

The BubbleFish: Bringing Digital Document Management into the Physical World

Graduate Course Examination paper: "Trends in HCI"

Daniel Fallman
Department of Informatics, Umeå University
The Interactive Institute, Tools for Creativity Studio
dfallman@informatik.umu.se

Abstract

This paper introduces the BubbleFish project, in terms of the main idea, related work, and a set of possible prototypical implementations. The BubbleFish is a digital document manager that uses Augmented Reality to let its users maintain digital documents in the physical environment, as an alternative to the currently dominating file/folder approach in desktop based computer systems. The user moves documents in and out of a virtual stage, currently the computer screen, and is able to freely position the documents spatially, e.g. on the desk, in a shelf or a drawer. The framework of interaction provided follows the principles of direct manipulation, but takes the issue of directness at the interface further by providing an interaction space that blends the virtual and the physical world.

Introduction

This paper will introduce the BubbleFish project. In short, I will present the BubbleFish concept for how to allow the desktop computer user to manage his or her documents in the physical space surrounding the computer. Management of all digital documents—be it word-processing documents, pictures, or organizational budgets—is currently restricted to putting files (i.e. documents) into folders (i.e. hierarchical separators); a hierarchical tree-like system promoted and implemented in basically all available operating systems. Digital document management was chosen as it should be considered as one of the most frequent everyday concerns of computer use: how to store, retrieve and maintain digital documents. As mentioned, in today's computer systems documents are stored as files within folder hierarchies. The problem is that we often find it intricate to organize digital documents using this approach. Rekimoto (1999) argues that it is difficult to correctly classify documents into specific folders as some documents may belong to two or more categories, or insufficiently match a user's subjective folder system. For the same reasons, finding a document is also a difficult task, which is especially true for digitized data such as images, audio and video where there is currently no apparent mechanism available to perform searches on content (Rekimoto, 1999).

The starting point for this project was hence to question the currently applied view of how digital documents should be managed, and specifically find a new way of allowing users to manage digital documents in desktop computing. The paper will also consider related work in

the field of Human—Computer Interaction (HCI), and present three possible and simple implementations of the concept, each with their individual pros and cons.

Paper Context and Course Themes

The context of this project, which is currently expanded upon and run as a research project at Interactive Institute's Umeå Studio, is the graduate course "*Trends in Human—Computer Interaction*", given by Prof. John Waterworth in the Department of Informatics at Umeå University during the spring of 2000. During this course, three main themes (i.e. trends in HCI) were introduced through lectures, exemplified by papers in a comprehensive course reader, and discussed thoroughly by the class. To a large extent, this work has its origins in the course activities and in these themes. The themes or trends considered during the course were: *synaesthetic media*; *using 3D space*; and *Intimate Systems and Physical—Virtual World Bridges*. Specifically, in the BubbleFish project, I draw extensively on the trend presented as "Using 3D space" and "Physical—Virtual World Bridges".

Using 3D Space

The trend of "using 3D space" was discussed mainly in terms of how it is possible to extend the traditional 2D space of the computer screen with a third dimension (though often still presented with a two-dimensional output device, the computer screen), and what new interaction and information visualization challenges we then encounter. The use of three dimensions was also suggested as possibly beneficial for allowing "navigation of information", that is to some extent to draw on human skills of navigation in the physical world, which of course is three-dimensional. As an example, in the reader, Small (????) presented an idea of how to manage large amounts of text, drawing extensively on human navigational skills (but for a different purpose) using a 3D space. Other interesting material in the reader regarding the use of 3D space, which are also interesting to this project as they too present novel ways of interacting with information, were papers by Card et al (1991), which presented a 3D visualization of information (*The Information Visualizer*) primarily according to cost structure of information and searches, and Robertson's (1991) *Cone Trees* (see *Figure 1*), which extends the file/folder dichotomy by making the tree structure graspable and visualized in a 3D space, and the Perspective Wall by Mackinlay et al (1991), which presents 2D layouts in a 3D way that is thought to imitate the human eye to integrate detail and context.

