Towards scalable interfaces using spatial cues for document management

Eddie Edwards
Department of Computer Science
University of Canterbury
New Zealand
Supervisor: Dr. Andy Cockburn

November 2, 2001

Abstract

Document management is an unavoidable task. Previous research evaluated the effect of building document management systems exploiting spatial memory. It was proven that spatial memory enhances retrieval tasks for systems dealing with a small number of data items. We investigate the effect of spatial cues with scalable systems. We present EDM (Enhanced Data Mountain), a scalable picture management system based on the Data Mountain by Robertson, Tesle, Tversky & Mullet (1991). Our system exploits the use of spatial cues and spatial memory in retrieving images. This report explains the design methodology behind EDM, in addition to an evaluation that compares EDM with Microsoft's My Pictures (MP) system for managing pictures. Results show that our subjects were faster at retrieving pictures in the display when using the MP interface, but not significantly so. As expected, the more time subjects spent browsing through the picture groups, the faster they became at retrieval. Subjective satisfaction rating showed an overall preference favouring the MP system over the EDM system. A conclusive statement can not be made regarding the exploitation of spatial memory in scalable systems.

Contents

Chapter 1

Introduction

"It is through science that we prove, but through intuition that we discover."

—Henri Poincare

1.1 Overview

As computers become widely used in various areas of science, arts and entertainment, the need for effective document management systems becomes imperative. This report addresses the main issues involved in retrieving large numbers of documents using spatial cues.

Document management systems have evolved dramatically over the last decade. Until the the advent of the graphical user interfaces in the early 1980s, users used text-based interfaces to manage their documents. Nowadays, there are sufficient resources to support a complex – 3D, if required – graphical interface in real time. From simple text-based file management interfaces, document management has evolved to include a suite of specialised systems, each built specifically to manage one type of document.

The Data Mountain is a document management system that attempts to ease existing problems in document managing by exploiting the advantages of humans' spatial memory. It has been proven to be a more effective way of browsing and navigating bookmarks than the existing bookmarking systems (Robertson, Czerwinski, Larson, Robbins, Thiel & van Dantzich 1998). Applications of the Data Mountain not only apply to visualising web-pages, but can also be used to manage personal collections of pictures. Figure 1.1 shows Cockburn & McKenzie's (2001) implementation of the Data Mountain on which our system was based on.

As digital cameras become cheaper and more popular, it is estimated that the average user will have hundreds, if not thousands, of pictures on their personal computer. With this number of documents, an effective picture management system must be implemented.

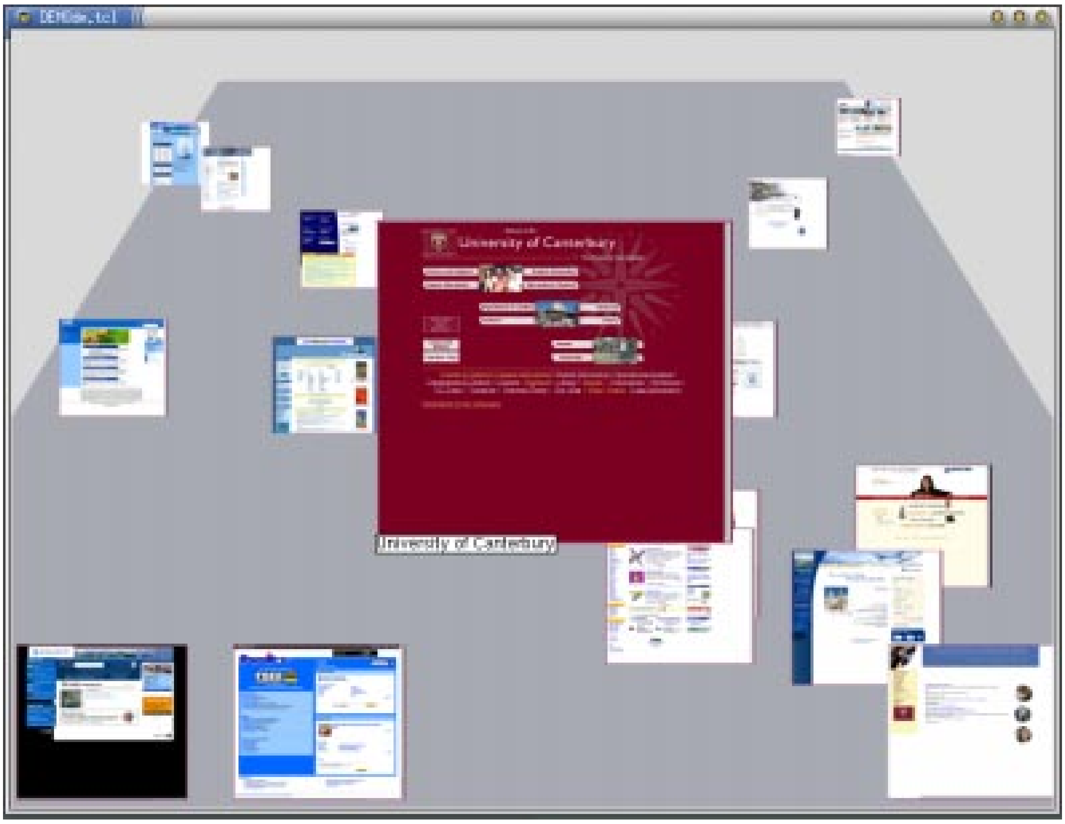

Figure 1.1: The Data Mountain developed by Cockburn & McKenzie (2001)

This research uses the problem of picture management as a platform for exploring scalable spatial memory. Our picture management system is based on the Data Mountain.

1.2 Motivation and Problem Definition

The results of the evaluation included in Robertson et al. (1998) indicate that the Data Mountain is significantly more effective that Internet Explorer's Favourites mechanism. This significant difference is mainly due to the Data Mountain's effective use of spatial memory. However, The results of the evaluation of the Data Mountain included in Cockburn & McKenzie (2001) indicate that the Data Mountain becomes significantly less effective as the number of pages increase. This system limitation motivated us to modify the behaviour of the Data Mountain to manage larger numbers of documents and hence resolve the scalability issues. We were also interested in discovering whether spatial memory hold in a system that does not show all data on one screen. Our system use scrollbars as the main means of navigation and consequently, not all the documents show on the screen at any one time. We were interested to know whether users would still be able to form associations between objects and their location on a large system that does not show all the documents at once.

1.3 Report Structure

Chapter 2 gives an overview of the previous work undertaken in the area of spatial memory in conjunction with document management systems.

Chapter 3 explains the iterative design process that resulted with the final prototype; EDM – the Enhanced Data Mountain.

Chapter 4 discusses the evaluation study undertaken to assess the effectiveness of the system along with the results, discussion, system limitation and potential areas of further work.

Chapter 5 summaries the results and proposes some ideas for further research.

Chapter 2

Related Work

"Copy from one, it's plagiarism; copy from two, it's research."
—Wilson Mizner, screenwriter.

This chapter presents the previous work that was carried out in the areas of document management system, concentrating on those exploiting spatial memory. It discusses the previous evaluations of the Data Mountain from which our system evolved.

2.1 Spatial Memory

Spatial memory refers to the memory of spatial information, such as the geographical layout of cities or the interior layout of a building. Spatial memory involves remembering the location of items, by forming associations between the items and their environment. One of the scientific papers that describes the different types of memory (spatial, long-term, short-term, recall, emotional, etc) is by Trumbo (1998). The author uses multimedia as a tool to illustrate the difference between the different memory types and uses. Trumbo states that multimedia/HCI designers face major challenges "The challenge for the multimedia designer is to create an environment of sensory elements that is memorable". It is important to create memorable multimedia and interfaces to solve the problem of overloading users' memory. Creating system depicting memorable environments eases the task of understanding those systems. It also allows users to form effective associations between the environment and the newer items.

Dumais & Jones (1985) conducted an experiment comparing the times and correctness ratio obtained from retrieving thirty documents (news articles) per subject using symbolic and spatial representation of the different documents. The aim of the experiment was to assess spatial memory in building computer interfaces. The subjects were presented by a statement that described a news article and they were asked to retrieve the documents they have previously filed (placed). The paper compares four filing schemes: name only, location only, name and location spatially separated, and name and location combined in the same area.

Surprisingly, results show that the location only technique performed significantly worse than the others. The authors explain the reasons behind this unexpected result in detail. Firstly, the experiment does not separate between recallability and recognisability of the news articles, which makes determining the exact reasons behind the failure of spatial memory uncertain. Secondly, the items being new articles may have affected the experiment as news articles are not naturally spatially recognisable. Third, the news articles were filed according to their textual contents and not their visual appearance. This is also believed to have affected the speed and accuracy of retrieval.

However, the authors assert that spatial memory can be used to guide users to the general location of the correct item, but cannot be employed to retrieve the exact location. Also, the authors do claim that filing documents on a computer has more degrees of freedom. In computers, filing may be based on several cues and matrices including size, colour, shape, name, location, content, etc.

Another evaluation study (Donelson 1978) investigated the effect of spatial memory in navigating through a Spatial Database Management System (SDMS). In this evaluation, the subjects used a joystick attached to a chair to navigate the database, to resemble the navigation of an airplane, thereby utilising the subjects' knowledge about space and direction. The author claims navigating and using the spatial database system is similar to managing one's items on a desktop. A shortcoming of this paper is the lack of an empirical evaluation. However, the authors conclude that "user response to SDMS has confirmed spatial management of information as an outstanding concept". In summary, the paper presents a new domain that can benefit from spatial memory.

2.2 Document Management

The most common Document Management Systems (DMSs) use hierarchical structures to organise data. Despite of the simplicity and intuition of the approach, hierarchical structures impose restrictions on the users' mental model of how the documents should be organised. Furthermore, these systems separate documents from the activities that users perform on them. As we will see, most DMSs try to contradict the metaphor and organise documents using new techniques aimed at more effective and efficient management. Our work follows the trend and presents a non-hierarchical organisation mechanism.

