity analysis (GSA) is a unifying framework for extracting structural patterns from a hypermedia information space [7]. A number of intrinsic inter-relationships in hypertext, such as hypertext linkage, content similarity and browsing patterns (see Chen [20]), are consistently incorporated into the generic framework. We have been exploring applications of GSA in a number of areas, including visualising information on the WWW [7], digital libraries [6] and characterising relationships between WWW connectivity and research ratings [9]. Ongoing projects using GSA are in the areas of information filtering, co-citation analysis and latent semantic analysis.

The architecture of the GSA framework consists of a number of computational models. Each of these computational models generates a virtual link structure based on a distinct characteristic. A virtual link structure can be generated based on an integration of some or all the component models. One can incorporate additional inter-document relationships into the framework, for example, based on co-citation counts between documents. The similarity between two documents can be measured psychologically or statistically. In hypermedia systems, some fundamental relationships are hypertext linkage, content similarity and browsing patterns.

In GSA, the spatial layout is determined by a spatial configuration, which maps similarities from the domain $\mathrm{N} \times \mathrm{N}$ to $\Re$ or $\Re^3$ as with multidimensional scaling (MDS). The difference is that a Pathfinder network is built on salient relationships determined by the triangular inequality condition, which is imposed to eliminate invalid links, whereas an MDS solution is derived from $\mathrm{N} \times \mathrm{N}$ measures of pair-wise relationships. For example, given any two objects, A and B, the triangular inequality condition simply requires that a valid link between the two objects must satisfy the inequality condition: $\mathit{weight}(A, B) \leq \mathit{weight}(A, C) + \mathit{weight}(C, B)$, where $\mathit{weight}(X, Y)$ can be conceptualised as a kind of semantic distance between objects X and Y, and C is an arbitrary object other than A and B. If this condition is not satisfied, it is regarded that this particular link between A and B represents redundant or inaccurate information about the structure, therefore this link between A and B must be removed.

A Pathfinder network can be displayed as a graph using a suitable graph-drawing algorithm. Force-directed graph-drawing algorithms belong to a class of algorithms for drawing undirected graphs (see e.g. Fruchterman and Reingold [21]). A force-directed graph-drawing algorithm typically searches for positions of nodes in a 2D or 3D space such that some global criteria can be optimised in a self-organised process. Because the idea is simple and intuitive, force-directed graph drawing becomes increasingly popular in information visualisation [21, 22].

One particularly popular algorithm is called the spring-embedder model (Kamada and Kawai, 1989). In this model, nodes are connected by springs. These nodes are pushing as well as pulling each other until the whole system settles. The satisfactory state is normally reached when the spring energy in the entire system reaches the global minimal. The term self-organisation is usually used to describe the process and to emphasize the role of various controlling and aesthetic heuristics involved.

General aesthetic layout criteria include minimising the number of link crossings and overlaps, symmetrical displays and closeness of related nodes. Although the spring-embedder algorithm does not explicitly support the detection of symmetries, it turns out that in many cases the resulting layout demonstrates a significant degree of symmetrical arrangements.

We also experimented simulated annealing-based algorithms to achieve high-quality spatial configurations at the cost of a longer cooling process. Simulated annealing can deal with more complex situations. For example, one can freeze a special group of nodes as seeds, anchors or reference points depending on visualisation metaphors. At the same time all the rest of the nodes can take part in the self-organising process.

Pathfinder networks originally developed in psychology are small in terms of the number of nodes and links in these networks, whereas our extended Pathfinder networks may be very large. The efficiency of this approach can be hindered by the size of a large network. One strategy is divide-and-conquer: a large information space can be rapidly split into smaller clusters with simple classification algorithms until those computationally expensive algorithms can be effectively used.

In addition to the layout heuristics, a good navigation map should allow users to move back and forth between local details and the global context easily, search across the entire graph and follow a path consisting of a number of connecting links. More advanced features may include simulation and animation through consecutive views. Some of these requirements can be readily met by Virtual Reality Modeling Language (VRML), especially using Java and VRML 2.0 [23–25].

In this paper, the virtual world is based on the latest three ACM CHI conference proceedings. A total of 169 papers were gathered from the recent three consecutive years, i.e. 1995, 1996 and 1997. Each paper has a list of keywords. For the purpose of this study, we extracted the keyword list from each paper. The similarity between two papers was calculated using the well-known Dice coefficient, which measures the extent to which the two papers overlap in terms of the keywords they used. Using the *Dice coefficient* also gave us a convenient basis for keyword search. We have been incorporating other content-based similarity generation schemes into the GSA framework, including the classic vector space models used in information retrieval $tf \times idf$ [26] and latent semantic indexing [27].

We generated a simple index and integrated search and visual navigation into the same virtual world. A similarity matrix of 169×169 was generated. The matrix is symmetrical. Then we used GSA to generate a VRML virtual world automatically with a number of parameters for the comparison purpose. For example, we used Pathfinder network scaling as well as a minimal spanning tree algorithm to reduce the number of links in the final visualisation. We also generated 3D as well as 2D solutions. The virtual environment has a total of 169 papers: 64 papers from CHI95, 51 from CHI96 and 54 from CHI97. We will discuss these virtual worlds in more detail in the following sections.

