The London Walkthrough in an Immersive Digital Library Environment

Leonidas Deligiannidis

University of Georgia
Dept. of Computer Science
Athens, GA 30602
ldeligia@cs.uga.edu

Robert J.K. Jacob

Tufts University
Dept. of Computer Science
Medford, MA 02155
jacob@cs.tufts.edu

Abstract

A new approach to browsing digital libraries is presented in this paper. This approach enables users to comprehend and digest large amounts of information easily. Our approach, which extends Shiaw's research [31], takes advantage of fully immersive Virtual Reality Environments to preserve context even when a user focuses on a single item. Currently, in web-based digital libraries, users can either get an overview of the context or focus on a single item, but not simultaneously, sacrificing the context view. Our approach eliminates this focus versus context problem where a small focus, such as a web page, competes for the limited screen space of a computer monitor with a large context, the collection of information.

Keywords: Virtual Reality, Human-Computer Interaction, Digital Library, Information Visualization

1. INTRODUCTION

We have implemented a fully immersed environment where users can navigate and interact with specific data (focus) as opposed to the complete data set (context), while not losing the context. The 3D synthetic environment is populated with 3D data artifacts extrapolated from 2D image collections found in the Perseus Digital Library [3].

The prototype is an interactive London walkthrough built from a 3D virtual world developed by Horn-Yeu Shiaw [31]. She took several historical Tallis maps and texture mapped them onto simple 3D objects for each building and plot. Users can learn about the 1765 fire that affected the modeled intersection. A sample of the Tallis maps in Perseus is shown in figures 1 and 2.

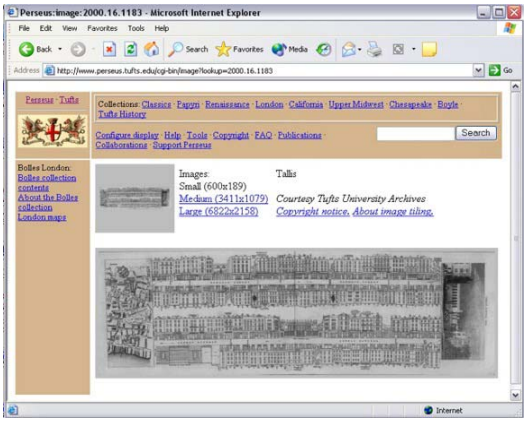

Figure 1. Sample of Tallis maps (web page)

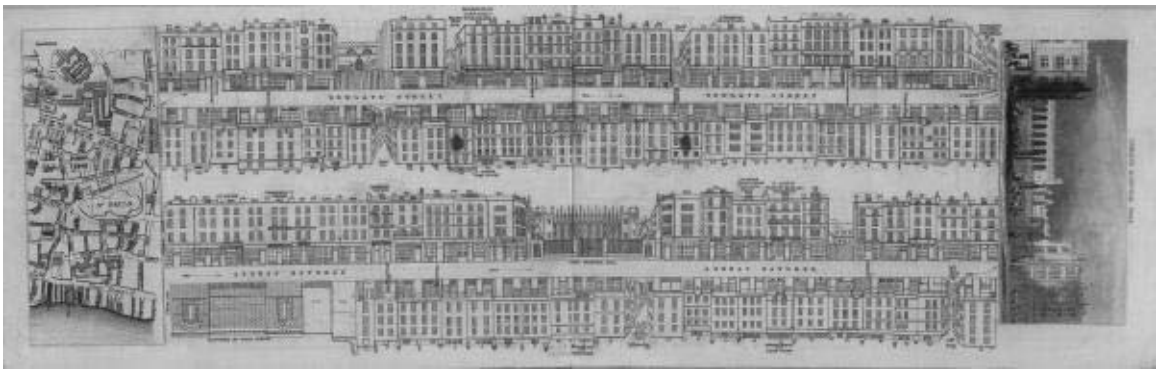

Figure 2. Sample of Tallis maps (enlarged)

The Tallis maps are created by John Tallis who was one of the most popular cartographers of the 19th century. He was renowned for his accurate designs of the maps, which are among the most popular 19th century maps available today. These maps are characterized by their topographical detail and political information.