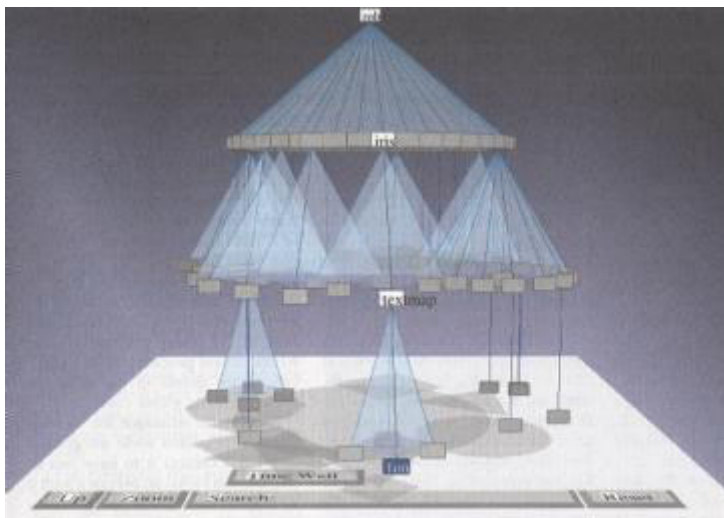


Figure 1. *Cone Trees* (Robertson, 1991)

Physical—Virtual World Bridges

The second theme the BubbleFish project focuses on is that of “Physical—Virtual World Bridges”, which is also capitalized on in the BubbleFish project. In the reading, this theme was particularly approached through the notion of augmented reality, where a computer generated picture is juxtaposed on the user’s experience of the physical world, so that the computer generated materials (typically images and sound) for the user tend to blend with the physical world. The notion of augmented reality was introduced by Ranson et al (1996) through the rapid scout project, Vallino’s (1998) paper on the potentials and challenges of augmented reality from the perspective of philosophy, and its similarities and differences to Virtual Reality (VR). The coming together of physical and virtual was also introduced by the notion of affective computing, where sensors for the physical world, here typically the user’s body, is used to inform virtual activities (Picard, 1997), and through the concept of implants of computer chips into the physical body (Warwick, 2000). The BubbleFish project especially draws on the notion of augmented reality to provide then blend between the physical and the virtual necessary to allow users to physically interact with virtual information. A possible alternative notion of how this could be achieved is through the notion of tangible computing, e.g. as introduced and exemplified by Ishii et al (1997).

Related Work

As have been argued, when using today’s file/folder based systems, it is difficult to correctly classify and hence be able to place a specific documents into a specific folder, as many documents belong to two or more categories. We have also said that, for the same reasons, finding a document is equally difficult, and especially so when digitized data such as images, audio, and video is concerned.

At large, two explanations to these problems have been explored in related work: first, that spatial organization and overview is poorly supported, and second that the issue of time is not properly conveyed. To overcome the first issue, Card et al. (1996) propose a 3D-based spatial information retrieval system, called the WebBook and the Web Forager, which together forms a work space primarily designed to give users rapid access to and comprehensible management of a large number of web pages (see *Figure 2*). Tesler & Strasnick (1992) presents a 3D based file hierarchy landscape (see *Figure 3*), where the user travels in 3D between different folders, providing a visual clue to navigation. Here files sit on top of pedestals, while subdirectories recede into the background (which is also worth mentioning because this system featured in the movie Jurassic Park). A similar system is the previously introduced *Cone Trees* by Robertson (1991), *Figure 1*, but here the user is not presented with a 3D landscape, i.e. in this sense a metaphor of a physical landscape where there is a horizon, a sky, and a horizontal plane on which is it thought to be “natural” to place objects. Rather, information is presented in a 3D space with no up and down, more of a being-in-space metaphor than those that rely on that of a landscape. Here, as there is no “natural” horizontal plane on which to present information, the application can take more advantage of the three dimensions, e.g. in terms of fitting more information in less space, but this benefit may come at expense of not having the set of “shared understandings” that may exist between the system and the user by use of the landscape metaphor (i.e. that there is an up and a down, left and right, and so on).

To account for the second issue, that time is not properly conveyed, Freeman et al. (1996) instead argue for a one dimensional system, called *Lifestreams*, where documents are ordered

by time, and where time hence becomes the replacement for files and folders (see *Figure 4*). Kullberg (1996) makes use of a traditional timeline within an abstract, virtual 3D space for presentation of historical information.



