PaperSpace:
A Novel Approach to Document Management by Combining Paper and Digital Documents

A Thesis Submitted to

The College of Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the

Department of Computer Science

University of Saskatchewan

Saskatoon, Saskatchewan

By

Samer A. Sallam

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5A9

Abstract

Personal document management systems provide good support for storing and organizing digital documents. However, there are no computer tools that support organization of paper documents on our desks. We ran a study of people's organization of their office desk space with respect to their digital workspace. This study resulted in a set of requirements for a media bridging tool. Based on these requirements, we built a prototype media bridging tool called PaperSpace that uses computer vision to link paper and digital documents. The system also tracks piles of paper documents on the real desktop, and links those papers to digital documents stored in the computer. Digital documents can be sorted and grouped according to the physical layout of the corresponding papers on the desk. The system automatically creates digital piles of documents in a simulated desktop that reflect the paper piles on the real desktop. The user can access valuable information through the system, such as printing statistics, location of a printed document on the desk, and past projects and their documents. A two week user evaluation of the system showed interesting usage scenarios and future trends for improving user interaction.

Acknowledgements

First of all, I like to thank my awesome supervisor Sriram Subramanian for his full support, suggestions, guidance, and trust in my work. I also like to thank Carl Gutwin for his priceless insights and road pointers. Special thanks are due to my friend and colleague Miguel Nacenta. I thank him for always being available for productive discussions of the project. I also thank him for being my first user of the system. I also thank Andy Cockburn for his suggestions in relation to the interview study.

I thank everyone who participated in the interview study and in the evaluation study. I also appreciate the patience and support that I received from the members of the Interaction lab where this research took place.

I dedicate this work to my dear parents for their love and support in this and every aspect of my life.

Table of Contents

Permission to Use .... **i**

Abstract .... **ii**

Acknowledgements .... **iii**

Table of Contents .... **iv**

List of Figures .... **vi**

Chapter 1 Introduction **1**

1.1 Problem 4
1.2 Solution 4
1.3 Steps in the Solution 5
1.4 Evaluation 6
1.5 Contributions 7
1.6 Thesis Structure 8

Chapter 2 Review of Literature **9**

2.1 Working with Documents and Desktop Management 9
2.1.1 Handling Paper Documents 10
2.1.2 Personal Information Management 14
2.2 Augmented Reality Systems 19
2.2.1 Augmented Paper Systems 19
2.2.2 Augmented Desktop Systems 22
2.3 Tracking Physical Artifacts 26

Chapter 3 User Study **29**

3.1 Approach 29
3.2 The Physical Space Organization 30
3.3 The Virtual Space Organization 31
3.4 Comparing Virtual and Physical Spaces 33
3.5 The Virtual-Physical Link 33
3.6 Requirements for a Media Bridging Tool 35

Chapter 4 PaperSpace Design **37**

4.1 System Overview 38
4.2 The Document Tracker 39
4.2.1 Visual 2-Dimensional ID Tags 40
4.2.2 Finding and Decoding the Tags 42
4.3 The Pile Tracker 46
4.4 The Document Manager 49
4.4.1 Adding Documents 50
4.4.2 Printing Documents 50
4.4.3 Finding Documents (search methods) 51
4.4.4 Documents Views 52
4.4.4.1 Document Query List View 52
4.4.4.2 Pile List View 52
4.4.4.3 Desktop View 53

4.5 Database design 54
Chapter 5 User Evaluation and Technical Performance 56
5.1 Technical Evaluation 56
5.1.1 System setup 56
5.1.2 Test-1 57
5.1.3 Test-2 58
5.1.4 Pile Tracking 59
5.2 User Evaluation 61
5.2.1 Method 61
5.2.2 Results 62
Chapter 6 Discussion 65
Chapter 7 Conclusion 68
7.1 Summary 68
7.2 Contributions 69
7.3 Future Work 70
References 72
Appendix A - Interview Structure 79
Appendix B - PaperSpace Database Schema 81
Appendix C - Pile Tracking Test 82