A serious problem with the Tallis maps is the focus and context problem because there is no context neither focus to the maps. Users need to choose between focus (a map, a section of a map, a building on the map) versus context (related maps and their relationships).

It is not difficult for users who are interested in specific maps to locate the right maps and understand them if this is their task. However, these maps are scattered in Perseus without context and it is up to the users to figure out the relationships between them. Placing all related maps on a single page/screen is not practical since the maps could be related to other maps based on different criteria. On the other hand, if only a single map is presented to the users at the time, the users will loose the contextual information of the other related maps. Regardless of how the information is presented to the users, it is difficult for the users to form a mental model of the information. Users have to browse multiple maps, going back and fourth, without appreciating the geographical connection between the maps. A lot of effort and energy needs to be invested by the users to partially, if at all, understand the correlations between the maps.

Our approach, which is an extension/improvement to previous research [31], enables the users to understand the relationships between the maps with minimal effort on their part. This is achieved by presenting the information in a 3D immersive environment where users can simply walk, fly, and turn around, and interact with the VE, to observe multiple maps at the same time, or they can focus on a specific datum without losing the overall picture, the context.

2. BACKGROUND AND RELATED WORK

Due to the limited space of computer monitors, there is a constant tradeoff between 'a view of the whole data available, while pursuing detailed analysis of a part of it' [1 pp307]. This is the "focus versus context" problem and it is applicable in digital libraries. This a problem where the overview and the detail constantly compete with each other for a limited screen space. One solution to this problem is for the user interface design to afford a balance between the local and the global information; focus and context respectively [31]. According to S. K. Card et al. [1 pp307], there are three premises that can define a solution to this problem:

- 'First, the user needs both overview (context) and detail information (focus) simultaneously.
- Second, information needed in the overview may be different than that needed in detail.
- Third, these two types of information can be combined with a single (dynamic) display, much as in human vision.'

There is a need for focus and context simultaneously where focused information is different from contextual information. Focus and context must be provided in a single display. Digital libraries offer a lot of information by nature. On the web, users face the problem of viewing isolated pages (focus) of information while navigating via cyber-links to more and more context. Since a window can occupy most of the display, more and more windows (the result of following cyber-links) could lead to a cluttered display where users loose context very easily. It is apparent that a solution must be found where users can focus on a datum of interest without losing context.

Our new non-WIMP (Window, Icon, Mouse, Pointer) [15] [19] [24] interaction techniques solves this problem, by providing a larger display to the user in a Virtual Environment (VE). We provide tacit and non-command techniques in virtual reality where the users can interact with data objects in an information rich environment, naturally [2]. The VE is constructed from large data sets found in a real digital library, the Perseus library [12] [20] [21] [22] [23] [29]. The users can focus on a part of datum by navigating in the virtual environment while the context of the entire data set remains available in the surroundings.

Current user interfaces for the digital library are characterized as a process of search, browse, display, and study. However, as users start browsing and following links on the web, the contextual purpose becomes a moving target. Instead of the user to focus on a particular piece of information, the user branches out to different pages to find relevant information ending up in losing the context. At the end the user may end up with several open pages that usually clutter the display and in cases disabling the user to even focus. In that respect VR is a promising technology because it offers more virtual display real estate. User can be given the opportunity to hear pre-recorded messages/stories, touch and feel data elements or even swap views in and out to see the effect of the London fire, for example, as it is shown in the next sections of this paper. All of this however comes with a cost. Building the VE is a time consuming task. Images and maps from the digital library must be extracted and used to build the VE. This involves studying the maps, cutting the right pieces from drawing, and finally attach them in the virtual environment. Then we should also provide new interaction and navigation techniques to the users

so that the users can explore the VE and also be able to interact in some degree with it and the data that is being visualized. Our solution is primarily visual, users must be able to recognize the visualized data with little or no reading at all. Since most Head Mounted Displays (HMD) have low resolution, reading text while the user is immersed in the VE is difficult and can cause eye-strain. During the last two decades, a lot of research is focused on information environments that can afford both focus and context at the same time [13] [14] [16] [17] [25] [26] [27] [28] [32].