Research has been conducted into visualising hierarchical structures in a more appealing way, such as Cone Trees, shown in Figure 2.1. Cone Trees are 3D animated trees aimed at visualising hierarchical information with efficient use of

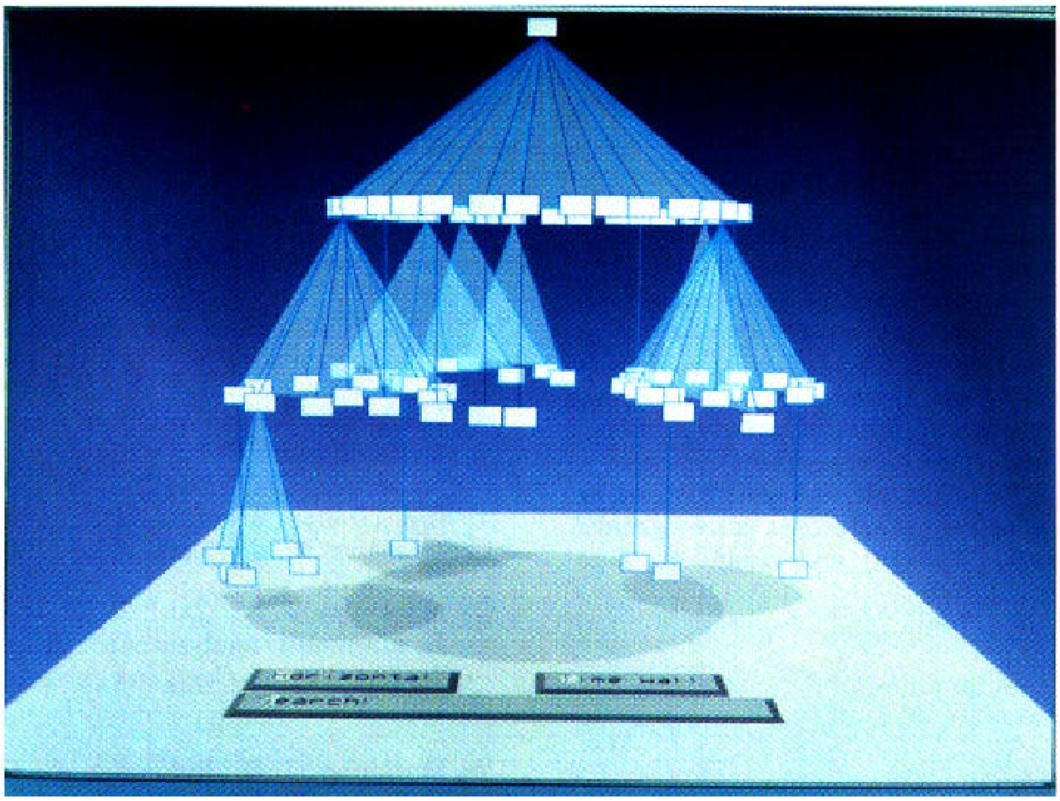

Figure 2.1: A screen shot of Cone Trees

screen space. Cone Trees extends the simple hierarchical structure by introducing the third dimension and hence the illusion of space. The animation techniques shift the load from the user's cognitive memory to spatial recognition. The approach has definite potential in visualising any kind of hierarchical structure, such as a file structure. Analytical space-efficiency comparisons between two-dimensional and three-dimensional trees favour a 3D tree. The authors claim that "the emerging technologies of 3D and interactive animation, and the human perceptual system can be effectively exploited to improve management and access of large information spaces." However, we see no proof of this claim. No empirical evaluations assess the effectiveness of Cone Trees with regard to ease-of-use and time to achieve certain tasks (Robertson, Mackinlay & Card 1991).

Research into developing new systems for window management, can also be applied to document management systems. Robertson, van Dantzich, Czerwinski, Hinckley, Thiel, Robbins, Risden & Gorokhovsky (2000) present a 3D window manager system called the Task Gallery, shown in Figure 2.2. The system depicts a room where windows are hung on the walls as drawings are hung on the walls of an art gallery. The user's navigations are designed to be similar to those of

Figure 2.2: A screen shot of the Task Gallery

navigating through a real art gallery. Users can walk around entering different rooms of the gallery. In addition, they can look up, down, left and right to see the windows that are hung on each of the walls, ceiling and floor. Some empirical experiments were conducted during the design and implementation phases to support the design decisions. However, there was a lack of experimentation showing the merits, or otherwise, of using the Task Gallery to using a normal windows management system. In summary, the paper presents a novel approach to window management which exploits spatial memory and human visual perception and knowledge of 3D spaces. Like our system, the Task Gallery uses a real-world metaphor (rooms and art galleries) to create a memorable environment aiding the user's cognitive ability and exploiting spatial memory.

Another system that relates to our research is Workscape, shown in Figure 2.3. The system exploits the advantages of the third dimension and the illusion of depth. The designers of Workscape assert that "the typical two-dimensional display of documents on a screen wouldn't provide the rich working environment we desired". Interestingly, the interface's 'look and feel' was inspired by Italian Renaissance paintings. These 14th and 15th centuries' paintings used depth cues and three-dimensionality in ways that were not seen before. The designers illustrate

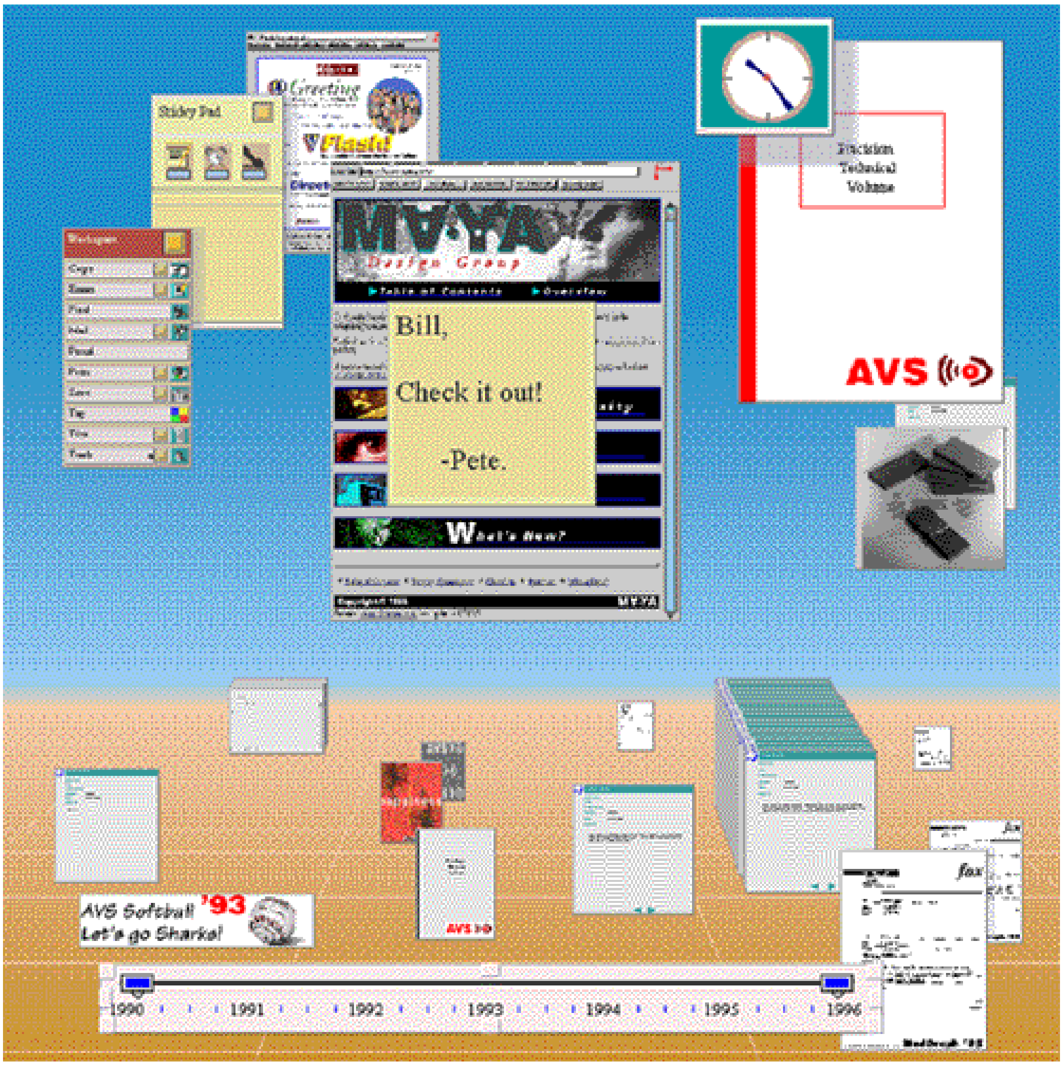

Figure 2.3: A screen shot of Workscape

the similarity between the visual design of their system and those paintings using graphical examples. The main advantages of the design of Workscape is that it is able to deal with hundreds of, always open, always displayed, documents at once. This feature makes Workscape usable for managing a larger suite of documents: both text and image (Ballay 1994). Workscape is closely related to the system we have developed. We also wanted to maintain hundreds, if not thousands, of always open documents at once. However, we also seek effective categorisation of documents which was a feature that was provided by Workscape.

As we will explain in detail in chapter 3, our system is based on the Data Mountain, an Internet bookmarks management system developed by Robertson et al.

(1998), called the Data Mountain. The Data Mountain is based on the use of 3D animation and spatial memory. The system depicts a mountain on which people can arbitrarily place the thumbnails of the pages being bookmarked. Texture and colours are added to the structure of the mountain in order to provide visual cues via landmarks optimising the use of spatial memory. The authors clearly indicate that the motivation behind their approach is to minimise the user's cognitive load by shifting it to the human perceptual system and the natural knowledge of 3D spaces. A user study compared the effectiveness of the Data Mountain, and Internet Explorer's (IE) Favourites bookmarking scheme. The Data Mountain proved to be more effective as a bookmarking system than IE's Favourites. Also, the user satisfication measures showed a reliable preference towards the Data Mountain. The authors claim that the reason behind the success of the Data Mountain is of a combination of three-dimensionality and spatial memory.

An empirical evaluation by Cockburn & McKenzie (2001), compared the three-dimensional Data Mountain to a two-dimensional Data Mountain, where the thumbnails do not get scaled if they are pushed up the mountain. The experiment also analysed the correlation between the number of thumbnails and the time to retrieve pages. The experiment showed that the reason behind the success of the Data Mountain is spatial memory only, as there was no significant difference between 2D and 3D Data Mountains. However, users preferred the 3D Data Mountain.