List of Figures

Figure 1.1 The PaperSpace System ... 3
Figure 2.1 An office full of papers, documents, and folders ... 10
Figure 2.2 Diagram of an organized and a messy office ... 11
Figure 2.3 Example of a cluttered desk ... 12
Figure 2.4 A snapshot of Vista, an interface for the Presto system ... 15
Figure 2.5 Snapshot of “MyLifeBits” ... 16
Figure 2.6 UMEA: to create and manage project-specific work contexts.. ... 17
Figure 2.7 A screen shot of the Stuff I’ve Seen system ... 18
Figure 2.8 Schematic diagram of the DigitalDesk prototype ... 20
Figure 2.9 A copy/paste interaction in PapierCraft ... 21
Figure 2.10 Illustration of the Intelligent Paper system ... 22
Figure 2.11 Overview of the EnhancedDesk setup ... 23
Figure 2.12 Using two cameras to cover a bigger area ... 25
Figure 2.13 An example of live card ... 25
Figure 2.14 A snapshot of PaperWindows prototype in action ... 26
Figure 3.1 Users' organization levels in the physical and virtual spaces. ... 33
Figure 4.1 System architecture diagram ... 39
Figure 4.2 Desk and camera setup ... 40
Figure 4.3 A sample of the visual tag ... 41
Figure 4.4 Sample Image of the result of applying the canny operator ... 43
Figure 4.5 Threshold sub-image of a candidate tag ... 44
Figure 4.6 Region adjacency graphs for a guide block and for a data block ... 44
Figure 4.7 An exact image of the tag after rotation and thresholding ... 45
Figure 4.8 A snapshot of the document manager’s GUI ... 49
Figure 4.9 Sample of the print confirmation dialogue ... 51
Figure 4.10 The Desktop View displays thumbnail images of documents ... 54
Figure 5.1 PaperSpace setup for the technical evaluation study ... 57
Figure 5.2 Papers used in Test-1 ... 58
Figure 5.3 Papers used in Test-2 ... 59

Figure 5.4 The three PaperSpace setups used in the user evaluation study. 62
Figure 6.1 Example of a tangible command bar. 67

Chapter 1

Introduction

Documents form an integral part of our everyday activities. We use documents in both paper and digital forms, with each form having its own strengths and weaknesses. Paper documents are easy to produce, handle, read and port whereas digital documents are easy to store, search, edit and retrieve.

In the digital world we are surrounded by files, folders, and email. Despite this abundance of digital information, the idea of a paperless office remains a myth today as it was decades ago. It seems that as we have much more access to information we need to print that information "in order for us to read it and make more sense of it" [57]. Instead of thinking about technology as a means to get rid of paper we can find novel ways to use technology to augment our use of paper, to support the management and organization of our paper space.

In the digital world, document management systems augment the computer's typical file system and help the user to better organize and make use of their digital documents. They enable users to store, retrieve, and use their documents more easily and powerfully than they can do within the file system itself. These systems are effective for managing the digital document space, but what about the paper space?

While in the realm of the paper world, managing paper documents is a typical activity performed by information workers which receives no computer support. Therefore, users develop their own strategies for organizing their paper documents. As the number of papers multiplies on the desk, the organizing task becomes more difficult and demands more cognitive effort to maintain. Many users end up having dozens of papers split across several stacks on the desk.

Most knowledge workers deal with both digital and paper documents, often moving back-and-forth between the two mediums. On many occasions paper documents are created from their digital version, but when a user prints a paper document the connection between the two versions is lost. Yet current technology has the potential to substantially bridge the gap between media, bringing the benefits of electronic file organization and search to paper documents, and also integrating the user's physical actions (such as ink annotations and the formation of desktop piles) into the electronic environment.

Researchers have investigated the management of paper documents to inform the design and development of computer tools that support the management of digital documents [12, 32, 43, 44, 57, and 62]. However, there is still a lack of understanding of how users deal with both spaces simultaneously. Furthermore, there has been little empirical research on how users employ tools that bridge the media gap.