3. LONDON WALKTHROUGH

The 3D models used to populate the virtual scenes in the prototypes are constructed from actual collections found in the Perseus Digital Library (Perseus) at Tufts University [9] [30]. Over the years, Perseus has become a leading web-based digital library in classics and Greek archaeology. Its collection includes many images of old maps, ancient sites, mythologies, gems, coins, vases and sculptures. The development and evolution of Perseus has been the topics of many research studies [4] [5] [6] [7] [8] [11]. Although we have used Perseus as our data source, the solution proposed in this paper is applicable to digital libraries in general. Large data sets can be visualized as an information landscape (context) where the users can navigate and interact with data of interest (focus).

The London walkthrough is a 3D cityscape of the 1840's developed as a VRML model by Shiaw [31]. There are 242 building-like boxes, texture-mapped with images of the buildings cut out from the original 2D building maps [31]. The buildings are lined up to form an intersection that is bent in space to match the contour of the street map. Three types of maps are combined to create the VE: buildings, plots, and street maps. The roofs of the buildings are oriented in a way that can reflect the plot map over the street map. In addition, on the roofs we display the name of each building visible from a bird's-eye view.

In a single virtual display, we are able to provide the context (the cityscape) and the focus (a building) by navigation in space. Using gestures we can provide to the user another rendition of the London walkthrough; we show the effect of the London fire that occurred in that intersection in 1765. The users can explore the burnt scene, as compared to reading about the damage in a text passage in Perseus. We can also allow the users to interact with our 3D menu, which is analogous to hyperlinks on a web-page. The menu selection can switch views between the burnt and the not-burnt streets, show textual information in the environment, and transport the user to relevant locations in the Virtual Environment.

4. THE PROTOTYPE

We implemented the VR prototype at the VR lab at UGA using Java3D and the JWSU toolkit [18], which is built on top of Java3D. For 3D tracking we utilized a Polhemus Fastrak; tracking the head and the two hands of the user. We also used two PinchGloves to implement the gestures. A stereo Head Mounted Display (HMD) was also utilized. The 3D environment, the London walkthrough, was designed in VRML that we then loaded into Java3D using the Xj3D loader [10].

Pictures of buildings cut off from drawings are textured mapped onto 3D boxes in the VE. Coloring of the original images help us distinguish the borders of each building. The buildings are aligned in exact proportion to match the contour on the street maps. We used two versions of the London walkthrough. The first version shows the model as it used to be before the 1765 fire, and the second version shows it after the fire; the difference is in the coloring and the markers we placed in the model to indicate wind direction and sequence of events. Both versions were loaded initially but only one of them was visible at the time; the user can toggle between the two as it will be described shortly. Because both versions were identical in scale and transformation, switching from one to the other, appears that simply the appearance of the buildings change while the user stays at the same place.

Since the scale of the model is much greater that the physical laboratory, users need to travel in the virtual environment to explore different parts of the virtual environment. We implemented two travel techniques, *fly* and *drive*. The direction of the user's right hand indicated the direction of travel. The difference between the *fly* and the drive technique is that in the *drive* the user can not change the vertical axis's values (axis Y in Java3D). The user travels only on the X-Z plane.

Figure 3 illustrates the internal state machine implemented in the application. There are five states in total. IDLE is a special state and is used simply as a starting point. All gestures implemented follow the Pinch-Release paradigm; touch two fingers and then release them to perform an action, except when moving to FLY or DRIVE states. A release of the fingers in these two states moves you back to the IDLE state. For example, to fly one needs to touch his/her right hand's thumb and ring fingers. While the fingers are in contact the user is flying in the virtual environment. Upon the release of the fingers, the travel technique is deactivated. In figure 3 the releases of the fingers are not shown, to avoid cluttering the diagram, except for the DRIVE and FLY states, where we do show the finger releases.