A number of design limitations were identified from the experiments of Cockburn & McKenzie (2001) and Robertson et al. (1998) and the resolution of these limitations were applied to our picture management system. From Robertson et al.'s (1998) experiment we noted that a significant number of users requested a 'grouping' feature that would allow them to group related items. The coloured textured circles occupying different spaces on the mountain's body, did not offer an obvious visual cue that aids the users in grouping the related bookmarks. From Cockburn & McKenzie's (2001) experiment we have noticed that the Data Mountain would not be able to manage more than a 100 page efficiently. We have also noticed that the use of 3D is preferred by the majority of users, although it does not influence the efficiency or effectiveness of the retrieval tasks.

2.3 Prior Work on Picture Management

A few papers discuss image indexing, and query-based image retrieval systems. Of these papers, the most relevant paper to our research is the paper by Rodden, Basalaj, Sinclair & Wood (2001) which discuss the organisation of pictures. They question whether organising by similarity assists in browsing and locating images. The effectiveness of organising images according to their similarity versus presenting pictures randomly – sorting by name for example, is compared subjectively. They found that similarity of organisation aid users in retrieving images. However, one draw-back of grouping similar images is that images appear to merge making it harder and slower to identify images. It would have been beneficial to see some empirical timing tests giving insight on efficiency merits, or otherwise, of grouping similar pictures. In general the paper presents a good research experiment that is "of interest to anyone designing a system that involves presenting sets of images to users".

To date, no picture management systems have been scientifically proposed. However, since the photographic industry was revolutionised by the commercialism of digital camera in 1995, the need for effective picture management systems became the next logical step. The lack of effective picture management systems contributed to our incentive to carry out our research.

Chapter 3

System design

"Inventions have long since reached their limit, and I see no hope for further development."

—Julius Sextus Frontinus [1]

Our aim is to enhance the Data Mountain, which was proven to be an effective document management system. We expect that in a few years, digital camera owners would have thousands of pictures on their desktop computer. We aim to scale the Data Mountain to handle the requirements of digital camera owners.

There are some essential properties that we intended to have in our Enhanced Data Mountain (EDM) system:

Scalability. Our main goal is to design a system that can potentially handle as many pictures as required. For experimental use, we have limited our prototype to handling 1000 thumbnails.

Spatial cues. Our aim is to incorporate the power and speed of spatial memory by supplying effective spatial cues.

Visual cues. In order to effectively group pictures, pictorial cues have to be provided to aid the user in finding groups and retrieving pictures.

Effective Grouping. In order to deal with thousands of pictures, grouping of pictures becomes imperative. Users have to be able to group their pictures in any arbitrary way. The system has to support the flexibility of grouping.

Ease of navigation. Ease of comprehension and use of the prototype is paramount. Navigating the system has to be simple and easy.

Metaphor correspondence. It is important that the system corresponds to the metaphor it depicts. The current metaphor is the metaphor of a mountain.

[1] Highly regarded engineer in Rome, 1st century A.D.

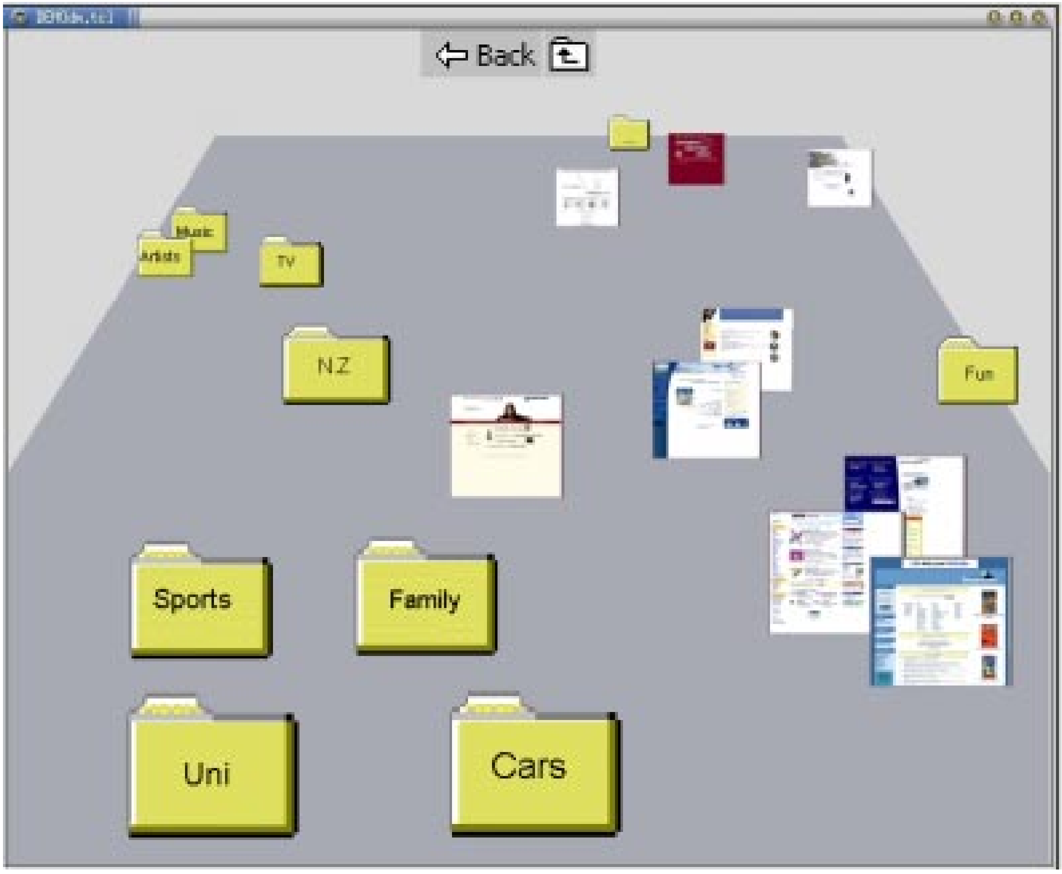

Figure 3.1: A 'drawing' of the proposed hierarchal Data Mountain

3.1 Design Phases

3.1.1 Design Alternatives

Our initial intention was to build a scalable system that would allow users to manage hundreds of documents. We chose to explore two alternatives: The hierarchical Data Mountain and the rotating Data Mountain.

The hierarchical Data Mountain

The 'hierarchical' Data Mountain is a system that looks like the Data Mountain shown in Figure 1.1, but with the notion of grouping, and directories embedded in the system's design. Figure 3.1 shows a 'drawing' of the proposed design. The figure shows how folders may be used to convey grouping of similar items. Users are presented with a Data Mountain that contains folders, each folder representing

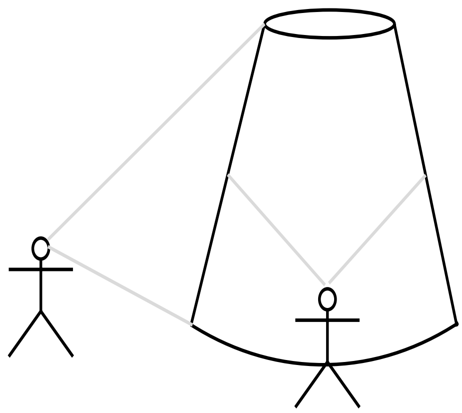

Figure 3.2: Two user views of the rotating mountain

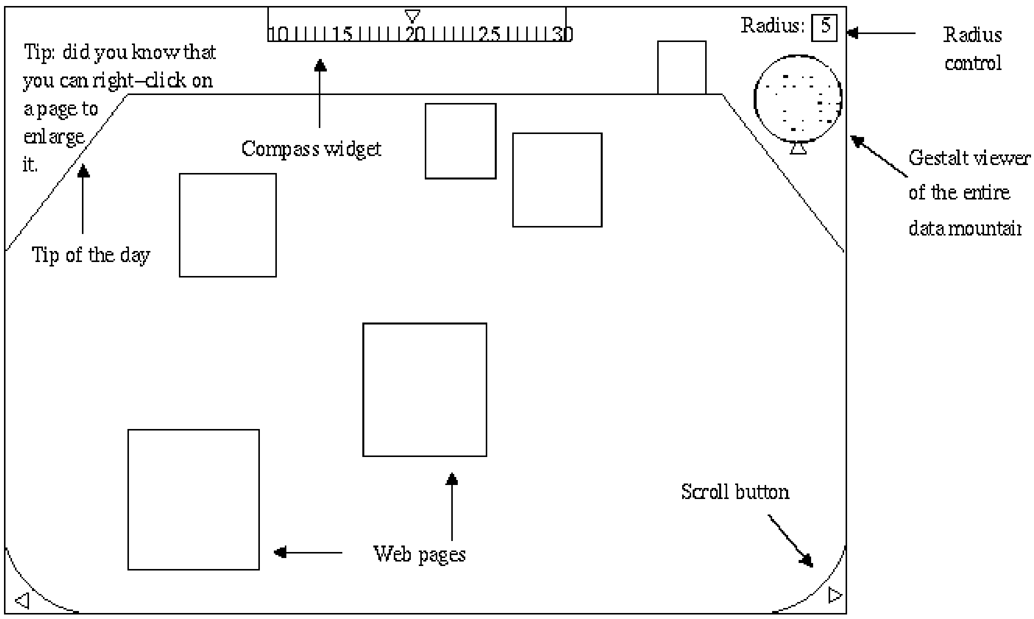

Figure 3.3: A 'drawing' of the proposed rotating Data Mountain

an independent Data Mountain holding related documents. Double-clicking on a folder opens the Data Mountain held by that folder, resulting in a figure similar to Figure 1.1 with some additional controls, an 'up one level' and a 'back' button. The basic operations of the Data Mountain are still supported. Holding down the left mouse button while dragging a thumbnail moves the selected thumbnail. Holding down the right mouse button on the selected thumbnail magnifies it.

Scalability. The system can maintain thousands of pictures with unlimited grouping. The system can also support groups within groups.

Spatial cues. The system allows the placement of thumbnails and folders within the spatial space provided. The system provides spatial cues for each level individually. once a folder is clicked, the position of folders and thumbnails are no longer visible making the spatial cues ineffective.