Figure 2. *The Web Forager* (Card et al, 1996)

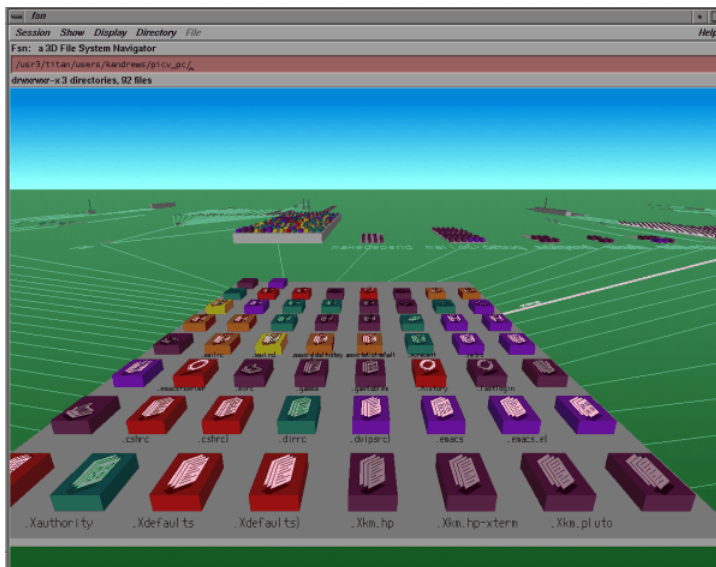


Figure 3. *File System Navigator* (Tesler and Strasnick, 1992)

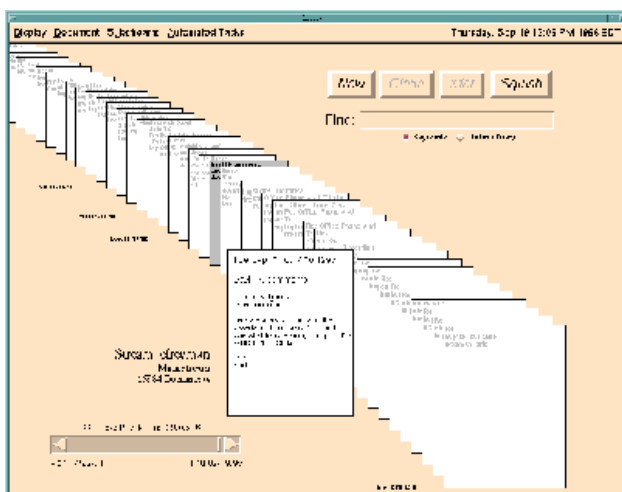


Figure 4. *Lifestreams* (Freeman et al, 1996)



Figure 5. *Dynamic Timelines* (Kullberg, 1996)

Rekimoto (1999) considers these two approaches as complementary, providing a design entitled *TimeScape* that combines a spatial desktop metaphor with chronological navigation in time (see *Figure 6*). Here, the user can spatially arrange information, e.g. documents, on the computer desktop. The system keeps track of all changes, so any document can be moved or deleted at any time, but still retrieved by travelling back in time. Rekimoto argues that the combination of being able to spatially arrange information on the desktop, and the possibility of travelling back and forth in time free the user from the troubles of the file/folder approach.

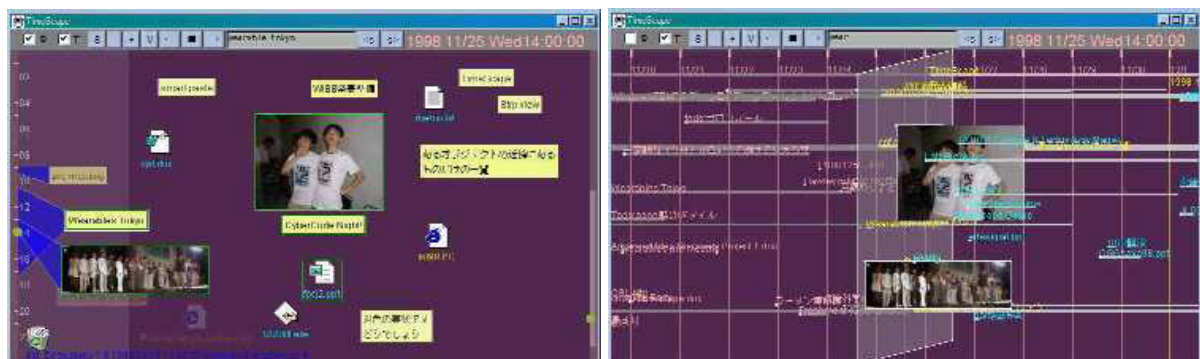


Figure 6: *TimeScape* (Rekimoto, 1999)

