2D vs 3D, Implications on Spatial Memory

Monica Tavanti
Department of Information Science, Uppsala University Sweden
monica.tavanti@dis.uu.se

Mats Lind
Department of Information Science, Uppsala University, Sweden
mats.lind@dis.uu.se

Abstract

Since the introduction of graphical user interfaces (GUI) and two-dimensional (2D) displays, the concept of space has entered the information technology (IT) domain. Interactions with computers were re-encoded in terms of fidelity to the interactions with real environment and consequently in terms of fitness to cognitive and spatial abilities. A further step in this direction was the creation of three-dimensional (3D) displays which have amplified the fidelity of digital representations. However, there are no systematic results evaluating the extent to which 3D displays better support cognitive spatial abilities. The aim of this research is to empirically investigate spatial memory performance across different instances of 2D and 3D displays. Two experiments were performed. The displays used in the experimental situation represented hierarchical information structures. The results of the test show that the 3D display does improve performances in the designed spatial memory task.

1. Introduction

In everyday life, when we face the task of retrieving objects, one of the strategies used is to encode their spatial positions. Experimental results [7] established that processing spatial locations of objects is an effortless and unintentional process. Hasher and Zacks [5] demonstrated that the spatial locations of objects is processed automatically.

Many graphical user interfaces allow users to spatially organize information. A typical example is the desktop metaphor. This 2D environment is not an actual desktop and doesn't even look like one [8] but it nonetheless contains files that not only have names but are located *somewhere*, *inside* a certain folder, *next* to another file, *above* an icon, etc.

A new trend in GUI design is the production of 3D interfaces intended to support the storage and retrieval of textual and abstract data. The common belief behind this trend is that realistic 3D representations of the real world allow a more direct connection between information environments and their electronic representations.

Creating 3D spaces are intended to provide cues that naturally trigger natural cognition and actions. In contrast, 2D representations are thought to be more unnatural and require training to be used.

There is, however, a general lack of comparable experimental results assessing the supposed superiority of 3D in relation to 2D displays. This study tries to clarify whether and how 3D displays can better assist spatial cognitive abilities, specifically spatial memory. Two experiments involving 40 subjects were performed. The results of the experiments show that a realistic 3D display better supports a specific spatial memory task, namely learning the place of an object.

The following sections will first introduce some basic considerations regarding 3D hierarchical representations and the background work. Then, the two displays used in the experiments are described, followed by the specifics of the experimental design. For each experiment, there is a discussion of the results. The work is then summed up in the conclusions section at the end.

2. Hierarchical representations

In visualization tasks involving abstract data, it is very common that users are required to access structured information arranged in a hierarchical fashion.

There are two conventional ways of representing hierarchical data.

The first could be called *symbolic* because it uses names and special characters to represent the hierarchical structure. One example of a symbolic description of hierarchical data is the path names in DOS. The standards used in DOS need to be learnt by the users. For instance, the *slash* sign is used in combination with verbal labels to specify the sub-tree of the directories.

The second could be called *diagrammatic* because it tries to convey the structure to the users by means of visual expressions. A typical example can be found in so-called "tree views." In "tree views," the elements (folders or files) of the trees are icons linked through thin dashed lines; the depth of the nested elements is expressed through their positions along the *x* axis, etc.

Diagrammatic forms have also been implemented in 3D, for instance, in the form of 3D trees [10] [6] , where

Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)
1522-4048/01 $17.00 © 2001 **IEEE**

the leaves of the trees (files or folders and their labels) are at the base of an inverted cone.

It should be noted that such 3D representations are still abstract in that they require the user to learn certain conventions, since they do not resemble the things they refer to [11].

At a user interface level, those displays re-allocate the body of notations typical of 2D diagrammatic trees into a 3D environment (e.g., there are thin lines that connect the items of the structure at different levels of depth, etc.).

However, if we accept the use of 3D, a less abstract way of representing the hierarchical structure becomes available to us. If we position icons in space, we arrange them so that they form visible clusters and with some trickery these clusters can be made to convey a hierarchical structure, by more realistic means.

This more realistic hierarchical representation of information will be described and empirically evaluated in this work.

3. 2D vs 3D: background

By "3D" in this paper we specifically refer to 2D perspective projections of 3D environments. These 2D projections may also include other pictorial depth cues such as shading etc.

This work was inspired in part by some previous studies. The first study conducted by Cockburn and McKenzie [2] is an evaluation of the Cone Trees interface [10]. The research compared an implementation of Cone Trees to a normal tree-styled interface; subjects were engaged in text-based search and location tasks. The results showed slower time performances for the Cone Trees interface.