Visual cues. It was suggested in the design to add background images to the mountain, which would have provided visual cues. Again, clicking on the folders changes the background images distorting the visual cues provided.

Effective Grouping. The folders provide an easy hierarchical structure allowing an arbitrary number of groups.

Ease of navigation. Double clicking on the folders provide an easy method of traversing the hierarchy downwards. The two buttons at the top (see Figure 3.1) allow upwards traversal of the the hierarchy. Although the system provides easy navigation for simple tasks, the hierarchical design provides detail, but does not offer the overall context of the structure which makes performing advanced navigation difficult. For example, moving from one folder located deep within the hierarchy to another can take time and mental effort. In most cases, the user has to navigate all the way up to the root level and then locate the desired folder which may be nested deep within the hierarchy.

Metaphor correspondence. Opening a folder opens another mountain within the current one. This feature violates the metaphor of a real mountain.

The rotating Data Mountain

The second alternative was to make the Data Mountain not just an elevated plane, but to extend the users' mobility, allowing them to rotate around the mountain. The idea of using the third dimension as a means of expanding the available working space was inspired by the research on Cone Trees. Figure 3.2 shows two possible user viewpoints of the mountain. The same basic controls of moving and magnifying thumbnails are provided. Additionally, a 'pan left' and 'pan right' buttons act as the main means of rotating (scrolling) around the mountain. Additional features such as a compass widget (see Figure 3.3) will be added to the top of the screen to indicate the users' relative position and heading, preventing them from losing orientation. A gestalt viewer was also considered to act as a faster way of viewing other sides of the mountain (see Figure 3.3). Initially, our aim was to make the Data Mountain a system capable of handling an unlimited number of documents, but with a limited number of faces, the mountain would eventually become cluttered. To overcome this problem, a radius-control feature has been added to the design, allowing users to increase or decrease the radius of the mountain to suit their personal needs. We also thought of using the additional space to the left of the mountain to provide tips on how to use the system if it were to be commercially deployed.

Scalability. The radius of the mountain controls the supported number of pictures. There is no forced limit on the radius of the mountain. However, for the evaluation purposes, we limited our prototype to 1000 pictures. The radius-control feature was never implemented.

Spatial cues. The rotating behaviour of the mountain may degrade the performance of users as parts of the data space are not visible at any one time.

Visual cues. Background images appearing behind the mountain were proposed. Background images provide visual cues allowing the formation of associations between the thumbnails and their location on the mountain.

Effective Grouping. At this point, the association between the proposed background image and the thumbnails provides a loose grouping mechanism.

Ease of navigation. Scrollbars provide a familiar and easy way of navigating the spatial space provided. The gestalt viewer provides a quick way of rotating around the mountain and it facilitates moving from one group to any other via single click.

Metaphor correspondence. This design corresponds well to the metaphor of the mountain. The mountain is now a full 3D mountain where users can rotate around it.

Enhancing and refining this design was supported by the facts listed above. Also, because this design adapts the photo album metaphor better than the 'hierarchical' alternative.

3.1.2 Implementation and Refinements

The next few sections discuss the initial implementation of the 'rotating' Data Mountain and refinements that were made to the design. The different phases of the design are also discussed.

Initial implementation

In addition to the compass widget, some visual cues were included in the initial implementation, shown in Figure 3.4, to ensure that users maintained their orientation in relation to the mountain. Background images were added to achieve this goal, as these allow users to form associations between the background image and the locations of the different thumbnails. Thumbnails that semantically relate to the background image can be grouped together into the space where the related background is visible. For example, users may place their holiday and travel thumbnails on the side of the mountain that shows a background image of a resort island, beach or sea. This solution added realism to the mountain's general structure and appearance.

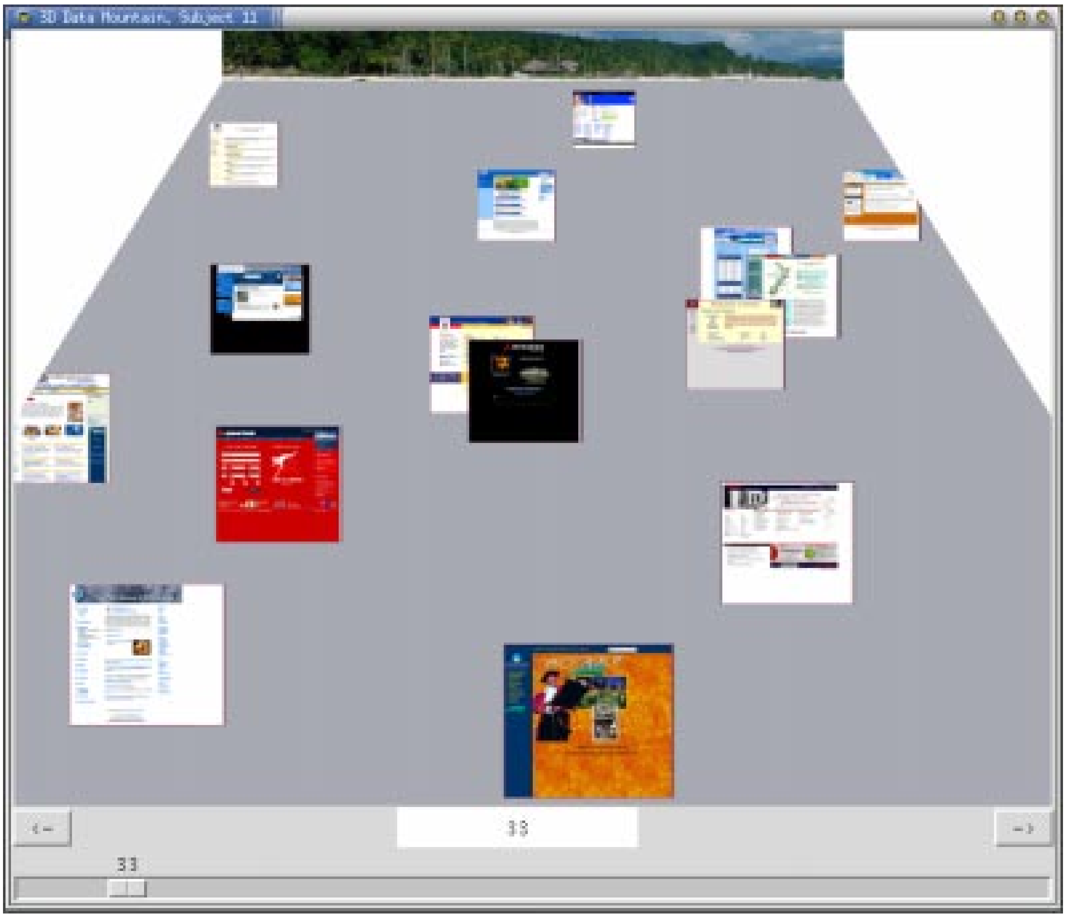

Figure 3.4: Initial implementation

The scrolling buttons were implemented as designed. The scale widget was moved to the bottom in order to group all the controls in one part of the screen. A scrollbar's knob was also included to provide rapid view-positioning. The user moves the knob in the scrollbar to determine the contents of the viewing area. The knob's default position is at the very beginning, or 0 degrees; the maximum position is 359 degrees. The position of the knob in the scrollbar changes as the buttons are clicked to correspond to changes made to the position of the mountain.

At this stage, when the mountain is scrolled, the thumbnails move in the opposite direction, giving the impression that the user is rotating around the mountain. The thumbnails disappear when they reach the edge of the mountain.

Scalability. In terms of scalability, our initial implementation can support around 1000 thumbnails without making the display cluttered. This limit was enforced only for experimental purposes.

Spatial cues. The added background images hints the location of the user in relation to the mountain.

Visual cues. Background images do not offer a strong visual cues because the space allocated to them is small.

Effective Grouping. This was not achieved by this implementation because the background images were not very visible and hence it was difficult to make associations between the background and the thumbnails.

Ease of navigation. A scrollbar acts as the main navigation tool for this interface. The knob of the scrollbar, provide a faster way of rotating around the mountain. However, users have to associate areas of the mountain with their angle. This approach load the user's memory and therefore is not efficient.

Metaphor correspondence. The design corresponds to the mountain metaphor. However, the behaviour of the thumbnails violates the metaphor. The thumbnails disappear suddenly when scrolled to the either edges of the mountain. In real life, objects diminish and disappear gradually into the distance.

More spatial cues

The initial system had a few, easily remedied, limitations. Most, if not all, software packages place the controls at the top of the screen, not the bottom. For consistency we relocated the scrollbar to the top. The compass widget was unnecessary as the angle of viewing shows just above the knob of the scrollbar (see Figure 3.5). The space allocated to the background image was enlarged to improve its effectiveness. We had to choose between expanding the vertical space above the mountain or making the images visible from the sides of the mountain. However, the first alternative would mean that the distance between the controls and the thumbnails on the mountain would increase. This, would have affected the performance of the system as users would have to move their mouse larger distances, which would take more time (Fitts 1954).

When the background images were made visible from above the mountain and from its sides, we found that the mountain became inconspicuous, especially when colourful backgrounds were used. To solve this problem, a border was created around the mountain. To solve the problem of having the thumbnails suddenly disappear, we have used stippling techniques (see Figure 3.5) on both sides of the mountain to make the background semi-transparent. This feature make the thumbnails being scrolled out of the mountain gradually vanish instead of suddenly disappear.

Scalability. These refinements did not affect the scalability of the system.

Spatial cues. Again, the current view is signalled by the background image and the angle of viewing located on top of the knob of the scrollbar.