To examine the nature of the link between the two document spaces we interviewed several researchers from different disciplines. Our results show that when managing documents users most often dispense more resource on one medium over the other and so they miss out on the benefits of one of the mediums. Also, most users relied on their memory or developed ad-hoc strategies to maintain the connection between the two mediums. Importantly, many participants complained about the difficulty to cope with the other space.

Therefore, one possible solution is a system that can somehow keep track of paper documents. To achieve that, the main challenge is in accurately recognizing those documents. Many researchers have studied the tracking of artifacts on our desks [52, 28, 60]. The outcome of those studies proves that real-world artifacts can be recognized and tracked with acceptable accuracy. Thus, we can use similar techniques to recognize and track paper documents.

To experiment with a media-bridging environment, we developed PaperSpace [58] a document management system that uses computer vision to track documents on the desk and maintains their links to their digital counterpart (see Figure 1.1).

PaperSpace is inspired by [34] and [56] described in Chapter 2. A key difference is that our system supports real-time tracking of multiple documents with very high accuracy. We restricted the current prototype to supporting the activities in an academic research lab. Researchers work with many papers related to their work which make them suitable for testing our prototype.

Figure 1.1 The PaperSpace System: a media bridging tool for managing paper and digital documents.

We deployed the system in different research labs and observed the experiences of three users in a 15 day field study. The results showed that the system improves users' coping with documents when dealing with the paper and digital workspaces simultaneously. We also found that users developed different methods for interacting with the system than we predicted.

1.1 Problem

The problem we address in this thesis is that current computer systems do not support the management and organization of personal paper and digital documents in the user's workspace.

With the increasing use of papers in the office, piles and clutter of papers become more cumbersome to organize and keep track of [43]. The process of finding a paper in a heavily occupied desk where dozens of papers and notes spread over the desk surface requires high memory load and is time consuming. The problem is even worsened when juggling between the paper and digital media.

Those clutters of papers on our desks mainly come from printing digital documents stored in our computers. By printing documents, papers accumulate on our desks over time. When working with such papers, it is a common situation where we have an interesting paper on the desk for which we need to retrieve its digital file, for example, to send it to someone, or to copy/paste some of its content. Unfortunately, this process could take anywhere from one second to minutes in some cases. Part of the problem is that once a document is printed it looses its link to the digital source.

Furthermore, an old print of a document may hold useful annotations. So we like to easily find such a document. Currently, people sometimes resort to reprinting of papers already on their desks, either because they are not sure if it is on the desk or because it takes more time to find than to reprint. In some situations, we can not find a paper in time or we do not remember that it was printed before. Reprinting unnecessarily means wasting time, paper, and ink resources.

1.2 Solution

The solution we propose is to support the management and organization of documents by coupling the physical paper space with the virtual document space.

To understand the nature of the physical-virtual link, we need to study the way people currently manage their physical and virtual spaces. This will also provide us with guidelines for designing an effective solution prototype.

Our solution involves building a system to recognize and track documents on the desk. The system maintains a link between the identified documents on the desk and those stored in the computer. At the same time, the system keeps a history of the user's recent activities and provides a rich user interface for accessing the information in a timely manner.

The link between paper and digital documents is established by printing specially designed 2-dimensional tags, on the paper's margins, that can be identified by using computer vision techniques. These tags encode information about the corresponding digital document. Tracking documents allows for extracting information about the arrangement of papers on the desk and reflecting that information on the document views available to the user.

Thereafter, we store information about both paper and electronic documents in addition to history of user activities in a relational database system. Important information such as what document has been printed, when, where and by whom, is saved in the database. The current location of printed documents could simplify the finding task.

These components are presented in a rich user interface with multiple views and query methods. The variety of views enables the user to visualize their current work from different perspectives and allows for different manipulations of the desk's contents. Document queries can be performed on both paper and electronic documents. To sum up, the system provides a variety of features that allows effective management of the user's document space.

1.3 Steps in the Solution

The solution involves two main phases: identification of the user requirements for maintenance of the physical-virtual link and implementation of the system that fulfills these requirements. We further describe our research in the following steps:

1. We drew out the requirements for a document management system in the context of both physical and electronic workspaces. We obtained these requirements from two sources:

a. The collection of previous studies of office desk organization and personal document management.

b. A user study based on semi-structured interviews with knowledge workers in their natural work settings.