The Separation of Physical and Virtual Spaces

When working with the BubbleFish project, I recognize and draw extensively on these previous efforts, but I also intend to establish a third explanation: the separation of physical and virtual spaces. Currently, the user has to organize digital documents into hierarchical folders, while at the same time manage physical documents such as papers, books, print-outs and Post-IT notes, which are often spatially organized (Malone, 1983). In today's offices, these are two very separate activities that require different approaches. The purpose of the BubbleFish is to give physical shape and location in the room to digital documents, which allow the users to organize them using the same method as physical documents are maintained.

The BubbleFish Document Manager

By use of Augmented Reality, the BubbleFish lets its users manipulate and move virtual representations of documents in and out of a computer space—the virtual stage—where digital documents are manipulated by applications, and allows them to be placed somewhere in the room. In the current prototype a computer screen represents the stage, but through AR it could also be entirely virtually defined.

Representations of digital documents appear as bubbles, and are positioned spatially by the user in 3D somewhere in the room surrounding the virtual stage, on the user's desk, on a shelf, beneath the ceiling, to the left of the desk light, on top of a pile of other bubbles etc. The user is able to directly manipulate bubbles in the physical space, and to arrange them according to own preference in a way similar to how the user would handle physical documents, for instance according to document category, active tasks and frequency of use. The act of dragging a document out of the virtual stage by hand and placing it in the room equals the save command. When the user wishes to put a stored document to use, to open the document, the corresponding bubble is grabbed from somewhere in the room, moved and released in front of the virtual stage.

Figure 7 is an early visual and conceptual illustration of how the BubbleFish document manager would introduce a number of virtual but seemingly physical objects into the user's physical work place.



Figure 7: An illustration of the BubbleFish implemented and in operation within a work place

Defined Spaces

Users of the BubbleFish may also configure defined spaces in the room for special purposes. For instance, a physical trashcan may also be used to obliterate bubbles, and a physical outbox might become the defined space to place emails to be sent. The office window could be defined as a shared space, to share bubbles with other users.

Different Views

The use of augmented reality (AR) to give physical shape to digital information for the purpose of managing documents seems more appropriate than having actual physical objects to represent the documents, which would be an alternative approach. The reason for this design choice is that making objects appear as physical, rather than actually being physical, easily allows for different views on the stored documents.

In the aquarium view, the default mode, the spatial location of the bubbles is completely controlled by the user. While the current prototype does not support additional views, future versions of the BubbleFish would benefit from e.g. a search view, in which the bubbles may be reconfigured by the system according to search results and visualized in different ways. The user should also be able to define task views, in which certain bubbles are reconfigured according to the needs of a specific task. The time view would reconfigure the bubbles in terms of their age.

Framework of Interaction

The document manager is one of many possible implementations of the framework of interaction provided by the BubbleFish, which is consistent with the design principles of direct manipulation argued by Shneiderman (1983). However, the BubbleFish holds a substantial amount of directness at the interface not generally found in other systems that rely on direct manipulation. Here, the users do not merely control representations of themselves in a virtual space, e.g. through a mouse pointer. Rather, the BubbleFish provides representations of itself outside its realms, and lets the user manipulate these representations directly in the physical world. Hence, a blended space of the physical and the virtual is provided in which both the user and representations of digital information subsist and interaction takes place.

The BubbleFish should hence be seen as an example of how the virtual and the physical world can be blended by the use of existing and straightforward technology to produce practical and, expectantly, usable applications within HCI.

Three Possible ways of Implementation

Now that we know the basic idea behind the BubbleFish, I like to present three different ideas on how this concept for digital document management could be implemented. Each of these implementations draw on use of different technologies for allowing physical and virtual blend, and each have their own pros and cons. First, we could implement them with video projectors and camera tracking. A second way of implementation would be to use the ARToolkit and a pair of Sony Glasstron personal glasses. A third way is to use the Sony glasses with a system based on tracking of the user's head and hands. I will now go through each of these implementation ideas to explain them more in detail and discuss their strengths and weaknesses.