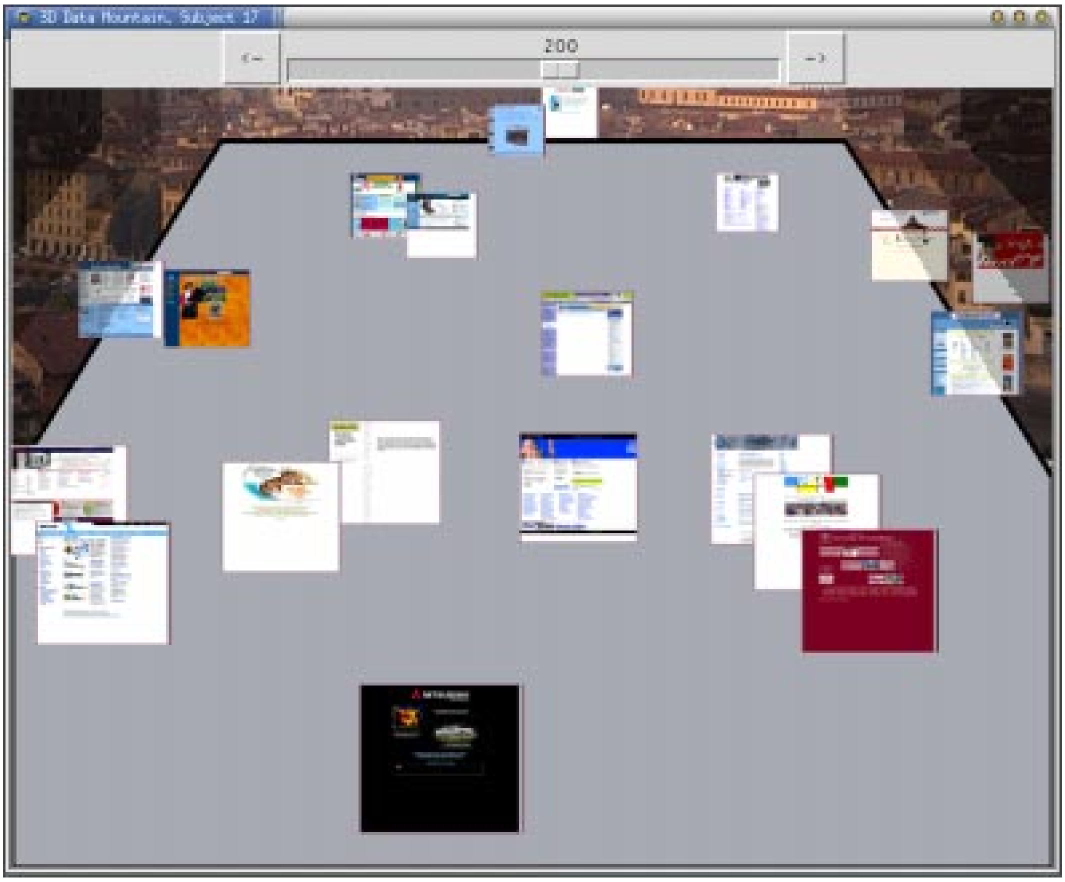

Figure 3.5: Background image of a city appears behind the Data Mountain.

Visual cues. The additional space allocated to the background images was not enough to make effective associations between the thumbnails and the background images.

Effective Grouping. Grouping was enhanced as a result of the enhancements made to the background images but again, the grouping was not effective.

Ease of navigation. Moving the controls to the top made the system look consistent with other systems.

Metaphor correspondence. The thumbnails behaviour was modified to disappear gradually when scrolled towards the edges of the mountain. The current behaviour is close to the mountain metaphor. Now, the thumbnails appear to fly out of the mountain which does not correspond with the mountain metaphor.

Figure 3.6: The panoramic background

3.1.3 Final Prototype

In this section, the last two phases of our prototype design are discussed. We explain the reasons behind the failure of the – theoretically perfect – panoramic background. We include the rationale behind our design decisions.

Panoramic background

As the system developed, further modifications were made. The stippling technique discussed in Section 3.1.2, did not appear to be optimal, as the background images were still not visible enough to enable associations between thumbnails and parts of the background. A proposed solution was to make the mountain semi-transparent. However, we realised that this would complicate the system, and furthermore, a semi-transparent mountain does not fit the metaphor of a real mountain. To solve these problems, a major change had to be made; The grey canvas representing the mountain was removed (see Figure 3.6). This improvement allows the background images to be completely visible. Also, this has increased space allocated working which makes the manipulation of thumbnails more flexible.

A 360 degrees panoramic image was selected to serve as a background image. The initial design of the gestalt viewer, shown in Figure 3.3, did not fit the current implementation of the system. We developed a new design that was directly implemented as shown in Figure 3.6. Since, we were not showing the data in our viewer, we could not call it a gestalt viewer, we called it a 'shortcut viewer'. In our prototype, the shortcut viewer was composed out of ten buttons, each button depicts a 40 degree view of consecutive parts of the mountain. Clicking on any of the buttons takes the user to the view associated with that button. Each button has a scaled-down version of the background image found at that view. It was argued that users may lose their orientation because of the fast transition between one view and the other using the shortcut viewer. However, the knob of the scrollbar and the angle displayed above it give a clear indication of the relative position of the user.

Scalability. Removing the mountain provided more space for the placement of thumbnails. Thumbnails can now be placed anywhere within the rectangular working area, they are no longer restricted to the boundaries of the mountain.

Spatial cues. The panoramic background image provided some spatiality. However, effective associations between the thumbnails and their location could not be made because of the subtlety of the background.

Visual cues. Although the background image was completely visible, it was subtle and hence weakly associated to the thumbnails and their locations.

Effective Grouping. The shortcut viewer provided some sort of grouping, but the subtle background prevented effective grouping to take place.

Ease of navigation. The shortcut viewer provided a quick method to move from one view of the mountain to another via a single click.

Metaphor correspondence. At this stage, the metaphor of the mountain was no longer followed. This fact has made the third-dimension illogical because it does not carry any information. In addition, it makes the pictures at the top of the display smaller and therefore harder to match. However, having the diminishing behaviour of the thumbnails allows for more thumbnails to be displayed at any one time. It was proofed by Cockburn & McKenzie (2001) that the third dimension is generally preferred to its counterparts even if the additional dimension does not carry any information. It was decided to leave the diminishing behaviour of the thumbnail – and therefore the third dimension – as it is despite of its known disadvantages.

Figure 3.7: Final implementation

Final prototype

Although the background image provided a 360 degrees panoramic image, it was subtle (see Figure 3.8), and did not offer any clues as to the contents of each group. Therefore a number of individual images (see Figure 3.8), each representing a different group were chosen (from left to right, trees and gardens, people, cars, houses, malls and shopping, buildings, cathedral and city centre, the beach, scenic views, and trees and gardens again to make the transition from 359 to 0 a smooth one). This arrangement was chosen to suit the thumbnails used in the experiment. In a real system, these groups would be arranged according to the user's preference.

The use of colourful background images made it possible for thumbnails to merge into the background. This problem and the problem of visually similar thumbnails merging together were solved by adding a black border around each thumbnail and treating the background images with a 25% white stippling texture obscuring the background while still making it visible enough to be associated to the thumbnails. Also, having borders around the thumbnails ensures that they do not merge together, or with the background image, solving the problem identified by Rodden et al. (2001) (see Section 2.2).

This new arrangement of background images provides memorable landmarks, solving the problem of the absence of effective landmarks indicated by Robertson et al. (1998) (see Section 2.2).

The shortcut viewer provides a fast 'view-repositioning' technique. In addition, it facilitates the grouping of related thumbnails in one screen identified by the context of the background image (see Figure 3.7).

Scalability. This final refinement did not affect the scalability of the system. The system can handle up to 1000 thumbnails without cluttering the display.

Spatial cues. The categorised background removed some of the spatiality provided by the panoramic background. However, associating the content of the thumbnails with their location hinted by the background, is made easier.

Visual cues. The new background provides visual cues helping the user to form association between the location of the thumbnails and the background image.

Effective Grouping. The new categorised background images provided a natural and effective grouping. The thumbnails can be grouped according to their semantic meaning. The thumbnails' semantic meaning can also be associated with the background's semantic content. This associations results with an effective and memorable grouping scheme.

Ease of navigation. The shortcut viewer's icons provide easily recognisable images, each representing a certain group. This feature provides ease of categorisation and ease of navigation.

Metaphor correspondence. The system does not correspond to the mountain metaphor anymore. However, it corresponds more to the photo album metaphor. Usually, photos are grouped by their content within a photo album. Our system depicts a grouped photo album with each group (page of the album) containing closely related photos.

This system was then evaluated and a number of limitation were discovered. These are discussed in Section 4.4.

3.2 Implementation

The prototype runs under Solaris and Windows 2000 on PCs equipped with 256Mb, with best performance achieved with 512Mb. We realise that this is a large amount of memory for a picture management system. However, the prototype was written for experimental use only at this stage. Our algorithms need to be more efficient both in memory usage and processing speed. All application code was written in Tcl/Tk v8.0.

For the prototype and user study, images were used instead of web pages, as in previous evaluation studies of the Data Mountain. The 400 images used are thumbnails of 24-bit colour photos taken with a digital camera.

(a) Panoramic background image

(b) Categorised background image

Figure 3.8: The difference between the original panoramic image and the categorised replacement one.

Chapter 4

Evaluation study

"If the facts don't fit the theory, change the facts."

—Albert Einstein

The evaluation compared our system, the Enhanced Data Mountain (EDM) with My Pictures (MP), the picture management system available as part of the Windows package since Windows 2000 (see Figure 4.1).

The primary purpose of the experiment was to determine whether there were any differences between the two systems in terms of time to retrieve pictures. We were also interested in whether the spatial memory of users is still effective in a system that cannot be viewed on one screen. Finally, we were interested in the subjective assessment of the two interfaces.

4.1 Experimental Design

The experiment was a paired T-test for the independent variable 'interface type' with two levels (EDM and MP). The same set of 400 thumbnail images, taken with a digital camera, were used for all 20 subjects. The photos, of family, friends and scenic shots of Christchurch, New Zealand, were arranged in the same fashion across the two interfaces. The photos were grouped according to their contents, e.g. photos of people, photos of cars, scenic photos and so on. See Section 3.1.3 for details on the categories used.