2. Based on the results of step one, we built a system that helps the knowledge workers in the management and organization of their paper and electronic documents. This step was carried out in the following order:

    a. Designed an identification scheme for paper documents so that we can link these to the electronic documents.

    b. Developed an appropriate tracking method that will allow the system to track the location of papers on the desk -based on the identification scheme- and the spatial relationships among the papers.

    c. Build a database for storing the information about documents and their locations. The structure of this database will be based on the requirements of step one.

    d. Integrate the document tracker and database components into an interactive user interface that allows the user to add, edit, remove, and query documents. The purpose of this interface is to have sort of a working system that can be tested in a real work environment.

1.4 Evaluation

Our evaluation involved two methods as follows:

1. In the first step we tested the technical performance of the system in two phases:

    - Tested the accuracy and speed of identifying and recognizing individual tags.

    - Tested the performance of the whole system's accuracy and speed. This involved evaluating the accuracy of tracking piles of papers, and the consistency of their representations in the user interface.

2. A user study, which consists of a series of user interviews conducted while trying the system out, will provide us with a basic understanding of how the system could be steered to meet their needs. Users will be asked to use the system to complete a set of predefined tasks. These tasks span adding, printing, and finding documents. Of particular interest is to observe users while trying to find a document on the real desktop by using the virtual desktop view. This study also will include a set of questions at the end of the session to obtain user feedback and to verify our initial design requirements.

We ran a field study by deploying the system for three users in order to test the usability of the system and obtain a qualitative evaluation of its implementation. The participants were two professors and one Ph.D. student all of whom were involved in research studies. Our system can help the users in organizing and linking their research related paper and digital documents. The users will be asked to use the system for all their reference papers and to record information about system features that are useful, unnecessary, cumbersome, or those that need to be improved.

1.5 Contributions

The major contribution of this research is the expansion of the design space of personal document management systems by using the physical desk space in organizing and finding documents. The tracking of paper documents on the desktop opens a new paradigm upon which electronic documents can be viewed and organized. Minor contributions are:

- Identification of the requirements for supporting and integrating the paper and digital work spaces. There has been extensive work on supporting the digital work space, which is not the case for the paper work space and combining the two work spaces into an integrated system.

- Demonstrating through a working prototype the feasibility of building a system that manages the link between paper and digital document spaces.
- Identification of the benefits, tradeoffs and limitations of using such a system. A working paper and digital document bridging system has not been tested previously.
- A paper tracking method based on computer vision techniques. Previous computer vision methods for tracking papers on the desk top suffered from slow response time [34].

1.6 Thesis Structure

The rest of this thesis is structured as follows: Chapter 2 presents a summary of previous research related to our work. Chapter 3 presents our user study that was conducted to find the current strategies people use to manage both of their paper and digital document spaces. In Chapter 4 we describe the system prototype that we built to test a media bridging tool based on the findings of our user study. Chapter 5 describes the results of system evaluation. In Chapter 6 we discuss our findings and finally we present our conclusions and the possible future directions for this work in Chapter 7.

Chapter 2

Review of Literature

In this chapter we present a detailed review of previous work related to our research. The related research includes "working with papers and desktop management", augmented reality systems, and tracking physical artifacts. By desktop we mean the physical desktop and not the computer virtual desktop, unless otherwise stated. In the working with papers and desktop management section we first describe studies of how users manage and organize their offices, and then review the current tools of both personal information management and document management. The next section, augmented reality systems, surveys the state of the art in paper and desktop augmented systems. Finally we outline the available technologies for tracking physical objects on the desktop.

2.1 Working with Documents and Desktop Management

Since the advent of personal computers, electronic and paper documents continue to coexist in our offices. The idea of the paperless office remains a myth [57] despite the availability of powerful electronic tools for handling documents. To understand why paper documents still occupy a significant portion of office space (see Figure 2.1), several researchers studied paper documents, and how they are used on our desks [43, 44, 46, 62, 1, and 12].