The more detailed results of the study indicated that the Cone Trees display had the special characteristic of quickly providing a sense of the global structure and of the density of information, because all the data was available in a single screen. On the other hand, when the amount of information was very dense, users found it difficult to discern and to find the textual labels of each item, since they overlapped.

In the Data Mountain interface [9], this problem has been avoided. Data Mountain consists of a 3D inclined plane in which the thumbnails of documents (in the specific case, Web pages) are vertically positioned. Some of the pages can be occluded by other pages, but no thumbnail is completely hidden, so all the information is accessible at the same time. An experimental investigation [9] was carried out in order to compare Data Mountain to the Microsoft Internet Explorer Favorites mechanism, a typical 2D environment. The subjects' task was to store and retrieve Web pages using the two displays. The results suggested a superiority of Data Mountain (both for time and accuracy). The authors also speculated that spatial memory played a role in the 3D environment, since subjects explicitly stated that they remembered a page's location.

A follow-up study on Data Mountain [4] engaged 9 of the same the subjects in a similar task, after 4 months. This study also evaluated the role of thumbnails in the performance. The results showed that spatial memory performance was independent of whether the thumbnails of the pages were present or not on the display. As a matter of fact subjects performed well even with blank icons as retrieval cues. Moreover, after 4 months, subjects were still fairly good at retrieving the web pages they had stored during the previous study.

Additional studies [3] compared Data Mountain to a 2D version of the same display across a very similar task used in [9]. The results showed that subjects performed faster using the 2D display. Nevertheless, Data Mountain had the advantage of presenting the information in a single screen and in an original and more natural way.

In real world environments documents are usually kept horizontally on a surface (e.g. pages on a desk). In the Data Mountain interface, every document is simply tilted upwards and placed, vertically, on a surface. However, all tested versions of Data Mountain lack precise techniques to represent a hierarchical structure.

A study by Ark, Dryer, Selker and Zhai [1] made use of homogenous and comparable displays to evaluate 2D and 3D environments. The study compared reaction times in an identification and location task using different instances of the same display. It consisted of typical objects which can be found in an office (telephone, desk etc) represented in either a 2D or 3D fashion. These objects were placed either on a flat 2D background or in a 3D representation of the office. The authors refer to the 3D depictions of objects and background as an "ecological" layout, that is a more realistic and natural layout. The results of the study revealed that the 3D ecological layout improved subjects' performance; the authors also suggest that for tasks that require identifying and learning the objects' location, 3D realistic and ecological layouts improve users' performance.

The theoretical conclusions of these studies can be summarized as follows:

a) 3D displays can be exploited to visualize large sets of hierarchical data;
b) The perspective nature of 3D representations makes it possible to show more objects in a single screen, objects shrink along the dimension of depth;
c) if more information is visible at the same time, users gain a global view of the data structure;
d) there is experimental evidence that 3D ecological displays enhance subjects' spatial performances.

It is possible to use this set of ideas together with the notion of hierarchical organization to create more realistic and ecological representations of hierarchically structured information.

To do so, we created a 3D tree in which the spatial relationships of the constituting elements were expressed in a very natural way. The elements of the tree — simple rectangles — were vertically arranged in space, as in the Data Mountain display. But in our display (shown in figure 2), the data are arranged in a hierarchy.

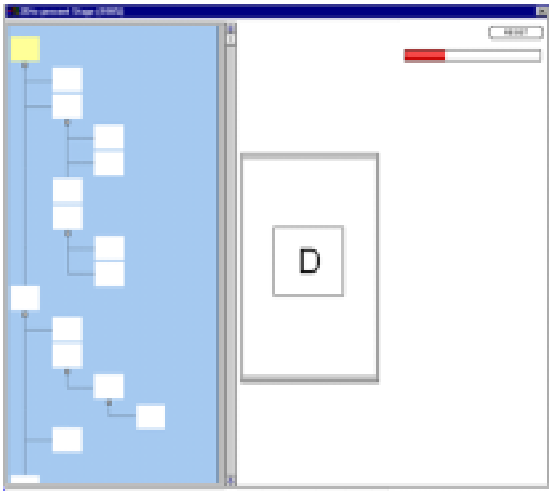

Figure 1: The 2D display

The *z* axis (depth) is used to indicate the position of the nested elements; the rectangles are placed in perspective from the user's point of view (so, the elements placed in higher positions of the tree are bigger, while the deeper ones are smaller). The rectangles in the tree are also properly juxtaposed so as to represent their logical distribution within the nodes. The display we created was experimentally compared to a more standard 2D tree (similar to Windows Explorer). The main purpose of the experiments was to evaluate one specific, but important, aspect of interface usage: the support of spatial memory. Does our proposed representation of hierarchical data better support users in spatial memory tasks (memorize the location of objects)?