Note that, in a real system, the users can control how the photos are grouped; some might choose to group photos by their contents and some might choose to group them by date. For the purposes of this experiment, grouping the photos by date would have caused difficulties because the subjects were not familiar with the photos and did not know when each was taken. It is also possible to sort photos by name. However, this approach seems counter-intuitive because in real life, it is unusual for people to name their photos. In addition, digital camera

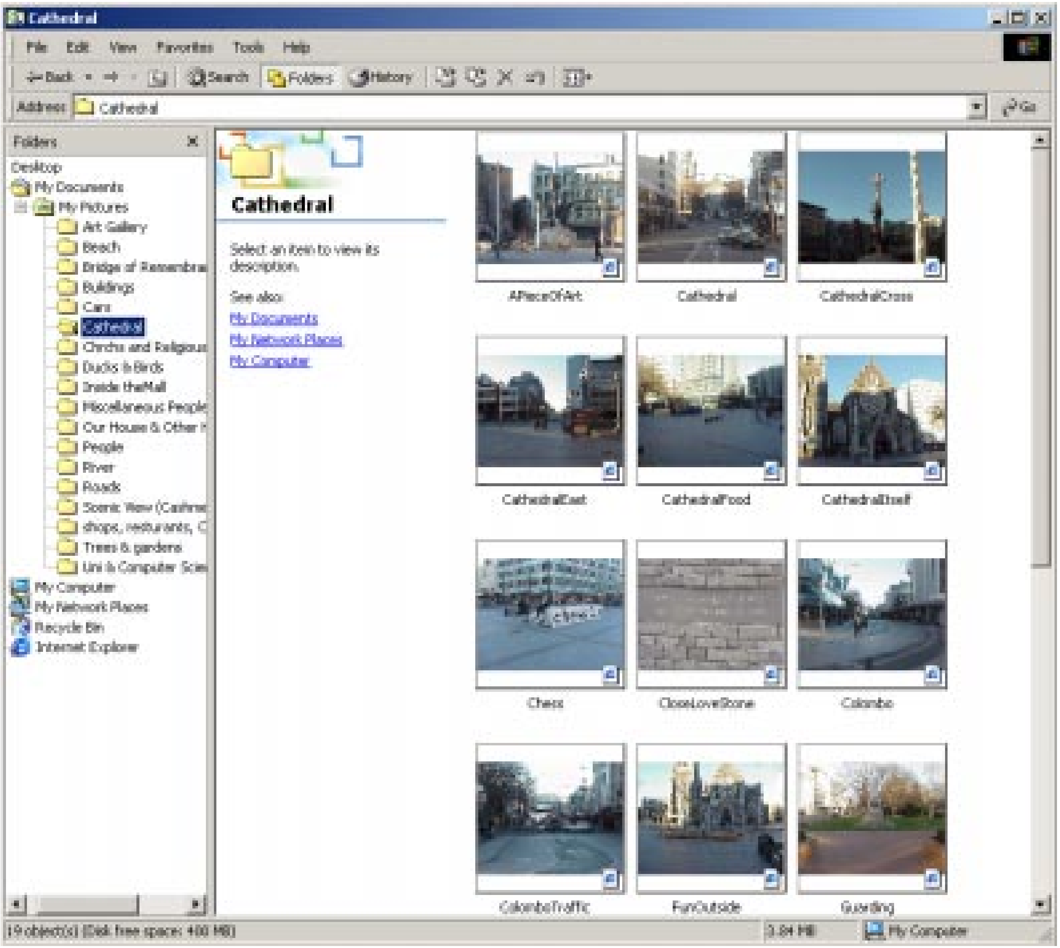

Figure 4.1: My Pictures (MP), Microsoft Windows' picture management system.

Interface type (within subjects)	
Enhanced Data Mountain (EDM)	My Pictures (MP)
S1-S20	S1-S20

Table 4.1: Experimental design. Interface factor within subjects.

generate automatic names for photos which makes renaming photos an optional task. Grouping the photos by their contents seemed to be a feasible grouping method especially as we allow some time for the subjects to study the grouping method and the groups used.

Evaluation procedure

For evaluation purposes, we disabled the ability to move thumbnails, as we wanted to test the same arrangement of images for all subjects. The evaluation procedure consisted of retrieving twelve random photos from each interface. The subjects were introduced to the features of each interface and then were given as much time as needed to study the arrangement of pictures, the location of each group, and in case of MP, the name of the group. The time was recorded and was called the 'time of browsing'. The users were then given at least three sample tasks to ensure their understanding of the procedure of the experiment. More sample tasks were given if the users did not feel confident with the system. Upon completion of the sample tasks, the subjects responded to one 5-point Likert scale question: Q1 "I will be able to quickly find the photos required". After answering Q1 the retrieval task started.

The retrieval task consisted of finding, as quickly as possible, twelve randomly selected photos from the display, one at a time. The users were shown the magnified image and its title in a separate window (see Figure 4.2). In the EDM system, software timed each task, with the clock stopping when the subject magnified the targeted photo (mouse right-click), while in the MP system, a stop-watch was used to time the tasks. To be consistent, the clock was stopped when the subject right-clicked the picture. After the retrieval task, subjects responded to two more Likert scale questions: Q2 "I was able to quickly find photos", and Q3 "Overall the interface is effective". The subjects then performed a further twelve tasks on the other system and were asked the same three Likert scale questions. On completion of all the tasks on the second interface the subjects were asked a final preference question; Q4 "Which of the two interfaces did you prefer?". Each evaluation session lasted approximately 40 minutes, including the time of browsing and training.

Learning effect control

We were aware that subjects might learn the location and grouping of picture thumbnails between the EDM and MP interfaces, therefore, we controlled the learning effect by having even numbered users carry out the experiment on the EDM and then the MP system, while odd numbered subjects used the MP system first, followed by the EDM.

Initially, we thought that the ultimate control of the learning effect would be to ask users to perform a single task on the EDM system, followed by a single task on the MP system, then repeat the process. However, this would have been potentially confusing for the subjects, and could have degraded their performance, as this would not have allowed them to be sufficiently familiar with either of the two systems.

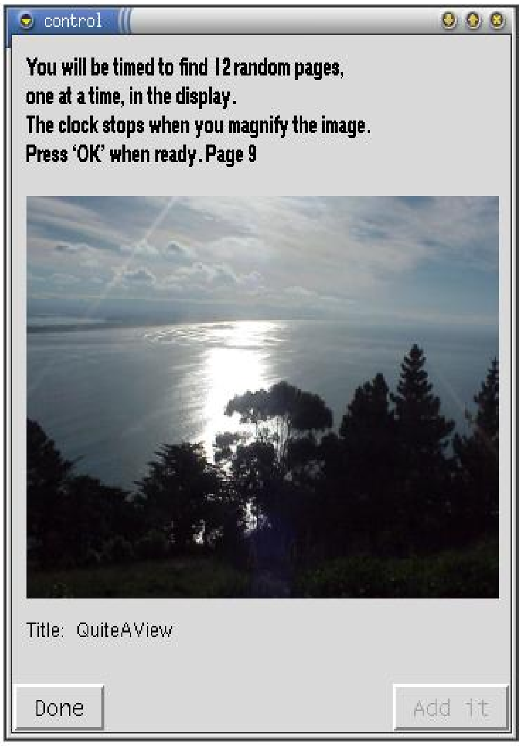

Figure 4.2: The control window showing one of the retrieval tasks

Because we automated and randomised the tasks, it was possible that subjects may have been required to perform the same task on each interface. Therefore, we arranged the tasks into two sets. The first ten subjects were asked questions from the first set on the EDM system and the second set on the MP system. The sets' assignment was reversed for the next ten subjects.

Subjects

The twenty subjects who participated in the evaluation were volunteer postgraduate (14), undergraduate (6) computer science students.

4.2 Results

Prior to the evaluation, we expected that in terms of efficiency, the EDM system will be significantly better. We also expected that EDM will be preferred to the MP system.

4.2.1 Image Revival Task

The mean times for the 'browsing' task in the EDM and MP systems were 515.5 (s.d 224.2) and 515.2 (s.d 180.7). This difference is not significantly reliable (Paired T-test, T(19) = 0.01, p = 0.99). The mean times for retrieving twelve images in the EDM and MP interfaces were 27.28 (s.d 31.69) and 17.65 (s.d 16.58) which gives a non-significant difference (Paired T-test, T(19) = 0.93, p = 0.87).

The subjects' comments reveal some of the reasons for no apparent difference between the two systems. With the EDM system, users spent a small amount of time identifying the group, and spent the majority of their time visually matching the images. Part of the problem is that the thumbnails were too small and hence difficult to visually distinguish and match. Ten subjects, out of the twenty participated in the experiment, mentioned that when searching, pictures are easily missed because they are too small to recognise. The opposite happened with the MP system; users spent most of their time determining the group, and most quickly located the desired thumbnail by matching its label. Most of the comments about the MP system indicated that the name of the folder does not effectively indicate the contents of the folder. Another common comment was that searching by name is a major advantage.

Cross-categorised photos also affected the efficiency of both systems. Some photos fit more than one category. For example, a picture of a garden on a hill may be grouped under the group 'trees and gardens' and also under 'scenic views and hills'. Problems such as this affected the time it took users to identify the group to which a picture belonged.

Thirteen subjects commented on the absence of titles over the thumbnails in the EDM system. We believe that this degraded EDM's performance because users could only visually match photos. The issue of visual matching and label matching is discussed in Section 4.4.

4.2.2 Browsing versus Retrieval

During the evaluation, we noticed that subjects who spent longer browsing through the images performed better than those who did not. This observation lead to analysing the data to determine whether there was a linear correlation between the time of browsing and the average time of retrieval. A strong correlation was found for both systems: for the EDM system a linear correlation (F(1,18)=172.9, R^2=0.91) and for the MP system a linear correlation (F(1, 18)=266.25, R^2=0.93). Figures 4.3 clearly shows, that for both interfaces, increasing the 'time to browse' significantly decreased the average retrieval time. These results strongly suggests that if the users had arranged the thumbnails themselves, the retrieval times would have been significantly faster. Trumbo (1998) suggests that activities that are rehearsed are moved to our long-term memory, while elements that are not, are often forgotten. We believe that this is an adequate explanation for our results.