Figure 2.1 an office full of papers, documents, and folders

Sellen and Harper [57] provide an inspiring study of paper documents use in the modern office. In “The Myth of the Paperless Office” they explain the rise of paper use by the unique affordances of paper. The properties of paper being thin, light, porous, opaque, flexible, and inexpensive, among others, allow for easy authoring, reading, reviewing and annotating. They also observed the important role of paper for collaborative working and sharing [25].

Paper and electronic documents are used by information workers interchangeably. Each of the two mediums has its own advantages. Electronic documents are easy to revise, store, and retrieve. Paper documents, on the other hand, are light, flexible, portable, and easier to read and annotate. In this section we discuss previous research of paper use and desktop organization.

2.1.1 Handling Paper Documents

Information workers use different strategies for managing and organizing paper documents. Personal information management systems, for electronic documents, are much less efficient than those techniques used by information workers in the physical domain. This subsection describes the important user studies that explore how information workers handle physical documents to improve the design of PIM systems.

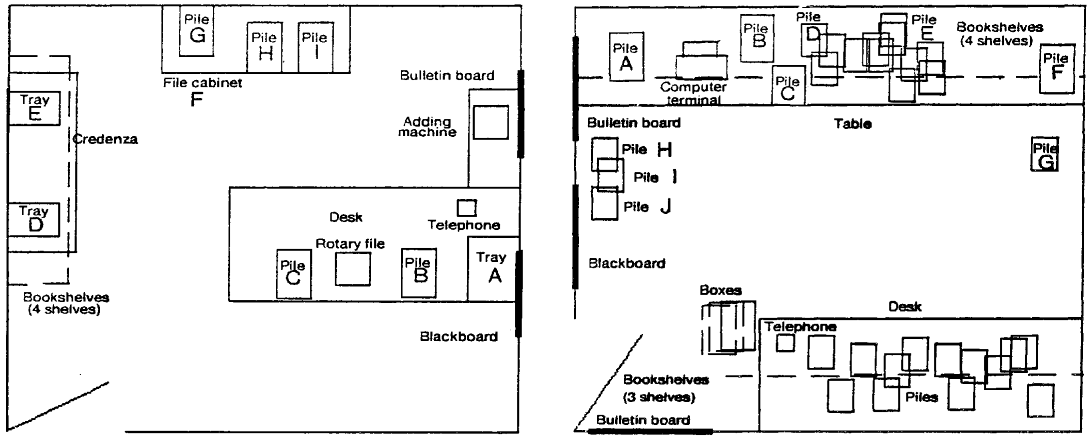


Figure 2.2 Left: diagram of an organized (neat) office. Right: diagram of a messy office, [43]

The earliest study on desktop management is Malone 1983's exploratory observational study of the office desk's organization. He identified two methods people use for organizing information: filing (systematic grouping of documents) and piling (miscellaneous grouping of documents) which was also confirmed by another user study [44]. Based on that classification, information workers can be filers or pilers, but most people do a combination of both with the emphasis on one or the other. Filers (filing-centered people) tend to have "neater" offices and spend more time organizing their office space. Pilers, however, have sort of messy offices, and organize their offices less frequently than filers (see Figure 2.2). An interesting finding is that workers organized their information not only to find things but also to help reminding them of things to be done. Malone introduced three reasons for piling information: the difficulty of creating labeled folders and binders, the cognitive difficulty of categorizing documents to serve as reminders, and to have easy access of frequent information.

An interesting design implication that emerges from this study is that computer systems need to assist users in finding and reminding of information. In the process of finding, the system should simplify the process of classifying information by allowing multiple and/or deferred classification and by automating the classification when possible. Reminding can be facilitated by using a variety of cues such as frequency of displaying information, size, location, and color of the icon representation.