Implementation 1: Video Projectors and Camera Tracking

This implementation builds on the notion of having one, two, or more (typically two) video projectors that project a computer generated picture onto large screens surrounding the user. User movement is tracked by a video camera set up behind the user, equipped with suitable computer vision software. The video camera is hence able to track a user as they are manipulating a bubble, i.e. a digital document, that is visible on the screen, and act accordingly. If the user holds her hand over a particular bubble for some specified time, that particular bubble goes into "move mode", and is able to be moved by the user by moving the hand to a particular

location, either a different location on one of the two screens, or activating it by dragging it to the computer screen.

A conceptual overview of this prototype is found in *Figure 8*.

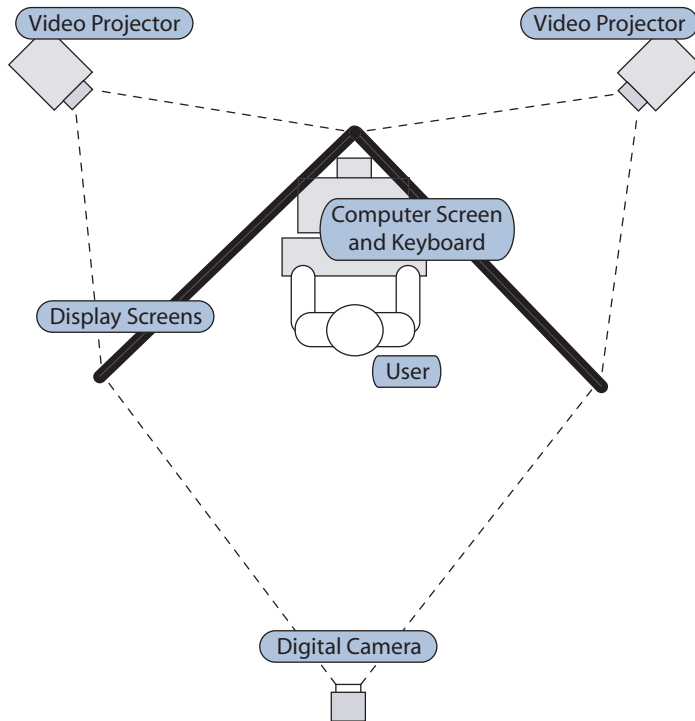


Figure 8. *Implementation 1 overview*

The benefits of this implementation would be that would use existing hardware and available computer vision software. On a level of interaction, it would also be unobtrusive to use, at it would not require the user be equipped with neither augmented reality glasses nor with tracking equipment mounted on the body.

On the downside, the display screens would not provide the user with a true three dimensional space on which to place the bubbles. Rather, this space would remain two dimensional. It is of course possible to fake three dimensions in a way similar to how it is achieved on a traditional desktop computer screen. However, as the user is directly interact with the bubbles by physical manipulation; faking 3D would by necessity introduce a number of interaction abstractions to the interface that would do no good to the proposed directness of the interface. This abstraction is of course also true for how we today interact with 3D spaces with a desktop computer, but here we already suffer from a number of interaction abstractions, so introducing 3D is simply a matter of adding abstraction—not introducing it. Regretfully, this would be the case with the implementation of this prototype. It seems likely that such an implementation would not benefit from using three dimensions. It would however still be interesting to implement this prototype using the two available dimensions, which probably in themselves would prove useful.

Implementation 2: ARToolkit and a Sony Glasstron Personal Glasses

This implementation would use the ARToolkit developed by the HIT Lab at University of Washington, a pair of Sony Glasstron personal display glasses and an off-the-shelf web cam head-mounted on the user. This is also the prototypical system which has been implemented in part. Benefits include use of existing hardware and software, that it is relatively cheap, and that it does not require a large or set location in which to operate as does the previous implementation.

However, the ARToolkit's approach to vision based AR has several limitations that currently constrain further development of the BubbleFish. The key concern is that it requires physical representations of the bubbles in terms of a printed pattern for each bubble, which in practice impedes free positioning of the bubbles in the 3D space. Likewise, the BubbleFish has no physical control over the printed patterns, and is thus not able to rearrange them e.g. to support views. In some sense, this implementation leans more towards the idea of tangible computing than augmented reality.

The following figures (*Figure 9-13*) demonstrate a prototypical set up for this implementation of the BubbleFish document manager.