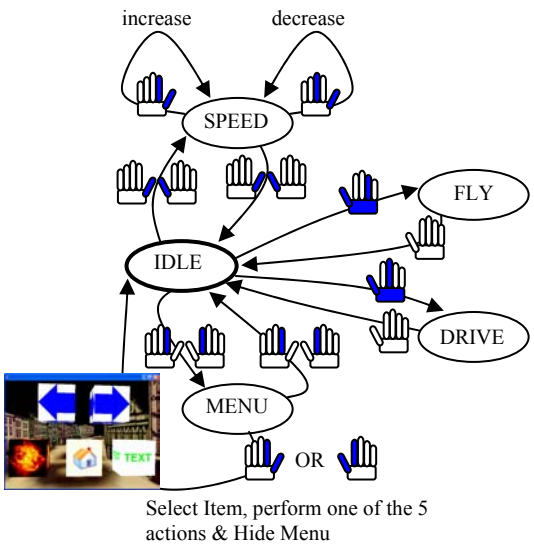

Figure 3. State Diagram

The user can also control the speed of travel as it is illustrated in the state diagram. When the MENU state is activated, a 3D menu appears in front of the user as it is shown in figure 4. When the menu is about to become visible, we store the position/orientation of the user's head, and we display the menu at that position/orientation; actually a little bit forward from the user's face. The user is still free to move around the menu and see the rest of the virtual environment; the menu does not occlude the scene. The user can now select one of the five menu items by 'pinching' one of the 3D menu items with his/her left or right hand. We made the 3D menu pop up in front of the user, instead of having it static at the bottom of the screen, for two reason: first because the user can select a menu item by simple pinching it, and second to provide more screen real estate.

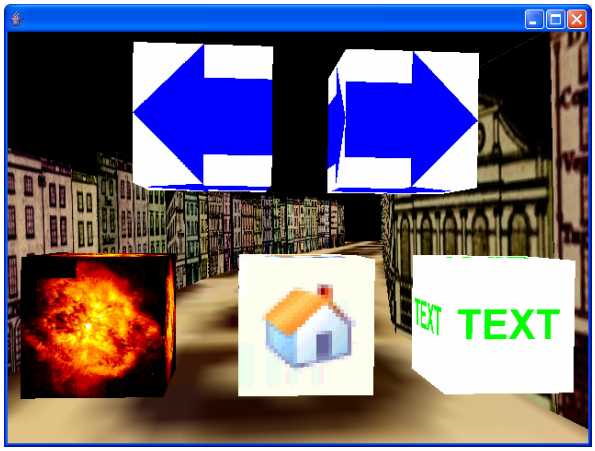

Figure 4. The 3D Menu

Upon selecting a menu item, the appropriate action is performed and the menu becomes invisible; this is illustrated in figure 3 in the state transition from state MENU to IDLE.

When the menu item with a picture of a house is selected, the user is transformed in the main intersection of the London walkthrough where the story about the 1765 fire begins. The two arrows at the top move the user to the next or the previous predefined 'bookmarks' to learn more about the 1765 London fire. The menu item with a picture of fire toggles between the fire and the non-fire version of the London walkthrough; this action take a noticeable time of 3 seconds. Figure 5 shows the main intersection before the fire, and figure 6 shows it after the fire (colored buildings are the ones that got affected by the fire). On the ground and also on top of the buildings the wind direction and how the fire moved from one point to the other, is shown.