Mander et al. [44] conducted a user study on how people deal with information flow in the physical workspace. The study involved interviewing users in their work area to find how they dealt with information from the moment it arrived until it was finally stored. The study shows that folders are used mainly for archiving purposes, since they require explicit classifying of documents. Many users considered folders as inconvenient and required considerable time for filing documents. Piles, however, had a less rigid categorization requirement, and were more flexible; piles allow temporary placement of documents, and are easy to change and reorder. For some users, piles were considered self-revealing and browsable. Users could tell the contents of a pile from its appearance and used different methods for sifting through a pile, revealing the edges of items, partially lifting items off the pile to see the lower items, and spreading all items of a pile and look at the contents in parallel. This study among others, give convincing proof for supporting pile creation and interaction in computer interfaces.

Figure 2.3 Example of a cluttered desk

In their study of knowledge workers, Kidd [32] suggested the need for "informing" systems rather than the passive filing systems that store large amounts of information. The clutter-up found in knowledge workers' office serve several purposes (see Figure 2.3). The physical space (desks and floors) serves as a temporary layout for holding uncategorized, raw input and ideas. The visual layout of things on their desk provides powerful context information to help them reset their mind on a recent idea or set of ideas.

A long term ethnographic study of personal document management of a group of information workers was conducted to improve electronic technologies by observing the advantages that paper offers [12]. One interesting difference between paper and digital tools is the flexibility paper offers in adopting a personalized management approach while digital tools only allow a single method of structuring documents (using folder hierarchies). Also, there is no need to explicitly classify paper documents as they move around the desk space, which is contrary to the computer system in which a document must be explicitly named (categorized). The authors identified different document management needs based on the flexibility and changeability of activities. In the "administrative" activities, where tasks have pre-structured procedures, inflexible, and do not change often, it is easy to categorize documents and document management systems can support this type of activity effectively. On the other end, there are the "research" activities as the authors call it. This type of activities is highly flexible and changeable according to circumstances and is more difficult to support by DMS.

In conclusion, [12] provide insightful implications for the design of document management systems:

- Documents should be grouped according to their task-related context.
- A DMS should support defragmentation of related documents that originate from different sources (email and the different file formats).
- Grouping and regrouping of documents should be easy to perform.
- A DMS should allow flexible and custom restructuring of documents.

From the previous studies we can see that PIM systems may be improved by allowing user interaction similar to the way people handle their physical documents. It is clear that piles are important in their own respect as they were used by most users. Thus, support for grouping of documents into piles would be a great asset to PIM. This would automatically solve, although partially, the classification problem as well; users can just throw documents into piles without having to categorize them. Piles also support regrouping of documents simply by moving documents between different piles.

2.1.2 Personal Information Management

Personal information management (PIM) refers to the activities people perform in order to acquire, organize, maintain, retrieve and use information items such as documents (paper-based and digital), web pages and email messages for everyday use to complete tasks (work-related and not) and fulfill a person's various roles (as parent, employee, friend, member of community, etc.). We are interested in studies of PIM that aim to improve current PIM computer systems.

In the context of business firms/organizations, document management systems (DMS) are a scaled up version of PIM. DMS provide support for the aquisition, storage, organizing, versioning, and access control of electronic documents in an organization or an enterprise. Although we limit our discussion in this section to personal information management systems, these systems also apply to document management in the context of personal space.

Many factors related to the structure of current computer operating systems make the management of personal information difficult. The standard file system is based on single inheritance file hierarchy. Researchers have found this scheme troublesome and limiting for most users [17, 10, and 51]. In a survey of PIM tools [51], users faced the following difficulties:

- The inconvenience of separation of files, email, and bookmarks. This forced users to use different tools for dealing with different types of data.

- Scarcity of tools that support short-term notes or remarks, and currently available tools are not practical.

Manual structuring and ordering of information, which can be automated to some extent, takes too much time.