3.4. Virtual Worlds Design

We developed a virtual reality-based user interface in Virtual Reality Modelling Language (VRML). The rationale is to represent the underlying semantic structure as a constellation of documents in a virtual world such that users can explore the structure of a complex network efficiently and intuitively. Major issues and experience associated with our design rationale are discussed in subsequent sections.

In our user interface, each document is visualised as a sphere. The link between two connecting documents is visualised by a cylinder connecting the two spheres. The length of the cylinder is the semantic distance between the two documents connected by the cylinder. The diameter of the cylinder is proportional to the degree of similarity between the two documents. In this paper, spheres and cylinders in a virtual world are randomly coloured on smooth surfaces (Table 1).

We have developed a number of VRML world generators to transform these representations into VRML worlds so that one visually navigates through these documents on the Internet. There are several advantages for using a virtual reality-based user interface, especially one that has been integrated into a widely available WWW browser such as Netscape Navigator. For example, Live3D is a VRML viewer bundled with Netscape Navigator 3.0 or later; it supports several direct manipulation tasks on a virtual world, such as walk, spin, slide, look and point. When users click on a sphere in the virtual world, the document associated with this sphere will be downloaded to the user's computer, and users can access the document directly from their own browsers. Usability issues specifically associated with such user interfaces will be discussed later in this paper. We intend to produce an atlas of interconnected documents to allow users manipulate its components easily.

New ways of interaction become possible with the virtual reality-based interface, such as seamlessly walking through, which effectively overcomes the traditional focus versus context problem. Traditional information retrieval strategies are augmented with an intuitive direct manipulation metaphor. Users can move closer to the virtual world for more details as well as move further away for a global view.

3.5. Levels of Detail in VRML

Existing empirical evidence suggests that graphical overviews are significantly useful for users to browse easily and effectively (see e.g. Chen and Rada [28]). Virtual reality-based

Table 1. Visualisation model

Visualised objects	Geometry	Visual attribute	Semantics
Paper	Sphere	Radius	File size
		Colour	Year of publication
Link	Cylinder	Radius	Document–document similarity
		Length	Minimal semantic distance
Query hit	Cylinder	Height	Query-document similarity
		Colour	Distinct keyword

user interfaces provide a promising means of helping users in accomplishing these cognitive needs. For example, by explicitly representing salient relationships between two documents in a virtual link structure, users are able to see the connectivity patterns in the entire information space. Virtual link structures of different natures, be they hyperlinks, content similarity or navigation patterns, can be readily presented to users as virtual worlds for them to explore intuitively.

Users may benefit from the virtual reality-based user interface in a number of ways. For example, direct manipulation-based user interfaces are often easier to learn and to use than other types of user interfaces such as command-driven and menu-driven.

We utilised a technique provided in VRML called level of detail (LOD) in order to provide an intuitive and seamless interface for users travelling across virtual worlds in various situations. As the user approaches to an object in the virtual world, the virtual world displays increasingly more information about the object. For example, by moving close to a cluster document, the full structure within the cluster is displayed in the virtual world.

As noted by Fairchild *et al.* [18], a graphical user interface should be designed such that users will be able to identify individual elements in the large space easily, to make sense of the spatial structure in visualisations as a meaningful context, and to recognise salient relationships between elements.

On the other hand, it is not easy to measure the impact of a graphical representation of the structure of a complex network. It is difficult to determine the effects of such visualisations on the performance of users on a range of tasks. We are currently investigating the relationship between such visualisation and individual differences, especially the interaction between users' spatial ability, cognitive styles and various visualisation representations.

The design of virtual reality-based user interfaces is still at early stages. A widespread use of VRML 2.0 will add more interactivity to virtual worlds on the Internet. Based on our experience in this design, we have identified some problems that should be addressed in future work for refining the integrated and iterative design process.

3.6. Survey Knowledge

Visual navigation in our virtual environment starts with an overview from the distance. Users then approach the centre of the virtual world for further details. Users have a number of options, such as *walk*, *spin* and *point*. In this section, we start with how an overview of an underlying information structure is presented to the user who is visually navigating in our virtual environment.

Figure 2 shows an overview of the entire information structure derived from the 169 papers on Human-Computer Interaction. Each paper is visualised as a coloured sphere. The colour indicates which year's conference proceedings the paper belongs to: red (1995), green (1996) and blue (1997). The virtual world can be generated with keyword search hit bars or without them. Feathers of hit bars will be discussed in the following section.