Figure 9: *Implementation 2 overview*



Figure 10: *User equipment, Sony Glasstron glasses and head-mounted camera*



Figure 11: *The stage and a bubble*



Figure 12: *The Trashcan, a defined space*

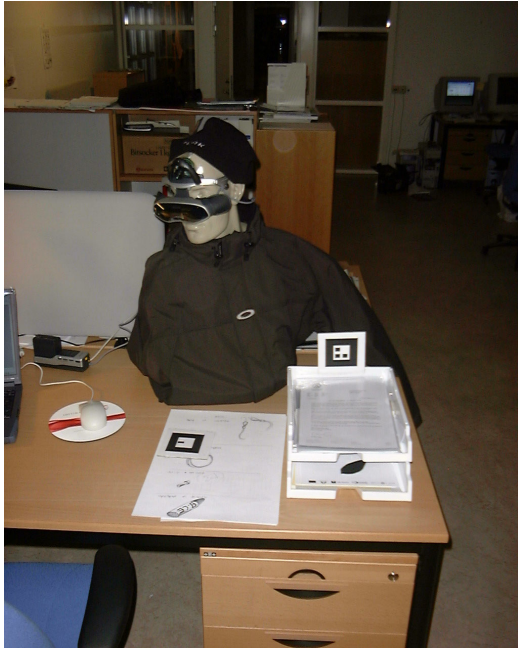


Figure 13: *The Outbox, another defined space*

Implementation 3: A System Based on Tracking of the User's Head and Hands

Finally, for a number of reasons, the most complete implementation of the BubbleFish document manager would probably rely on tracking of the user's head and hands, and where a computer generated image is presented by see-through augmented reality glasses, such as the Sony Glasstrons. The primary benefit the user would have of this set up is that it would allow the bubbles to have entirely virtual representations, and to be placed and manipulated freely by both the user and the BubbleFish application in a three dimensional space, overcoming one of the limitations of the previous implementation. Hence, this implementation would support both defined spaces and views. In addition, there would not be any need for additional hardware than those mounted on the user.

The less beneficiary side of this implementation idea is that it suggests a substantially more complicated technical solution than any of the two previous implementations. The software generating the user illusion of being surrounded by bubbles would need to be fast and accurate in order to keep the illusion working and the user from being annoyed by lags and misalignments.

The combination of two sets of trackers, one for the head so that we would know in which direction the user is looking, and one for the hands (or at least one hand), so we know if the user's hand is in close vicinity to that of a bubble. The combination of these two tracker systems with the problems of providing the user with a fast screen updates and feedback is a challenge both to today's augmented reality equipment, and to computer vision and augmented reality researchers. Undoubtedly however, this implementation would, if implemented in an unobtrusive manner, would be able to draw on the full range of features of the BubbleFish system as they have been introduced. *Figure 14* introduces this implementation through a conceptual sketch.

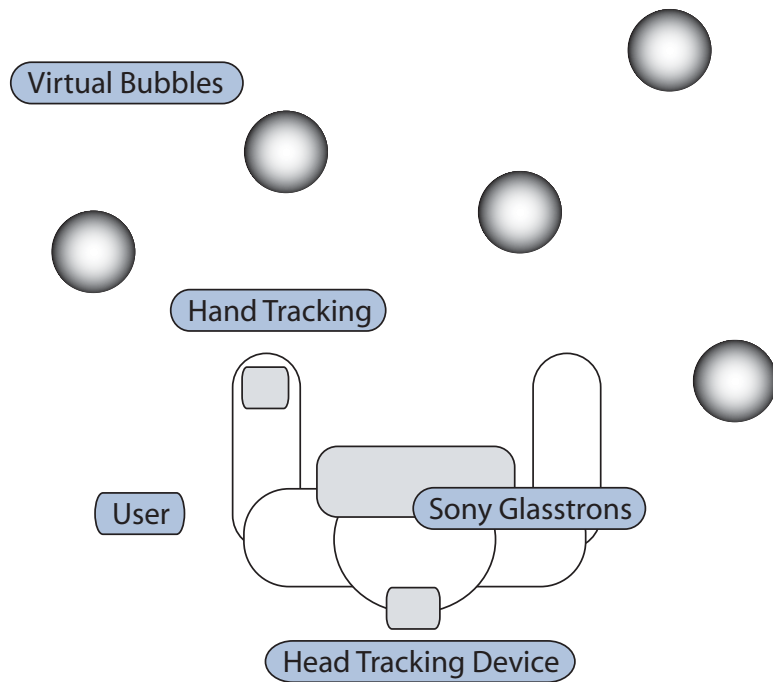


Figure 14. *Implementation 3 overview*

Future Directions

A possible future direction would be to add sensors for the physical environment, so that the user can keep certain bubbles in a drawer or a cupboard, which only become visible when open. Sensors would also allow defined spaces, such as a trashcan, to be physically moved but still retain its virtual functionality.