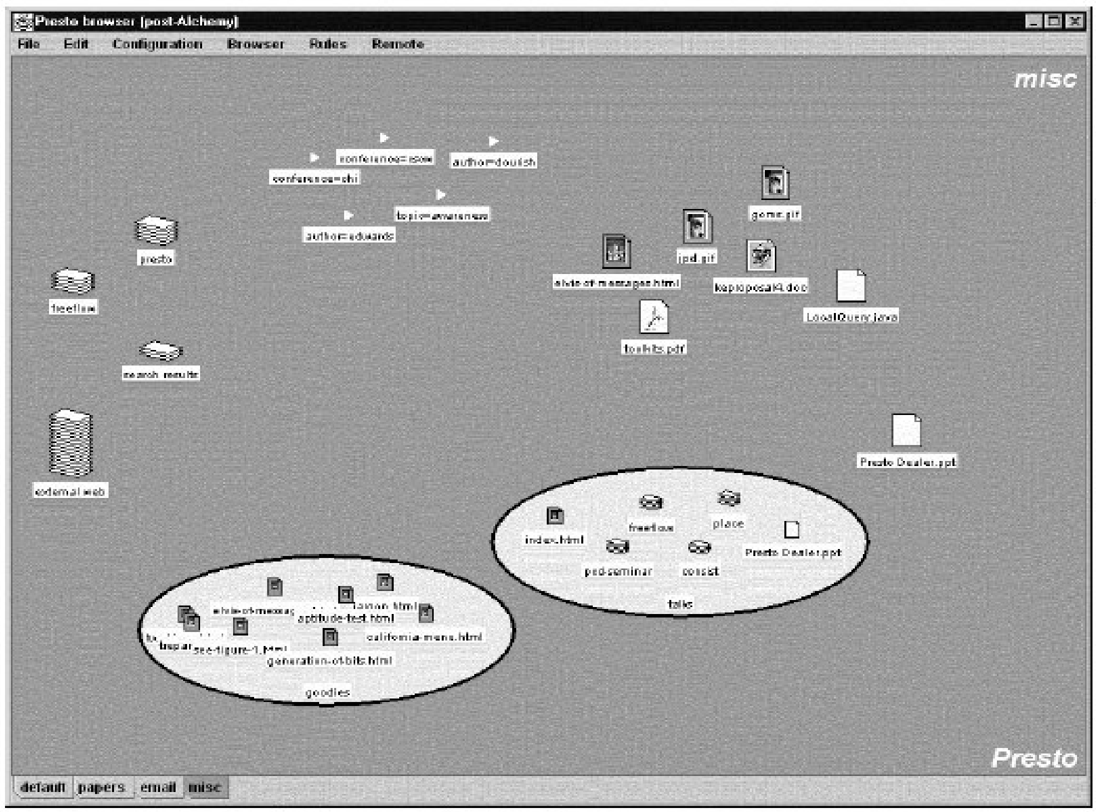


Figure 2.4 A snapshot of Vista, an interface for the Presto system [16]. This interface offers multiple representations of the same document database. Documents may apear at multiple places at once. Documents are displayed individually or in collections and piles.

Researchers followed different trends to address those difficulties. Some researchers extended present tools while others developed new approaches to PIM. The “Presto” system [16] is a new non-hierarchical approach in which the structuring of documents is based on their attributes that can be generated automatically or can be added by the user. Numerous attributes hold meaningful document information for different users. Presto provides dynamic document collections generated from user queries on document attributes (see Figure 2.4). Documents can also appear in multiple collections. A critical task in Presto is assigning attributes to documents. Task relevant attributes are important so the system can generate task related collections. This requires high cognitive effort similar to that for hierarchical structures. Furthermore, the users interact with the system by forming queries, and this requires supplying appropriate keyword. Nevertheless, their new concept for user interaction with dynamic collections of document space would improve usage and increase PIM flexibility.