Figure 3 is a screen shot when the user moves closer to the centre of the virtual environment. Papers were labelled with their ID numbers. One of our design motivations was to detect emerging patterns in terms of the topics addressed continuously over

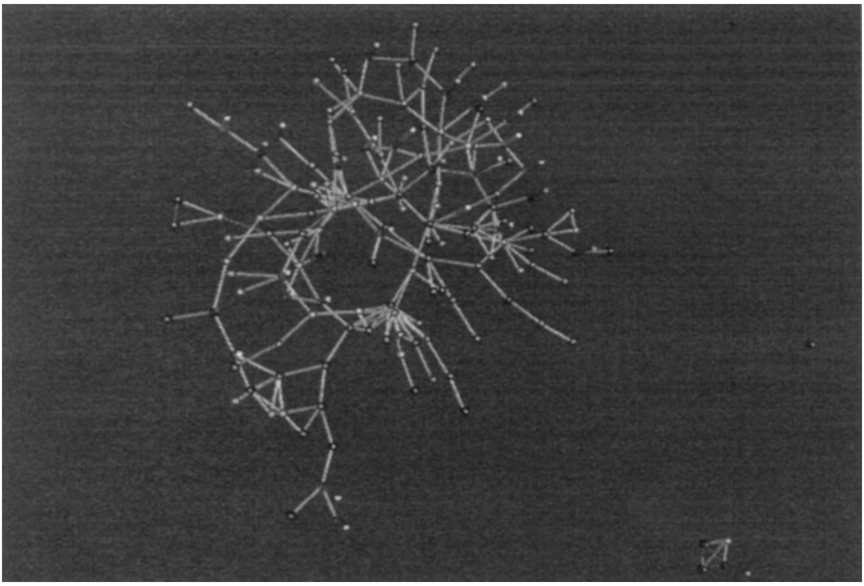

Figure 2. Overview of the entire collection of papers from the three year CHI conference proceedings

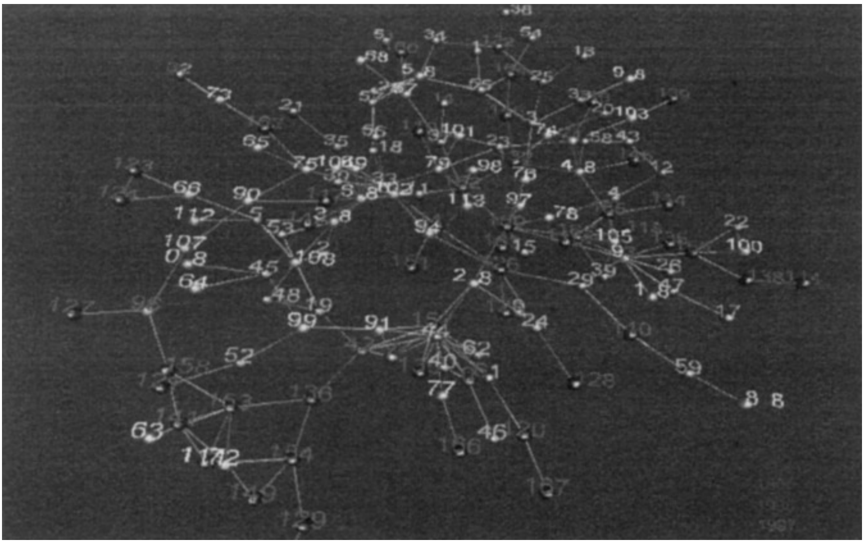

Figure 3. Overview with colour-encoded years of publication

a number of years. A large cluster of papers from a single year or a large cluster of papers equally from every year would be potentially interesting patterns.

Figure 4 shows a landscape view with additional vertical bars indicating the presence of a keyword in a query. Since the Dice coefficient-based similarity is a measure on the