4. Experiment I

The hypothesis formulated for this experiment was that the 3D tree we created would be more effective, in a task involving spatial memory, than a conventional 2D representation of the same tree.

As stated above, it is obvious that in a 3D representation, more objects can be present in a screen than in a 2D representation. Thus, the 3D display was a static environment in which all the elements of the tree were represented in the same screen, while the 2D display was embedded in a scrolling window although the actual 2D view port used was of exactly the same size in both conditions.

The independent variable was "interface type" with two levels (3D and 2D). In order to avoid potential carry-over effects between the two instances, interface type was a between-subjects factor.

4.1. The displays

The displays can be described as follows:

a) The 2D tree (shown in figure 1) was comprised of a scrolling window whose visible part roughly corresponded to 55% of the entire tree (329 pixels of width * 666 pixels of height). The tree consisted of 27 white rectangles, divided into four nodes (sub-trees) and articulated into four levels of depth. Dashed lines connected the rectangles to signify the nested structure of the tree. Subjects were requested to click on any rectangle to uncover the alphanumeric character associated with the rectangle. The character was shown on a small window on the right of the display. Every time a rectangle was clicked, it became highlighted and the corresponding character appeared in the small window. On the upper-right-hand corner of the display, a time bar indicated the time flow.

b) The 3D tree (shown in figure 2) was composed of a window whose size corresponded to the visible part of the 2D window. The tree consisted of 27 white rectangles, divided into four nodes (sub-trees) and articulated into four levels of depth. In this version, the nested structure of the tree was expressed in terms of depth, so that higher levels were represented by larger rectangles, while deeper levels were represented by far smaller rectangles. Also, the logical distribution of the rectangles in the four sub-trees was respected and represented. A system of shadowing was created to provide a more realistic three-

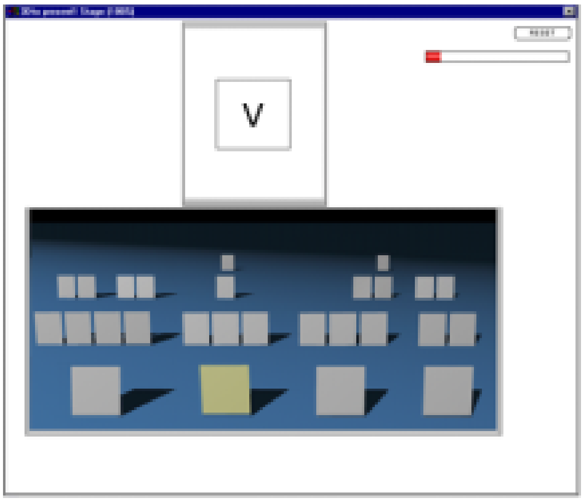

Figure 2: The 3D display

Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)
1522-4048/01 $17.00 © 2001 IEEE

dimensional effect. As in the 2D condition, subjects were requested to click on any rectangle to uncover the alphanumeric character associated with the rectangle. It was shown in the same window as in the 2D case, except that the window was disposed above the 3D display.

4.2. Method

As mentioned above, the content of each element of the trees was an alphanumeric character. There were important reasons for choosing this method. The first was that spatial memory plays a significant role in information storage and retrieval tasks. For such tasks, people mainly deal with textual data, so we had to insert a textual substance as the main object of the task. However, labeling any element of the tree with a word would involve a risk. Introducing semantics within the context of the experiment could have shifted the experimental focus; for instance, subjects may have different levels of familiarity with the subject matters, which might affect their ability (favorably or adversely) to remember the words.

The alphanumeric characters used in the test were 27; we used the Swedish alphabet, which actually contains 27 characters. Another point related to the alphanumeric characters is that they were not shown directly in the rectangles, but in a separated window. This choice can be justified as follows.

In the 3D tree, the rectangles are different, i.e., they have different sizes and different background luminance. A separated visualization window made it possible to keep to a single size and the same background luminance for all the characters in both the 3D and 2D displays. Furthermore, a separated visualizing window forced subjects to sharply concentrate their attention on the mnemonic task.