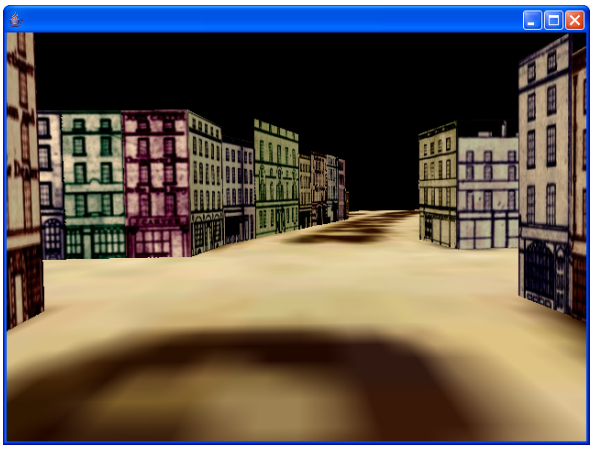

Figure 5. London walkthrough <u>before</u> the fire

Figure 6. London walkthrough <u>after</u> the fire

The menu item with the text "TEXT" on it toggles the visibility of the additional text we placed on top of the VRML files that tell the story on how the fire got

started, how it progressed, etc. Figure 7 shows one of the text templates that appear in the virtual environment (VE) when the "TEXT" menu item is selected and makes the textual information visible, while the user is in the fire version. Textual information can be displayed in both the fire version and the non-fire version of the London walkthrough.

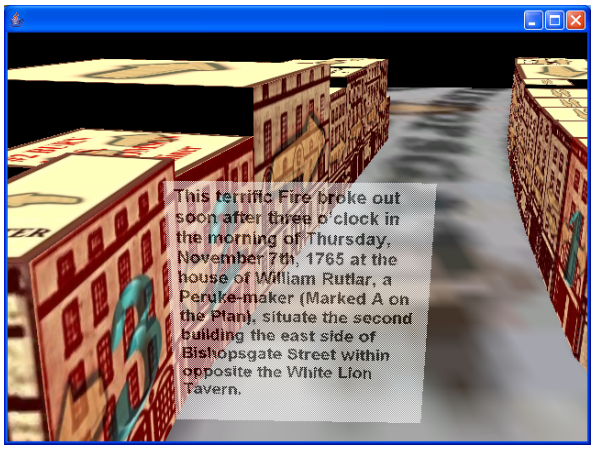

Figure 7. Textual information in the VE

The textual information is constructed from JPG images that are texture mapped onto semi-transparent three dimensional objects. The user can still see the environment behind the textual information while reading the information presented. We are planning on modifying this feature and have 3D text objects instead, which will enable a user to 'pinch' certain words that will transform the user to a specific place, which is similar to hyperlinks on a web browser.

Figure 8 is a bird's eye view of the London walkthrough. On the top of the buildings we display the names of each building. It is also visible the fact that the buildings match the contour of the street map.

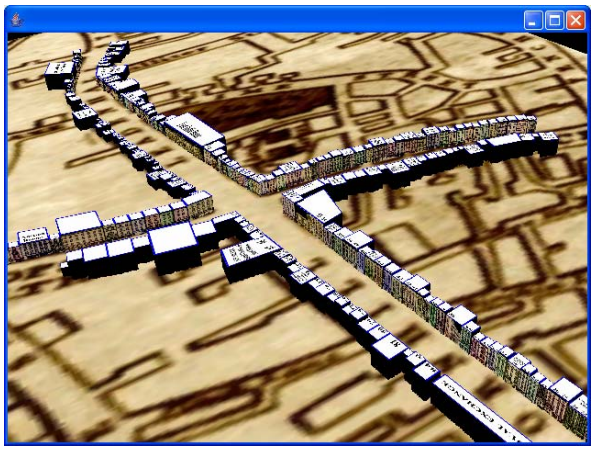

Figure 8. Bird's eye view

5. USER REACTIONS

Preliminary studies (with a few students) showed that the navigation techniques and interacting with the menu items were easy and seemed to be more natural than interacting in a VRML enabled browser. We plan further studies to find out if our proposed technique is indeed better than desktopVR and measure the difference.