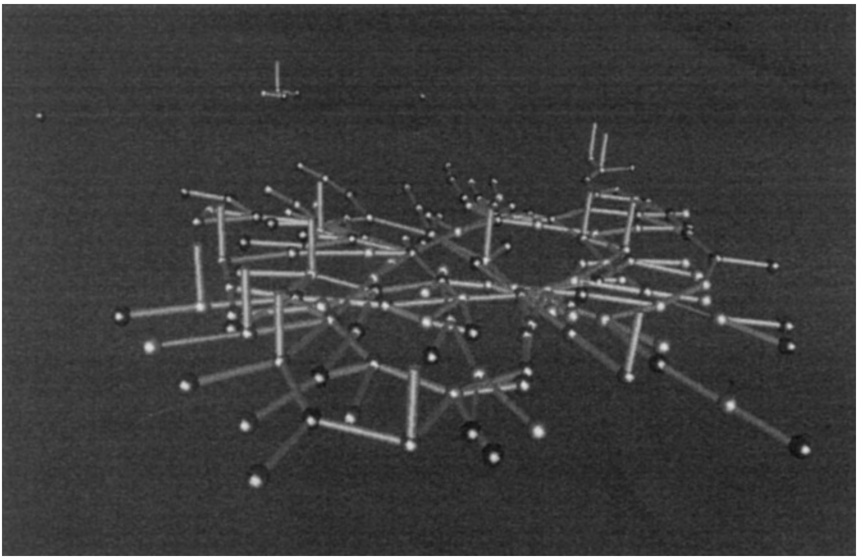

Figure 4. Landscape view with keyword search hit bars

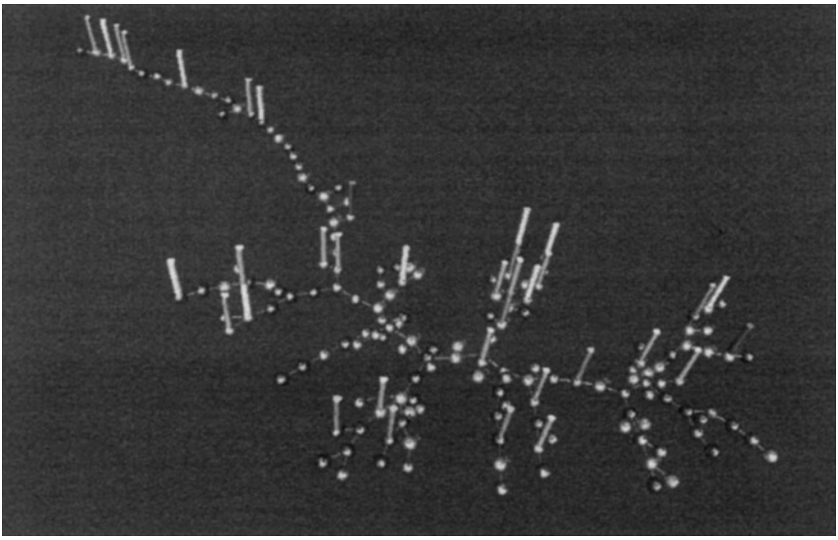

Figure 5. Overview of the papers as a minimal spanning tree (Nodes = 169, Links = 168)

overlapped keyword sets, it does not follow that papers having the particular keyword in common will be clustered in the spatial layout.

Figure 5 shows a visualisation of the CHI collection simplified as a minimal spanning tree. The minimal spanning tree contains 168 links, comparing the 208 links in our Pathfinder network. The effect of the minimal spanning tree seems that papers were grouped together tighter.

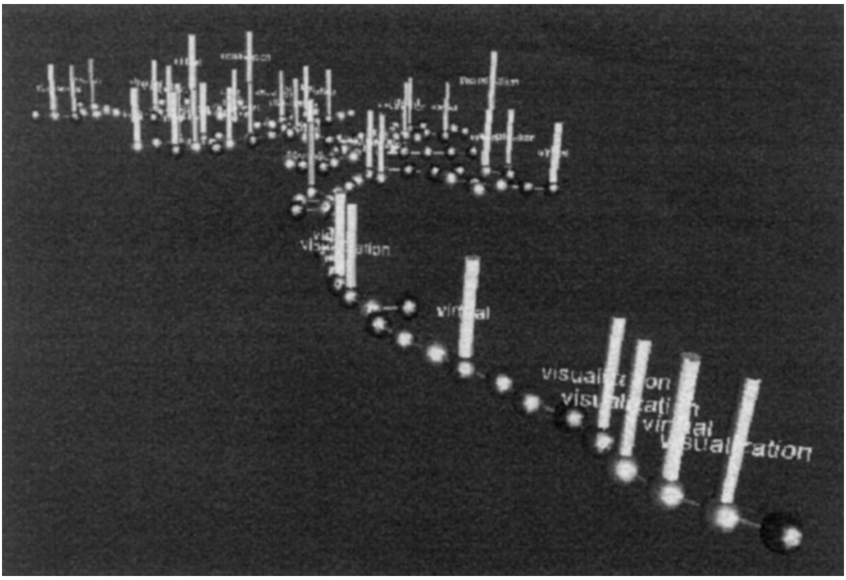

Figure 6. Navigating along the minimal spanning tree

The minimal spanning tree reduced the number of options that users may have for their visual navigation because the structure is simplified. Figure 6 shows a view from the user who was about to navigate along the minimal spanning tree.

3.7. Reference Points

There are two general types of network scaling that might be called homogenous type or heterogeneous type. In a homogenous network, all the nodes are of the same type; for example, the network contains papers and nothing else. In a heterogeneous network, one may deal with different types of nodes; for example, the network not only contains papers, but also contains sample queries (even though many studies have regarded queries as a special type of documents). There is a similar notion in association with multidimensional scaling, known as *unfolding* in psychology, in which subjects and stimulus are embedded into the same space.

In a heterogeneous network, one may mix papers and keyword-based descriptors in the same 2D or 3D spatial configuration. Then users are able to utilise not only explicitly represented inter-document relationships, but also term–document and term–term relationships in the given collection in a meaningful context.

3.8. Route Knowledge

Figure 8 shows a spatial layout of the virtual environment based on a query consisting of five distinct keywords: *design*, *heuristic*, *spatial*, *virtual* and *visualisation*. The skyline is formed by search hit bars. The higher the bars, the more keywords the paper has in common with the query.