4.2.1. Subjects. 20 graduate students from the University of Uppsala, Sweden, participated (10 females and 10 males). Their ages ranged from 25 to 48 years (mean: 30.9). All of them were native Swedish speakers, used computers on a daily basis and took part in the experiment as volunteers.

4.2.2. Equipment. The study was run on a Pentium II machine, with 128 MB of memory and a 19-inch display with 1024*768 resolution. All of the tasks were written in Macromedia's Lingo. The 3D display was made using Cinema 4D, version 5.10 XL.

4.2.3. Tasks. Subjects were requested to click on the rectangles in any order to uncover all the characters, following a personal strategy. The subjects' goal was to memorize as many characters' positions as possible.

The test was articulated in three phases. First, the subjects had to explore the displays (2D or 3D, depending on the group), knowing that they had two minutes to complete the exploration of the display. In the second phase, subjects had to fill in a short questionnaire (which could take from one to two minutes). The questionnaire contained items that were unrelated to the task. It was used to distract any possible rehearsing of the characters.

The third phase constituted the task phase. Subjects were given 5 minutes to complete this part of the test. During this phase, subjects were given a version of the display that was identical to the one used in the exploration phase (a 2D or 3D display, depending on the group). This version of the display was constructed to present the sequence of characters (one by one) to the subjects. Ten different combinations of the sequences were randomly arranged. Each sequence's combination was randomly assigned to one of the subjects.

Subjects were asked to associate any of the alphanumeric character presented in the display to a rectangle. They only had one guess per character. No feedback concerning their decisions was given.

4.2.4. Procedure. Subjects were randomly assigned to one of the two groups (10 subjects for the 2D condition and 10 for the 3D condition), with the constraint that the number of females and males in both groups was to be the same.

Through written instructions, they were informed of how to execute the tasks, and they were allowed to ask questions about the instructions, but only before the beginning of the test.

5. Results

As mentioned above, the goal of the experiment was to investigate whether display (2D or 3D) effects spatial memory performances. In order to structure the results, two parameters were processed: *a)* the number of correct responses (that is, all the characters that were associated with the correct position during the task phase); this parameter was used to evaluate the general performance of the subjects interacting with the display; *b)* the association of an alphanumeric character to the correct depth level in the tree.

Since this was an exploratory study, a significance level of .05 was chosen as the decision criterion, although 2 statistical tests were performed.

5.1. General performance

A comparison of performance using the two layouts is shown in figure 3. Due to the non-homogeneous variance between the two groups, a non-parametric test (the Mann-Whitney U test) was carried out. The analysis revealed a statistically significant effect of the application (U=10.5, p<.01). The average number of correct responses for the 3D condition was 10.2, while the average for the 2D condition was 4.2. Performances were in other words reliably superior with the 3D display.

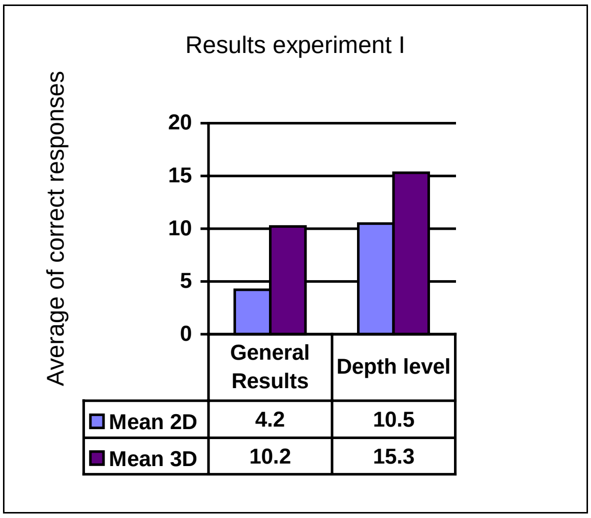

Figure 3: Means of the scores of the two groups

5.2. Position in the depth level

A second analysis was performed to compare how many characters had been correctly associated to the appropriate depth level of the tree. As already mentioned, the trees of the displays were articulated in 4 levels of depth. The mean of correct scores using the 3D display was 15.3, while in the 2D case, the mean was 10.5. The difference in performances is statistically significant (U= 20, p<.05).

6. Discussion

The results of the experiments seem to confirm that there is an improvement in performance with the 3D realistic display. The results related to the correct location of the alphanumeric character in the depth level reveal that this played a role in increasing performance. It is feasible to deduce that this result could be due to a special characteristic of the rectangles disposed in the 3D display. As a matter of fact, all of them are arranged in rows and their sizes are homogeneous according to the depth in which they are situated (as shown in figure 2). This mapping of the size of the rectangles may be the reason for the improved performance. In other words, it is reasonable to argue that an ecological representation of the logical relationships existing within a hierarchy, as illustrated in figure 2, may enhance the ability to locate horizontal structures in nested data.