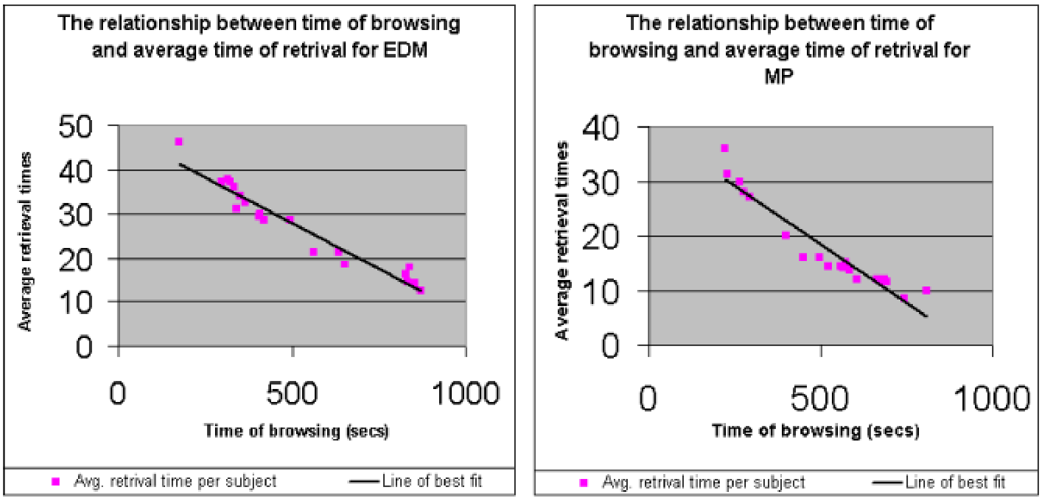

Figure 4.3: The correlation between the time of browsing and average time of retrieval for EDM(left) and MP (right).

Question	**EDM**	**MP**	**Reliably different?**
Q1. I will be able to quickly find the photos required.	3.55 (0.83)	3.85 (0.99)	No
Q2. I was able to quickly find the photos.	3 (0.79)	3.25 (1.07)	No
Q3. Overall the interface is effective.	3.4 (0.68)	3.65 (1.18)	No

Table 4.2: Mean (standard deviation) responses to 5-point Likert scale questions. 'No', not significant at the .05 level.

4.2.3 Subjective Measures

Although none of the first three Likert scale questions revealed significant differences between the two interfaces (Mann-Whitney U tests) (see Table 4.2.3), there was a significant difference in the subjects' answers to Q4 – the overall interface preference question. Subjective responses to Q4 showed a higher preference for the MP interface. Sixteen subjects preferred MP, two preferred EDM while two thought both interfaces were equally preferable. The difference in rating is significant (Chi-square test $\chi^2 = 4.5, df = 1, p < 0.05$). Many of the users commented that because of the deterministic nature of the tasks, a deterministic – search by name – method is preferable. Many users mentioned that "EDM would be good for browsing rather than searching".

4.2.4 Expert Timing

A very simple test to see how fast an expert, who arranged the thumbnails, could use both systems was undertaken. The author of this report performed 30 retrieval tasks as quickly as possible on both systems. The mean (standard deviation) times for retrieving 30 images using the EDM and MP interfaces were 7.67 (6.59) and 9.62 (9.3). The means and standard deviations for the two interfaces were lower and the EDM times were lower than those of the MP system suggesting potential for EDM.

4.3 Confounding Factors

There were several possible confounding factors in our experiment.

Self-organisation

The major issue that has affected both systems was that the subjects were presented with a forced grouping that they were not allowed to change. we speculate that allowing users to place and group thumbnails might have altered the results of the evaluation. We presume that this degraded the performance of both interfaces similarly. Section 4.2.4 clearly shows that the performance times of both systems improves when used by someone who arranged the thumbnails.

Deterministic tasks

Another confounding factor is that the retrieval tasks focused on retrieval of specific pictures, that is to say, the tasks were in the form of “find this certain picture” instead of general browsing question of the form of “find a picture of a car”. EDM appears to have greater potential in the area of browsing than specific retrieval.

Two machines

The evaluation was separated into two parts: the EDM part and the MP part. The EDM system running on a Solaris Unix machine, while the MP system ran on a Windows 2000 machine. Subjects had to move from one machine to the other in the middle of the experiment. However, we believe that this did not affect the results because we gave the subjects the time needed to focus and browse through the images, and we provided the sample tasks to ensure that the subjects were fully comfortable with the systems.

EDM's memory use

Because of the number of pictures dealt with, the EDM system used an enormous amount of memory. This made the system swap some of the images in memory onto the hard disk. However, in our implementation, we made sure that all the background images and the visible images were always in memory, ensuring that scrolling would be done in real-time. However, the enlarged thumbnails – appearing on right clicking – were not necessarily in memory. This affected the time it took to enlarge a thumbnail. However, to ensure that we logged the correct times, the software recorded the time at which a thumbnail was right-clicked before attempting to enlarge it. This way, the time it took to find a thumbnail was not affected by the time it took the computer to find the image on disk, move it to memory and then show it.

Subject pool

Our subjects (computer science students, and graduates), are more familiar with computers than most computer users. However, we believe this is not likely to have affected the performance of either systems.

4.4 Discussion and Further Work

To summarise, the mean task completion times for the retrieval tasks were higher when using the EDM system than when using the MP system. These differences, however, were not statistically significant. The mean task completion times for the expert retrieval tasks were lower when using the EDM system. Again, these differences were not statistically significant. However, these results show that the EDM has potential as a mainstream picture management system.

It is plausible that any performance enhancement the EDM interface provided was counter-balanced by the absence of titles for the images, which forced the subjects to visually match them. Supplying a title as well as the thumbnail would boost the system's performance. Further work to eliminate this problem could include excentric labelling to add titles to the thumbnails. Excentric labelling is a technique that provides dynamic labelling for any form of data. When the cursor stays idle for a certain period of time – normally, one second – excentric labels appear and follow the cursor. Shift-clicking freezes the labels, allowing the user to select an individual object by either clicking on the object itself or on its frozen label. Excentric labelling can also be useful when there are thumbnails partially or completely obscuring other thumbnails. (Figure 4.4 shows a sketch of how excentric labelling may benefit our prototype.)

A 'sketch' of another proposed design for incorporating labels in the EDM system is shown in Figure 4.5. This design combines both the benefits of visual cues and

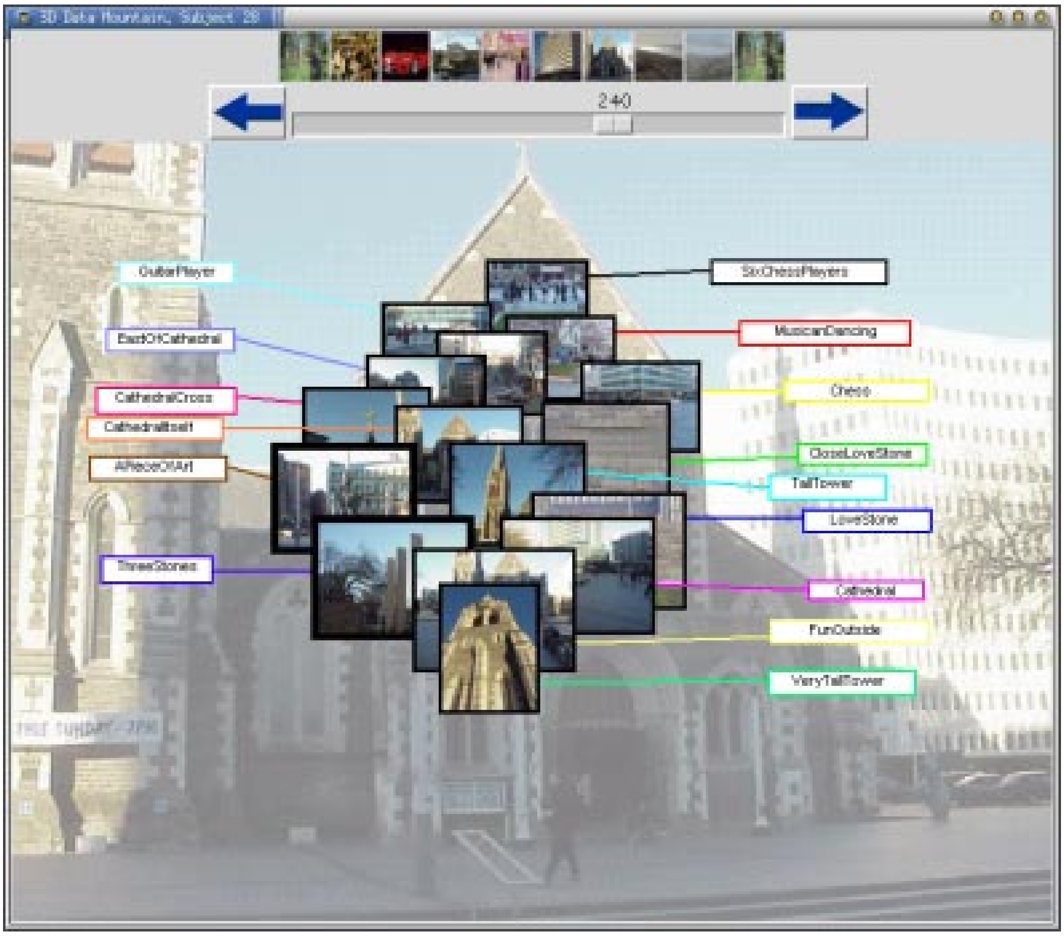

Figure 4.4: A sketch of how excentric labelling can possibly enhance EDM

label matching. On the right side of the working space, a dynamic alphabetical list of the visible thumbanils' labels is stored. Clicking on a title highlights the border of the corresponding thumbnail and vice versa. This design is very likely to enhance the performance of EDM and eliminate the problems of visual matching that was commented on by many users.

An idea which would eliminate the problem of cluttering, is to make the scrolling vertical as well as horizontal. In this case, the diminishing behaviour of the thumbnails (three-dimensionality) would be disabled. This approach would give greater flexibility and more scalability to the system.