6. CONCLUSION

Previous research [31] focused on the solution to the focus and context solution in digital libraries using desktop VR. In this paper we are extending the solution by utilizing immersive VR and preliminary results show that it is a better solution most importantly because in an immersive VR users get the feeling of being there. Secondly, interaction and navigation is performed more naturally, instead of clicking buttons in VRML enable browser, and dragging the mouse. The interaction techniques presented are user-driven. The users can freely navigate in the VE, focus on specific datum while the context stays in the surrounding.

7. ACKNOWLEDGMENTS

We thank Horn-Yeu Shiaw for her collaboration in this work, for providing the VRML files on which our virtual environment is based, and for her comments and guidance. We also thank Gregory Crane, founder of the Perseus Digital Library, for his support and collaboration throughout our work. And we thank the students at the University of Georgia who served as our first usability testers.

This work grows out of a collaborative project among the HCI groups in the Computer Science Departments at UGA and Tufts, the Perseus group in the Tufts Classics Department, the Tufts Psychology Department, and the Digital Collections and Archives Group at Tufts.

This research was principally supported by the National Science Foundation (NSF) and National Endowment for the Humanties under the Digital Libraries Initiative Phase 2 Program. It was also supported in part by the NSF Interactive Systems Program, Office of Naval Research, Berger Family Fund, the Tufts Selective Excellence Fund, and the Office of Instructional Support and Development at UGA.

REFERENCES

[1] "Readings in Information Visualization: Using Vision to Think,". Written and edited by: Stuart K.

Card, Jock D. MacKinlay and Ben Shneiderman, Academic Press, 1999.

[2] C. Chen. (1998). "Augmenting User Interfaces for Digital Libraries with Virtual Reality," In Proceedings 31st Annual Hawaii Int'l Conference on System Sciences, © IEEE,1060-3425/98.

[3] Crane, G. R. (Ed.) The Perseus Project. http://www.perseus.tufts.edu/, Tufts University, Medford, MA. 2003.

[4] D. Neuman. (1991). "Evaluating Evolution: Naturalistic Inquiry and the Perseus Project," Computers and the Humanities, 25. 239-246.

[5] E. Mylonas (1993). "The Perseus Project," In Scholarly Publishing on the Electronic Networks: The New Generation: Visions and Opportunities in Not-for-Profit Publishing: Proceedings of the Second Symposium, ed. Ann Okerson, 81-88. Washington, DC: Office of Scientific and Academic Publishing, Association of Research Librarie.

[6] G. Crane and E. Mylonas. (1998). "The Perseus Project: An Interactive Curriculum on Classical Greek Civilization," Educational Technology, 28. 25-32.

[7] G. Mar.ionini and G. Crane (1994). "Evaluating Hypermedia and Learning: Methods and Results from the Perseus Project," ACM Transaction on Information Systems, (12): 5-35.

[8] G. R. Crane. (1996). "Building a Digital Library: the Perseus Project as a Case Study in the Humanities," In Proceedings of 1st ACM Int'l Conference on Digital Libraries, 3-10.

[9] G. R. Crane. (Ed.). (1992). "Perseus 1.0: Sources and Studies on Ancient Greek Culture," Yale University Press, New Haven. CD-ROM.

[10] Home page of Xj3D http://www.xj3d.org.

[11] J. E. Hughes. (1998). "The Perseus Project: Implications for Learning and Teaching," Ph.D. Educational Psychology, Michigan State University. Able Media.

[12] J. Head. (2000). "Design Wise: A Guide for Evaluating the Interface Design of Information Resources," 2nd Printing, Information Today, Inc.

[13] J. Lamping and R. Rao. (1996). "The Hyperbolic Browser: A Focus+Context Technique for Visualizing Large Hierarchies," Journal of Visual Languages and Computing, 7(1), 33-55.

[14] J. Lamping, R. Rao, and P. Pirolli. (1995). "A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies," Inc Proceedings of CHI '95, ACM Press. 401-408.