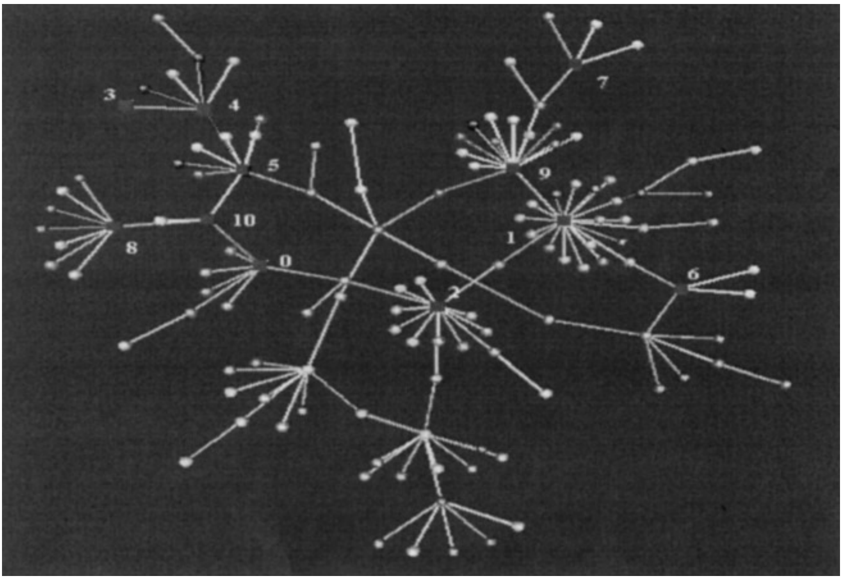

Figure 7. An example of a Pathfinder network with reference points

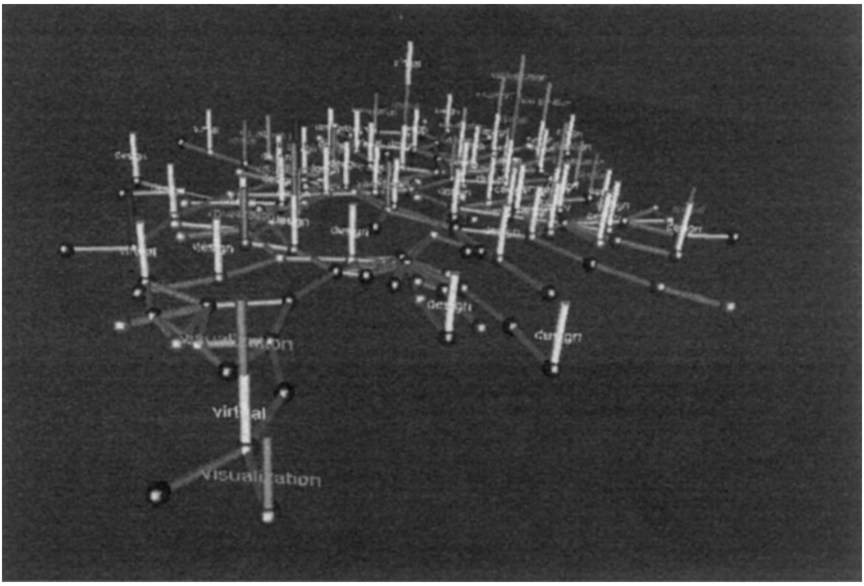

Figure 8. Landmarks and routes for visual navigation and exploration

Links preserved by the Pathfinder network are explicitly displayed. A route from one paper to another has the minimal cost, or the strongest connecting strength. The presence of a route in the virtual environment therefore suggests to the user that papers on the route between two relevant papers may be worth browsing. Papers from different years were coloured differently. Papers in CHI95 were red; CHI96 were green and CHI97 were blue. This colouring scheme was designated to detect emerging trends in research questions and application domains addressed by papers in consecutive years of conferences. For example, if we see a group of papers gathered together in blue (i.e. papers from the latest conference), it suggests that new topics emerged in the conference series. If a group of papers clustered in the network and they have every colour but the blue, then this pattern suggests that this particular area has not been covered substantially by these conferences.

The complexity of node placement relies on the number of pairwise similarity measures to be taken into account. Pathfinder networks tend to have a smaller number of links because of the triangular inequality condition. Links violating this condition were removed from the final network. Minimal spanning tree solutions may reduce the number of links even further. However, we have not yet investigated the significance of choosing among rival MST solutions. The structure of a Pathfinder network is unique if the triangular inequality condition is imposed over the entire network. In this case, the Pathfinder network is the union of all the possible MST solutions. These Pathfinder networks are subsequently used as the blueprint in our virtual reality modelling.

3.9. Landmarks

Predominant landmarks in the virtual environment correspond to the result of keyword search. Documents that contain a search keyword will be displayed with additional cylinders, or hit bars. Hit bars are coloured and labelled to enable users to distinguish between them easily. If a document contains more than one keyword, there will be a multi-part hit bar with each part in a distinct colour.

The highest hit bar naturally indicates the best result of the search. Figure 8 shows an example where the search query includes four keywords: *design*, *virtual*, *visualisation* and *spatial*. Among the 169 papers, *design* appeared 44 times as a keyword; *visualisation* 17 times, *virtual* 11 and *spatial* 9. One paper was found containing the maximum number of keywords: *design*, *spatial* and *visualisation*. One can easily tell this paper from the skyline of the virtual environment. It also appears that the neighbouring documents are likely to contain more keywords than elsewhere, for example, the area near to our present point of view. This may suggest that the structuring mechanisms work reasonably well in grouping documents on similar topics near to each other.

Live3D supports a menu option called point, which can be used to bring a selected object up front. Once the user identifies the document with the highest search bar, he/she can then use this point function to examine the local structure closely. For example, some documents nearby may not contain these particular keywords, but since they are grouped together they might have something in common, and worth exploring. Of course, the user may simply want to click on the search bar, download and view the whole paper directly.