The mean task completion times for the browsing tasks were not significantly different. However, subjective comments revealed that EDM is preferred for browsing tasks. The subjective satisfaction measures showed that there is no significant difference in the speed or effectiveness of the two interfaces. However, the final

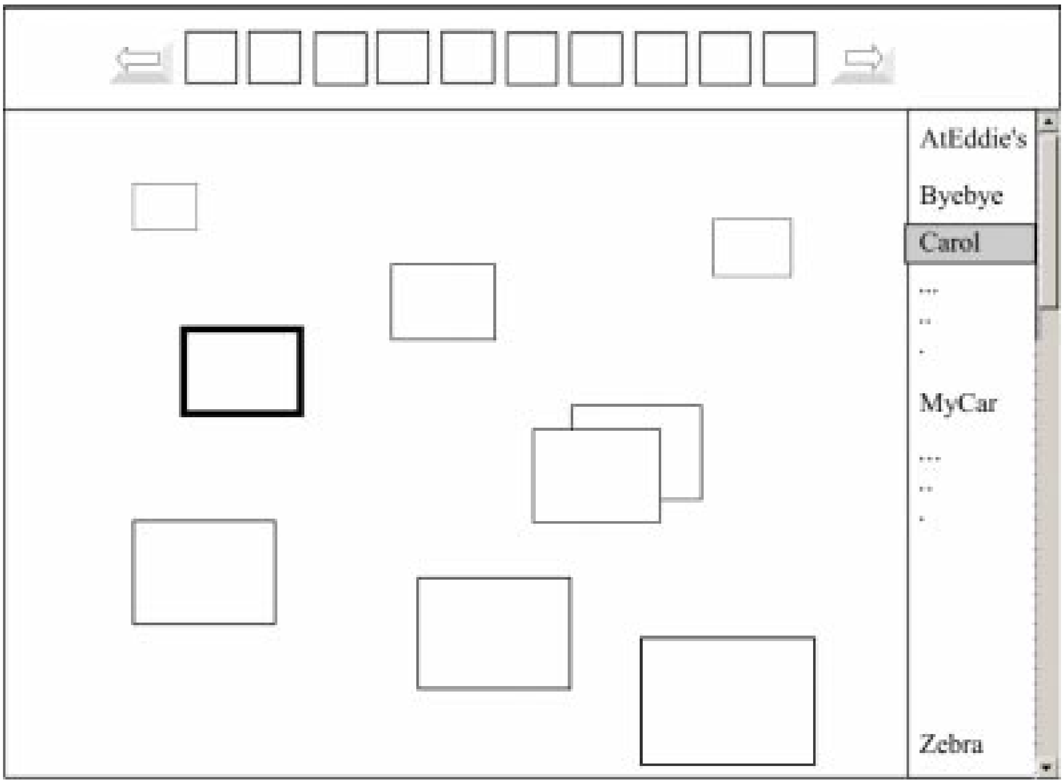

Figure 4.5: Combining visual and label matching

overall preference question revealed a significant preference of the MP system.

After the evaluation started, it was discovered that the scrolling behaviour implemented violates the natural metaphor of 3D objects. In order to give the illusion of perspective, objects 'closer' (at the bottom of the screen) should move horizontally faster than 'far' ones (at the top of the screen). This limitation did not appear to affect the performance of our system. However, we believe that it might have made the system slightly harder to learn and accept.

In all, we conclude that using spatial cues did not significantly enhance the performance, although reliance on visual matching may have counter-balanced the performance gains resultant from spatial cues. This is an area for further research. The effect of spatial cues in different systems also needs to be evaluated. Another possible area of study is evaluating how visual matching compares to simple text matching. There are still a number of unanswered questions in the area of spatial cues and document management that deserves some attention.

Chapter 5

Conclusion

"I think and think for months and years. Ninety-nine times, the conclusion is false. The hundredth time I am right."

—Albert Einstein

Document management is a ubiquitous activity. From text files to rich images, people spend a large portion of their time on document management. Research has been undertaken to incorporate the power of spatial memory and visual cues within document management systems. Evaluations of systems that exploit spatial memory showed that spatial memory is a promising cue to organisation of items and especially to rapid retrieval. In order to preserve spatial cues, these systems deal with a limited number of items.

Our research concentrated on building a scalable document management system exploiting spatial memory. Our prototype was based on photo retrieval. However, our results are generalised to all document types. We performed an evaluation that compared our system to a state-of-the-art picture management system – Microsofts' 'My Pictures' (MP). Several subjects commented that EDM would be preferred for browsing tasks rather than deterministic tasks, which shows that visual cues do aid such tasks. However, no significant performance difference between EDM and MP were found in terms of the retrieval tasks. In general, there was a significant overall preference for the MP system. The results also showed that an expert user performed better using EDM. The results of our evaluation were inconclusive. We were not able to make a firm statement about the effectiveness of spatial memory in scalable environments.

Future work includes investigating the conditions under which spatial cues for document management provide performance enhancements. Performance differences between visual-matching search methods and text-based search ones also warrant investigation

Acknowledgements

Since I have a chance to say ‘thanks’, I really do intend to use it. First, GOD, I feel your presence everyday, you are such an awesome GOD. My parents, without your support, your understanding and your sacrifices, I would not be what I am now. Carol, you are an elemental drive of my existence, I love you. Dr. Andy Cockburn, thank you for your time, your constant inspiration and for being the most supportive, friendliest, down-to-earth, and coolest supervisor. Thanks to Jane McKenzie for her effort in reviewing this document, without her you would not have understood anything I wrote. My thanks goes to all the subjects who took part in the experiment. Finally, all of you who have helped me in one way or another, I'm truly thankful, you are the best.

Bibliography

Arnheim, R. (1954), *Art and Visual Perception: a psychology of the creative eye*, Berkeley: University of California Press.

Ballay, J. (1994), Designing workscape: an interdisciplinary experience, *in* 'Proceedings of CHI'94 Conference on Human Factors in Computing Systems, Boston, April 24–28', ACM Press, New York, USA, pp. 10–15.

Chen, C. & Czerwinski, M. (1997), 'Spatial ability and visual navigation: An empirical study', *New Review of Hypermedia and Multimedia* **3**, 67–89.

Cockburn, A. & McKenzie, B. (2001), 3D or Not 3D? Evaluating the Effect of the Third Dimension in a Document Management System, *in* 'Proceedings of CHI'2001 Conference on Human Factors in Computing Systems, Seattle, Washington, March 31–April 6', pp. 434–441.

Curiel, J. & Radvansky, G. (1998), 'Mental Organization of Maps', *Journal of Experimental Psychology: Learning, Memory and Cognition* **24**(1), 202–214.

Czerwinski, M., Dumais, S., Robertson, G., Dziadosz, S., Tiernan, S. & Van Dantzich, M. (1999), Visualizing Implicit Queries For Information Management, *in* 'Proceedings of CHI'99 Conference on Human Factors in Computing Systems, Pittsburgh, May 15–20', pp. 560–567.

Donelson, W. (1978), Spatial management of information, *in* 'Proceedings of the Fifth Annual Conference on Computer Graphics and Interactive Techniques', ACM Press, New York, USA, pp. 203–209.

Dumais, S. & Jones, W. (1985), A comparison of symbolic and spatial filing, *in* 'Human Factors in Computing Systems II. Proceedings of the CHI'85 Conference.', Amsterdam; North-Holland/ACM.

Ekstrom, R., French, J. & Harman, H. (1979), 'Cognitive factors: Their identification and replication', *Multivariate Behavioral Research Monographs* **79**(2), 1–84.

Fitts, P. M. (1954), The information capacity of the human motor system in controlling the amplitude of movement., *in* 'Journal of Experimental Psychology', pp. 381–391.

Mackinlay, J. D., Robertson, G. & Card, S. K. (1990), Rapid controlled movement through virtual 3D workspaces, *in* 'Conference proceedings on Computer graphics', pp. 171–176.

Miller, G. A. (1956), 'The magical number seven, plus or minus two: Some limits on our capacity for processing information', *The Psychological Review* **63**, 81–97. URL: http://www.well.com/user/smalin/miller.html.

Pierce, J. S., Conway, M., Van Dantzich, M. & Roberstson, G. (1999), Toolspaces And Glances: Storing, Accessing, and Retrieving Objects in 3D Desktop Applications, *in* 'Proceedings of the 1999 symposium on Interactive 3D graphics', pp. 163–168.

Robertson, G., Card, S. K. & Mackinlay, J. D. (1991), Information Visualization Using 3D Interactive Animation, *in* 'Human factors in computing systems conference proceedings on Reaching through technology', pp. 461–462.

Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D. & van Dantzich, M. (1998), Data mountain: Using spatial memory for document management, *in* 'Proceedings of the 1998 ACM Conference on User Interface Software and Technology, November 1–4. San Francisco, California.', ACM Press, pp. 153–162.

Robertson, G., Mackinlay, J. & Card, S. (1991), Cone trees: animated 3D visualizations of hierarchical information, *in* 'Proceedings of CHI'91 Conference on Human Factors in Computing Systems, New Orleans, May', pp. 189–194.

Robertson, G., Tesle, J., Tversky, B. & Mullet, K. (1991), 3D or not 3D: More is better or less is more?, *in* 'Proceedings of CHI'95 Conference on Human Factors in Computing Systems, Denver, May 7–11', pp. 174–175.

Robertson, G., van Dantzich, M., Czerwinski, M., Hinckley, K., Thiel, D., Robbins, D., Risden, K. & Gorokhovsky, V. (2000), The task gallery: A 3D window manager, *in* 'Proceedings of CHI'2000 Conference on Human Factors in Computing Systems, The Hague, The Netherlands, April 1–6', pp. 494–501.

Rodden, K., Basalaj, W., Sinclair, D. & Wood, K. (2001), Does organisation by similarity assist image browsing?, *in* 'Proceedings of CHI'2001 Conference on Human Factors in Computing Systems, Seattle, Washington, March 31–April 6', pp. 190–197.

Roskos-Ewoldsen, B., McNamara, T., Shelton, A. & Carr, W. (1998), 'Mental Represtations of Large and Small Spatial Layouts Are Orientation Dependent', *Journal of Experimental Psychology: Learning, Memory and Cognition* **24**(1), 215–226.

Trumbo, J. (1998), Spatial memory and design: a conceptual approach to the creation of navigable space in multimedia design, *in* 'Interactions', ACM Press, New York, USA, pp. 26–349.

Volbracht, S., Domik, G., Shahrbabaki, K. & Fels, G. (1997), How Effective are 3D Display Modes?, *in* 'Proceedings of the ACM SIGCHI'97 Conference on Human Factors in Computing Systems, Atlanta, Georgia, March 22-27', pp. 540–541.