To summarize, according to the results of the first experiment, the 3D display seemed to contribute in significantly improving spatial memory performance. However, this effect may not be attributable to the use of a realistic 3D tree, but to other factors. For instance, the subjects of both groups had two minutes to explore the displays and to memorize as many positions as they could. But the subjects who had to interact with the 2D display had to use a part of this time to scroll the window, while the subjects who performed the task with the 3D were able to exploit all the available time in the memorization task (since they did not have to scroll). This difference in the time allowance could be the reason behind the poorer performance of the "2D group."

This hypothesis was tested in a second experiment. The following sections will describe the new experimental conditions, the results, and the conclusions.

7. Experiment II

The second experiment aimed at testing whether the 3D tree was still superior in the spatial memory task, even when compared to a non-scrolling and static 2D display. The experimental design and the same conditions as in experiment I were retained.

7.1.The displays

An identical version of the 3D display was used in this experiment, while a new 2D display (shown in figure 4) was created to test our hypothesis. This 2D tree contained

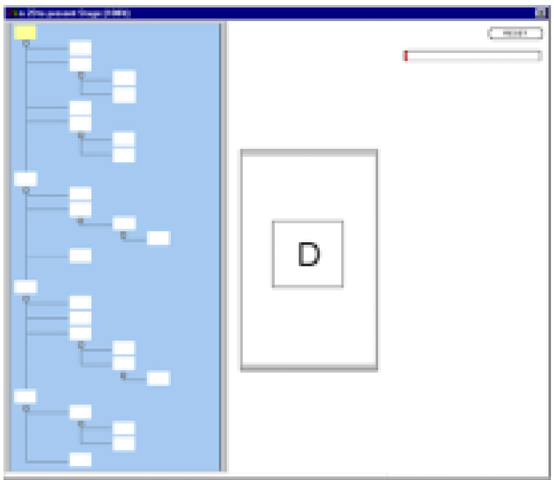

Figure 4: The 2D display

the same amount of rectangles (27) as in experiment I, but their sizes were shrunk proportionally to allow for all of them to be visible in a non-scrollable window. The remaining elements and properties of the display were equal to the 2D display of experiment I (i.e., the dashed lines connected the rectangles, clicking on a rectangle caused it to be highlighted etc).

7.2. Method

The same task and the same procedure as in experiment I were used, with the exception that subjects of both groups were given three minutes to explore the interface and that the non-scrolling version of the 2D display was used in the 2D condition.

7.2.1. Subjects. 20 undergraduate students from the University of Uppsala, Sweden, participated (10 females and 10 males). Their ages ranged from 19 to 30 years (average 24.95). All of them were native Swedish speakers, used computers on a daily basis and took part in the experiment as volunteers

8. Results

As mentioned, our focus of interest was to find out if layout (2D *vs* 3D) affects spatial memory performances when two equivalent static displays are used.

The parameters involved in structuring the results were the same as for experiment I: *a)* the number of correct responses (that is, all the correctly placed characters in the task phase); *b)* whether or not a placed character belonged to the appropriate depth level in the tree.

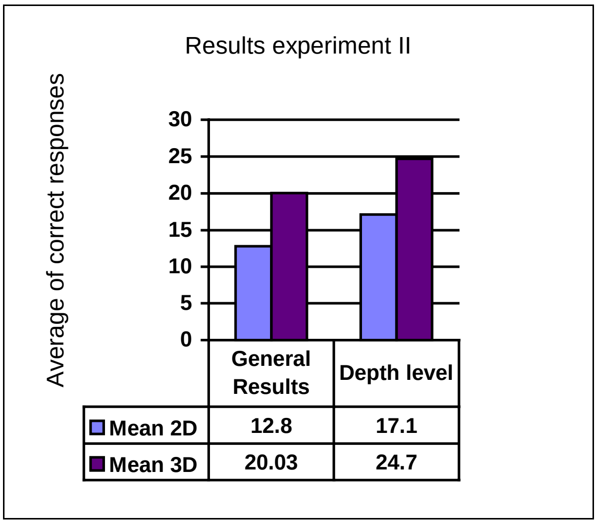

Figure 5: Means of the scores of the two groups

8.1. General performance

A comparison of performance using the two layouts is shown in figure 5. Because of the non-homogeneous variance between the two groups, a non-parametric test (the Mann-Whitney U test) was carried out. The analysis revealed a statistically reliable effect of the application ($U = 21.5$, $p<.05$). The average for the 3D condition was 20.03, while the average for the 2D condition was 12.8.