4. Visual Navigation

Furnas and Bederson [14] distinguished two types of zooming. In contrast to geometric zooming, where objects change only their size and not their shape when magnified,

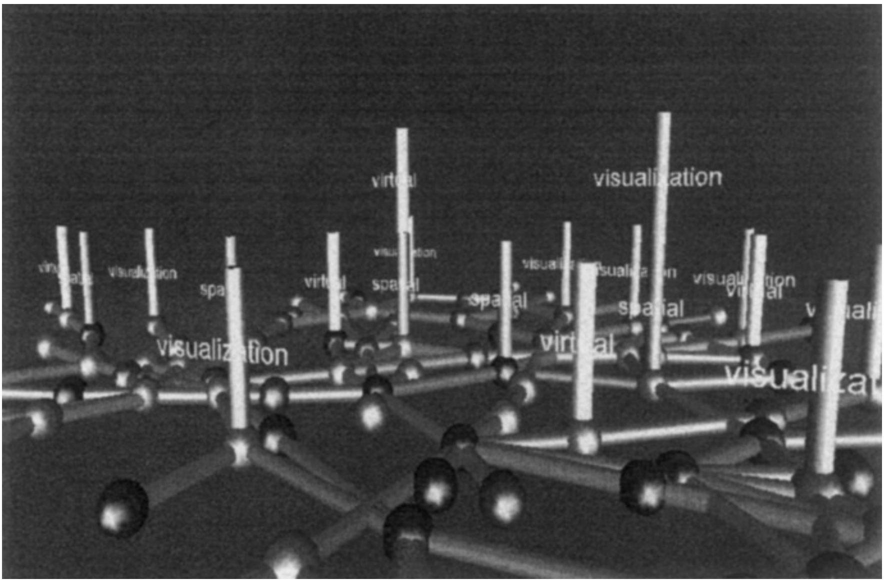

Figure 9. Approach to documents with the highest hit bars

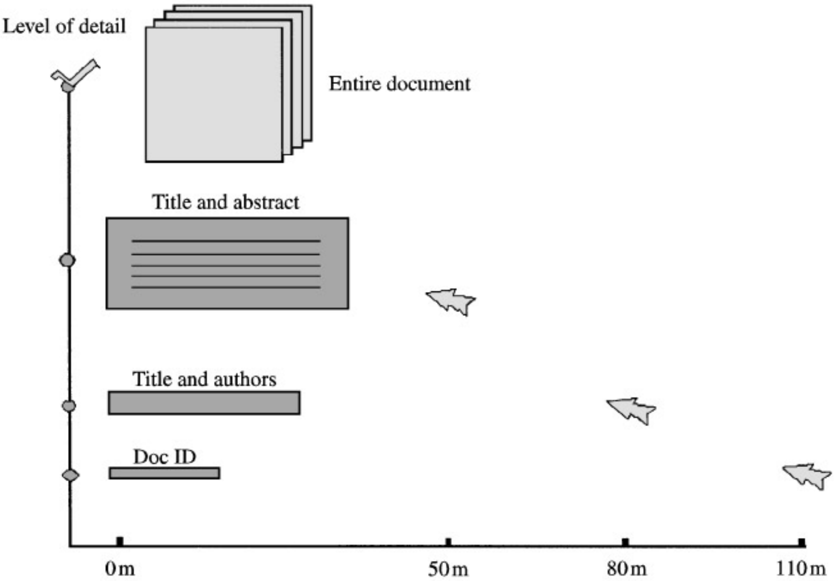

Figure 10. Space-scale diagram (a modified version based on Furnas and Bederson [14])

semantic zooming allows objects to change their appearance as the amount of real estate available to them changes. For example, an object could just appear as a point when small. As it grows, it could then in turn appear as a solid rectangle, then a labelled rectangle, then a page of text. Figure 9 is a diagram similar to space-scale diagrams

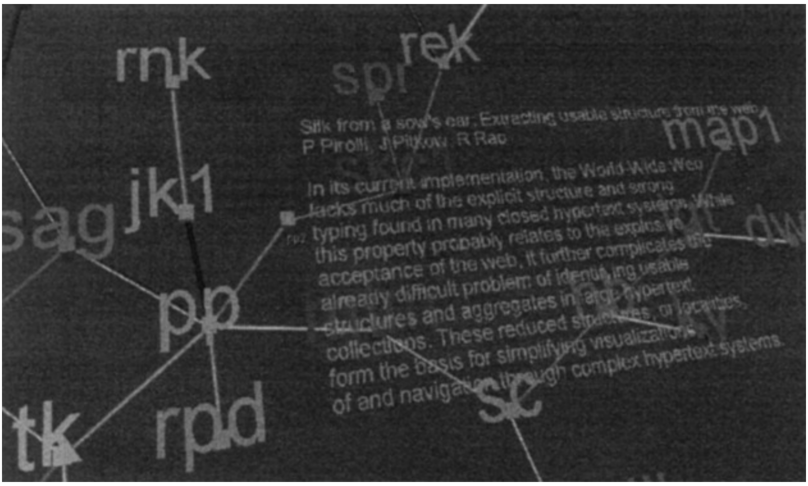

Figure 11. Transparent, multiple layered display for seamless transitions between local detail and context

described in [14]. The arrows indicate the direction of users' visual navigation. The horizontal axis is the distance between the user and the target object. As the user moves closer to the target, the target is displayed with increasingly levels of detail.