[15] L. Deligiannidis and R.J.K. Jacob, "DLoVe: Using Constraints to Allow Parallel Processing in Multi-User Virtual Reality", Proc. IEEE Virtual Reality 2002 Conference, IEEE Computer Society Press, March 2002.

[16] L. E. Holmquist and S. Björk. (1998). "A Hierarchical Focus+Context Method for Image Browsing," In SIGGRAPH 98 Sketches and Applications, 282, ACM Press.

[17] L. E. Holmquist. (1997). "Focus+Context Visualization with Flip Zooming and Zoom Browser." In Proceedings of Conferenceof Human Factors in Computing Systems (Atlanta, Georgia USA). ACM CHI'97, 263-264.

[18] Leonidas Deligiannidis, Gamal Weheba, Krishna Krishnan, and Michael Jorgensen, "JWSU: A Java3D Framework for Virtual Reality", Proc. of the International Conference on Imaging Science, Systems, and Technology (CISST03), June 2003.

[19] M. Green and R. J. K. Jacob. (1991). "Software Architectures and Metaphors for Non-WIMP User Interfaces," Computer Graphics, Jul., 25(3): 229-235.

[20] M. Green and R. J. K. Jacob. (1991). "Software Architectures and Metaphors for Non-WIMP User Interfaces," Computer Graphics, Jul., 25(3): 229-235.

[21] N. D. Gershon, S. G. Eick, and S. K. Card. (1998). "New Visions of Human-Computer Interaction," ACM Interactions, Mar., 8-15.

[22] R .J. Hendley, N. S. Drew, A. M. Wood, and R. Beale. (1995). "Narcissus: Visualizing Information," In Proceedings of InfoVis'95, IEEE Symposium of Information Visualization, NY. 90-96.

[23] R. J. K. Jacob, H.Y. Shiaw, and G. R. Crane. (2001). "New Human-computer Interaction Techniques for the Digital Library," Technical Report 01-1, Jan, Dept. of Computer Science, Tufts University, Medford, MA.

[24] R. J. K. Jacob, L. Deligiannidis, and S. Morrison. (1999). "A Software Model and Specification Language for Non-WIMP User Interfaces," ACM Transactions on Computer-Human Interaction, Mar., 6(1): 1-46.

[25] R. Rao and S. K. Card. (1994). "The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information." In Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems, Information Visualization 1, 318-322, Color plates on 481-482.

[26] S. Björk and J. Redström. (2000). "Redefining the Focus and Context of Focus+ Context Visualizations," In Proceedings of IEEE Information Visualization, The Interactive Institute, Sweden.

[27] S. Björk and L.E. Holmquist. (1998). "Formative Evaluation of a Focus+Context Visualization Technique," In Proceedings of HCI '98 (poster), The British HCI Society, Sheffield, UK.

[28] S. Björk, L. E. Holmquist, and J. Redström. (1999). "A Framework for Focus+ Context Visualization." In Proceedings of Information Visualization, © IEEE, 53-56.

[29] S. Gauch, R. Aust, J. Evans, J. Gauch, G. Minden, D. Niehaus, and J. Roberts. (1994). "The Digital Library System: Vision and Design", In Proceedings of DL'94, ACM Conference on Digital Libraries, 47-52.

[30] S. K. Feiner. (1988). "Seeing the Forest for the Trees: Hierarchical Display of Hypertext Structures." In Proceedings of of COIS'88, Palo Alto, CA. 205-212.

[31] Shiaw, Horn-yeu, "New Interaction Techniques for the Digital Library: Three-Dimentional Focus+Context Interactive Visualization", Ph.D Dissertation, Tufts University 2003.

[32] V. Giller, M. Tscheligi, J. Schrammel, P. Fröhlich, B. Rabl, R. Kosara, S. Miksch, and H. Hauser. (2001). "Experimental Evaluation of Semantic Depth of Field, A Preattentive Method for Focus+Context Visualization," Technical Report TR-VRVis 2001-021, VRVis Research Center, http://www.vrvis.at/, Austria.