8.2. Position in the depth level

In this experiment as well, the correct associations to the appropriate depth level of the tree was examined. The average for the 3D condition was 24.7, while the average number for the 2D condition was 17.1. The difference between the two performances is significant and reliable ($U = 14.5$, $p<.01$).

9. Discussion and conclusions

According to the results of the second experiment, it is feasible to sustain a general superiority of the 3D display for the chosen spatial memory task, even when compared to a 2D non-scrolling display.

The results of both experiments confirm that the 3D display better supported the task of correctly locating the characters on the depth level. For instance, the 3D version of the display uses a simple rule of perspective to express the nested relations among the rectangles: the deeper the rectangles, the farther (along the *z* axis) they go and the smaller they get. In contrast, in the 2D tree, those relations are stated through diagrammatic means. The ecological and natural style exploited in the 3D representation could well explain the superiority of the 3D display in both the experimental sessions. Thus, the creation of more natural ways to visualize hierarchical data should be strongly considered during the interface design process.

The results of the experiments also revealed that other issues related to the design of 3D displays need to be further investigated. It is, for example, reasonable to hypothesize that the natural appearance of the 3D display used in the tests did not actually affect the improved performance. For instance, the rectangles of the 3D display were characterized by different sizes; perhaps, the very size of every rectangle was used as a retrieval cue in the memory task. In other words, perhaps a single visual property of any rectangle (i.e., size) was used as a reminder for the memory task. If this hypothesis is true, the simple size differentiation of the elements disposed in the display would have been sufficient to better support the task and to give improved results.

In conclusion, the results of the experiment showed that subjects performed significantly better using the 3D display. However, it is still unclear to what extent single visual properties of the 3D space affected the performance.

This last question will be the object of future work.

1522-4048/01 $17.00 © 2001 IEEE

10. References

[1] Ark, W, Dryer, D.C., Selker, T., Zhai, S., Representation matters: the effect of 3D objects and a spatial metaphor in a graphical user interface, in H. Johnson, N. Lawrence, C. Roast (Eds), *People and Computers XIII, Proc of HCI'98*, Springer, 1998, pp 209 -219.

[2] Cockburn, A., and McKenzie, B., An Evaluation of Cone Trees. People and Computers XIV: *British Computer Society Conference on Human Computer Interaction* Springer-Verlag, 2000, pp 425--436.

[3] A. Cockburn and B. McKenzie, "3D or not 3D? Evaluating the Effect of the Third Dimension in a Document Management System". Proceedings of CHI 2001, ACM Conference on Human Factors in Computing Systems, CHI Letters 3(1), 434-441.

[4] M. Czerwinski, M. van Dantzich, G. Robertson, and H. Hoffman, "The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3D", Interact '99, Edinburgh, UK, August30, 1999.

[5] Hasher, L., and Zacks, R.T., Automatic and effortful processes in memory. *Journal of Experimental Psychology: Genera*l, 1979, 108, pp. 356-388.

[6] Hearst, M. and Karadi, C., Cat-a-Cone: An interactive interface for specifying searches and viewing retrieval results using a large category hierarchy. In *Proceedings of the Twentieth SIGIR Conference*, Philadelphia, Pennsylvania, Association for Computing Machinery, 1997.

[7] Mandler, J.M., Seegmiller, D., and Day, J., On the coding of spatial information. *Memory & Cognitio*n, 1977, 5, pp. 10-16.

[8] Nelson, T. H., The Right Way to Think About Software Design. In B. Laurel (Ed.), *The Art of Human-Computer Interface Design* Reading, Massachusetts, USA: Addison-Wesley Publishing, 1990, pp.235-243.

[9] Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D., and van Dantzich, M., Data Mountain: Using spatial memory for document management. In *Proceedings of UIST'98*, San Francisco, California, November 1998, ACM Press, pp 153-162.

[10] Robertson, G.G., Mackinlay, J.D. and Card, S.K., Cone Trees: Animated 3D visualizations of hierarchical information. In *Proceedings of SIGCHI'91*, 1991, pp.189-194.

[11] Scaife, M. and Rogers, Y., External cognition: how do graphical representations work? *International Journal of Human-Computer Studies*, 1990, 45, pp. 185-213.