Figure 11 shows a transparent, multiple layered display to enable the user to have a smooth transition across local detail and global context. When the user moves sufficiently close to the paper in the centre of the view, the title and the abstract of the paper will be superimposed and displayed over the existing view of the network structure. The user can still see how the neighbouring papers are interconnected as well as the abstract.

4.1. Seamless Integration

Using virtual reality will reinforce users' sense of being able to follow a link continuously, instead of having to jump from page to page. This continuity, as a central element of direct manipulation, can reduce users cognitive overhead and avoid getting lost in a complex structured information space. Chimera and Shneiderman [17] animated the screen update to reduce the need for a complete reorientation.

A number of users with our VRML-enhanced structural visualisation experienced a common problem. They had to explicitly switch back and forth between inspecting the virtual world itself and using the virtual world as a gateway to the WWW. Users can point and click a document sphere or a user profile cube in both the inspection mode and the gateway mode. However, the response is different in different modes.

This problem may cause users frustration and additional cognitive overhead if they have to change the mode of interaction explicitly. One possible solution is to capture different interactive events in different situations to avoid the confusion, for example, using double click for link following and single click for viewing the geometrical details.

4.2. Spatial Ability

There are more fundamental questions. For example, to what extent users can benefit from the spatial relationships shown to them? Will they find it easier to manipulate objects in a 3D virtual world? The evaluation of the quality of structural visualisation and virtual reality is not easy. Perhaps the following areas represent some of the most important aspects of human factors in association with spatial representations.

A number of empirical studies in hypertext systems focused on spatial ability and cognitive style, in particular, interaction between cognitive styles and structural maps in hypertext [29] and a spatial world [8]. Previous studies in hypertext suggested that spatial ability is a significant factor that affects users' satisfaction and performance with hypertext systems (see Chen and Rada [28]) for a meta-analytical summary). Users with high spatial ability completed tasks more quickly than users with lower spatial ability. Empirical studies also suggested that a good structural visualisation is particularly useful to people with lower visualisation ability. In the study by Campagnoni and Ehrlich [29], users with lower visualisation ability used the top-level table of contents more frequently than users with good visualisation ability. A natural starting point is to apply similar experimental designs to the study of the relationships between users' spatial ability and the structural visualisation in virtual worlds.

We recently conducted an empirical study in an attempt to understand the role of individuals' spatial abilities in information seeking in a virtual environment developed using some techniques introduced in this paper. A detailed report of the study can be found in Chen and Czeruinski [8]. In essence, we found that users' navigation strategies and their spatial memories were strongly shaped by the visualised semantic structure. In addition, users' spatial abilities were significantly correlated with the accuracy of their spatial memories (measured based on sketches) of the underlying semantic structure.

4.3. Direct Manipulation and Incidental Learning

Darken [12] noted that survey knowledge acquired from a map tends to be orientation-specific. In contrast, prolonged exposure to navigating an environment directly is more likely to result in survey knowledge that is orientation-independent. The virtual reality-based visual navigation therefore is likely to increase the opportunities for users to get familiar with the underlying information structure.

A space should have a basic organisational principle underlying it. For example, in a library or a bookshop, books on related topics tend to be organised near each other. We use similarity as a universal metaphor to structure spatial knowledge. Most importantly, a space must have in it a number of landmarks, which can be defined as distinct, recognisable locations or regions of a larger space often associated with landmarks. It has been suggested that if a map is to be used, it should show the organisational principle of the space as well as the design elements.

5. Conclusion

We have described a generic approach to the design of a virtual world in which a variety of structural visualisations are provided to support visual navigation. There is nonetheless a lot of work to do. Evaluating the usability of such design requires a re-examination of existing theories and methodologies of human–computer interaction.

The virtual reality-based user interface described in this paper offers an intuitive and effective means of navigating in a complex information structure. The design of the virtual environment takes the advantage of the widely used Web browser and the VRML standard. Users who have access to the WWW can immediately access such interfaces from their existing computers. The techniques described here can be applied to other subject domains.

The design practice described in this paper provides some valuable experience and lessons for developing an intuitive representation of the complex structure of a subject specific domain. In addition, a number of virtual link structures can be used together to provide users insights into the underlying structure from various perspectives. More work is needed in a number of areas, such as evaluating the usability of the new user interface, optimising clustering and graph drawing algorithms, investigating the role of individual differences in the use of spatial-oriented user interfaces, especially the spatial ability and cognitive styles. The approach described in this paper provides a promising way for building intuitive graphical user interfaces for users to browse seamlessly in a digital library.

Acknowledgements

The work is currently supported by research grant GR/L61088 from the Engineering and Physical Sciences Research Council (EPSRC) in the UK under the Multimedia and Networking Applications programme. The author would like to thank anonymous reviewers for their helpful comments.