Another possibility would be to explore providing shared spaces, e.g. a defined space which would be mirrored at other places, for instance making it possible for colleagues to share and manipulate documents.

Obviously, if one or more of these suggested prototypes are implemented, a substantial amount of usability evaluation will have to be carried out in order to state to what extent the BubbleFish framework of interaction is feasible in real-life situations.

Conclusions

The BubbleFish document manager uses AR to allow representations of digital documents, bubbles, to be placed freely in the room by a user. The user saves a document by physically moving it out of a virtual stage into the room, and opens it by moving it back into the stage. Defined spaces are provided to give physical objects and places in the room virtual functionality, and views that allow bubbles to be repositioned according to searches, tasks and time. Hence, the BubbleFish provides an interaction space that is a blend of the physical and the virtual. In this paper, the project's basis in the graduate course "Trends in HCI" has been explained. Furthermore, I have suggested and discussed three possible implementations.

References

- Card, S. K., Robertson, G. G., and Mackinlay, J. D. (1991). The Information Visualizer: An information workspace. *ACM Conference on Human Factors in Computing Systems (CHI '91)*, ACM Press, 181—188
- Card, S. K., Robertson, G. G., and York, W. (1996) The WebBook and the Web Forager: An Information Workspace for the World-Wide Web, *Proceedings of ACM CHI'96*, ACM Press, 111—117
- Freeman, E. & Gelernter, D. (1996) Lifestreams: A Storage Model for Personal Data, *ACM SIGMOD Bulletin*, ACM Press, March, 80—86
- Ishii, H. & Ullmer, B. (1997). Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *ACM Conference on Human Factors in Computing Systems (CHI '97)*, ACM Press, 234—241
- Kullberg, L. R. (1996) Dynamic Timelines: visualizing the history of photography. *Proceedings of ACM CHI'96*, ACM Press, 386—387
- Mackinlay, J. D., G. G. Robertson, and S. K. Card. (1991). The perspective wall: Detail and context smoothly integrated, *ACM Conference on Human Factors in Computing Systems (CHI '91)*, ACM Press, 173—179
- Malone, T. (1983) How Do People Organize Their Desks? Implications for the Design of Office Information Systems, *ACM Transactions on Office Systems*, 1(1), ACM Press, 99—112
- Picard, R. W. (1997) *Affective Computing*, MIT Press, Cambridge
- Ranson, D. S., Patterson, E. S., Kidwell, D. L., Renner, G. A., Matthews, M. L., Corban, J. M., Seculov, E., & Souhleris, C. S. (1996) Rapid Scout: Bridging the Gulf Between Physical and Virtual Environments, *ACM Conference on Human Factors in Computing Systems (CHI '96)*, ACM Press
- Rekimoto, J. (1999) TimeScape: A Time Machine for the Desktop Environment, *Proceedings of ACM CHI'99 Extended Abstracts*, ACM Press, 180—181
- Robertson, G. G., J. D. Mackinlay, and S. K. Card. (1991). Cone trees: Animated 3D visualizations of hierarchical information. *ACM Conference on Human Factors in Computing Systems (CHI '91)*, ACM Press, 189—194.
- Shneiderman, B. (1983) Direct Manipulation: A Step Beyond Programming Languages, *IEEE Computer*, 16, 57—69
- Small, D. (1996) Navigating Large Bodies of Text. *IBM Systems Journal*, Vol. 35, No. 3&4
- Tesler, J. D. & Strasnick, S. L. (1992). FSN: The 3D File System Navigator, *Silicon Graphics, Inc.*

Vallino, J. R. (1998) Interactive Augmented Reality, *Department of Computer Science, University of Rochester, Rochester NY*

Warwick, K. (2000) Cyborg 1.0, *WIRED*, February