References

1. J. Conklin (1987) Hypertext: an introduction and survey. *IEEE Computer*, 17–41.
2. G. W. Furnas (1986) *Generalized fisheye views*. In: *Proceedings of CHI'86*. ACM Press, New York, pp. 16–23.
3. I. Lokuge, S. A. Gilbert & W. Richards (1996) Structuring information with mental models: a tour of Boston. In: *Proceedings of CHI'96*. ACM Press, New York. Available as http://www.acm.org/sigchi/chi96/proceedings/papers/Lokuge/sag_txt.html
4. M. Sarkar, S. S. Snibbe, O. J. Tversky & S. P. Reiss (1993) Stretching the rubber sheet: a metaphor for visualizing large structures on small screens. In: *Proceedings of ACM UIST'93*, November 1993, ACM Press, New York, pp. 81–91.
5. R. Schvaneveldt, F. Durso & D. Dearholt (1989) Network structures in proximity data. *The Psychology of Learning and Motivation* 24, 249–284.
6. C. Chen (1998) Augmenting user interfaces for digital libraries with virtual reality. In: *Proceedings of the 31st Hawaii International Conference on System Sciences*, IEEE Computer Society, New York, pp. 148–157.
7. C. Chen (1997) Structuring and visualising the WWW with generalised similarity analysis. In: *Proceedings of Hypertext97*. ACM Press, New York, pp. 177–186.
8. C. Chen & M. Czerwinski, Spatial ability and visual navigation: an empirical study. New Review of Hypermedia and Multimedia (forthcoming).
9. C. Chen, J. Newman, R. Newman & R. Rada, How did university departments interweave the web: A study of connectivity and underlying factors. *Interacting with Computer* (forthcoming).

10. P. Thorndyke & B. Hayes-Roth (1982) Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology* 14, 560–589.
11. E. C. Tolman (1948) Cognitive maps in rats and men. *Psychological Review* 55, 189–208.
12. R. P. Darken & J. L. Sibert (1996) Wayfinding strategies and behaviors in large virtual worlds. In: *Proceedings of CHI'96*. Available as http://www.acm.org/sigs/sigchi/chi96/proceedings/papers/ Darken/Rpd_txt.htm.
13. A. Dillon, C. McKnight & J. Richardson (1990) Navigating in hypertext: a critical review of the concept. In: *Human–Computer Interaction—INTERACT'90* (D. Diaper *et al.* eds). Elsevier, Amsterdam, pp. 587–592.
14. G. W. Furnas & B. B. Bederson (1995) Space-scale diagrams: understanding multiscale interfaces. In: *Proceedings of CHI'95*. ACM Press, New York. Available as http://www.acm.org/sigchi/chi95/ proceedings/papers/gwf_bdy.htm
15. K. Utting & N. Yankelovich (1989) Context and orientation in hypermedia networks. *ACM Transactions on Information Systems* 7, 58–84.
16. J. Anderson (1980) Cognitive Psychology and its Implications. W. H. Freeman, San Francisco.
17. R. Chimera & B. Shneiderman (1994) An exploratory evaluation of three interfaces for browsing large hierarchical tables of contents. ACM Transactions on Information Systems 12, 383–406.
18. K. Fairchild, S. Poltrock & G. Furnas (1988) SemNet: three-dimensional graphic representations of large knowledge bases. In *Cognitive Science and its Applications for Human–Computer Interaction* (R. Guidon, ed.). Lawrence Erlbaum Associates, London, pp. 201–233.
19. H. Maurer (ed.) (1996) *Hyperwave: The Next Generation Web Solution*. Addison-Wesley, Harlow, England.
20. C. Chen (1997) Writing with collaborative hypertext: Analysis and modelling. Journal of the American Society for Information Science 48, 1049–1066.
21. T. M. J. Fruchterman & E. M. Reingold (1991) Graph drawing by force-directed placement. *Software—Practice and Experience*, 21, 1129–1164.
22. T. Kamada & S. Kawai (1989) An algorithm for drawing general undirected graphs. Information Processing Letters 31, 7–15.
23. R. Earnshaw & J. Vince (eds) (1997) *The Internet in 3D: Information, Images and Interaction*. Academic Press, New York.
24. R. Lea, K. Matsuda & K. Miyashita (1996) *JAVA™ for 3D and VRML Models*. New Riders Publishing, Indiana.
25. G. Walker, J. Morphett, M. Fauth & P. Rea (1996) Interactive visualisation and virtual environments on the Internet. *British Telecommunications Engineering* 91–99. Available as http://virtualbusiness.labs.bt.com/ msss/IBTE_CVE/part1.htm.
26. G. Salton, A. Singhal, C. Buckley & M. Mitra (1996) Automatic text decomposition using text segments and text themes. In: *Proceedings of Hypertext'96*. ACM Press, New York, pp. 53–65.
27. S. Deerwester, S. Dumais, T. Landauer, G. Furnas & R. Harshman (1990) Indexing by latent semantic analysis. *Journal of the American Society for Information Science* 41, 391–407.
28. C. Chen & R. Rada (1996) Interacting with hypertext: a meta-analysis of experimental studies. *Human–Computer Interaction*, 11, 125–156.
29. F. Campagnoni & K. Ehrlich (1989) Information retrieval using a hypertext-based help system. ACM Transaction on Information Systems 7, 271–291.
30. D. Canter, R. Rivers & G. Storrs (1985) Characterising user navigation through complex data structures. Behavioural and Information Technology 4, 93–102.