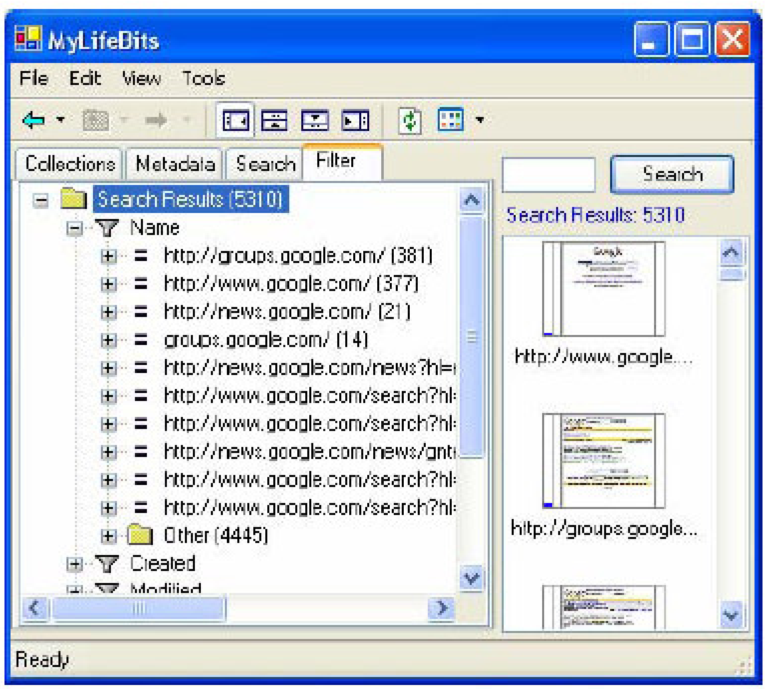


Figure 2.5 Snapshot of "MyLifeBits" showing top occurring names in search results [22].

Lifestreams [20] and MyLifeBits [21] use non-hierarchal approaches based on time as the main organizational element. Lifestreams stores and organizes all documents according to their creation time. Also, users can manually place items in the future as reminders. MyLifeBits aims to store a complete digital record of one's information (see Figure 2.5). It supports four different views: A timeline view displays documents according their related event time. A clustered-time view clusters documents by similar time. A detailed view shows a list of resources, and a thumbnail view.

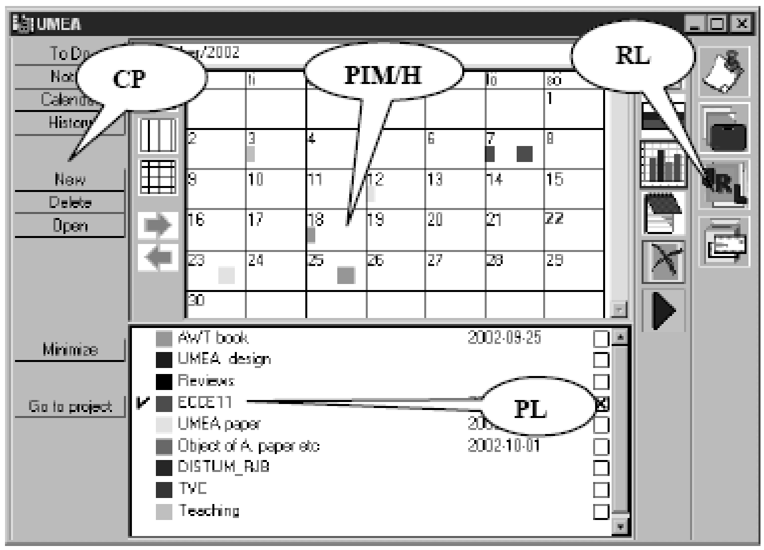


Figure 2.6 UMEA: allows users to create and manage project-specific work contexts. This system keeps all information related to a project in a single location [31].

The Umea [31] system use design implications from activity theory to support higher-level user activities. It allows managing and organizing project-specific work contexts. All activities related to a project (creating documents, reading emails, visiting web pages) are stored as a group (see Figure 2.6). As the user is working on a project, the system monitors her activities and tracks resources used, and automatically organizes and updates this information. This arrangement allows users to easily alternate working with projects without having to maintain the organization of project related resources manually.

The electronic file cabinet [40] introduces yet a complete shift from previous approaches. It creates a virtual file cabinet that mirrors a physical one. Documents are scanned into the system which in turn suggests a drawer to link the document to. This drawer recommendation is based on content comparison with other documents in the database, and documents can belong to multiple drawers at the same time. This system would be useful for the long-term type of documents in an office. But for the more active type of documents on the desktop, we need an easy way to add and track documents.

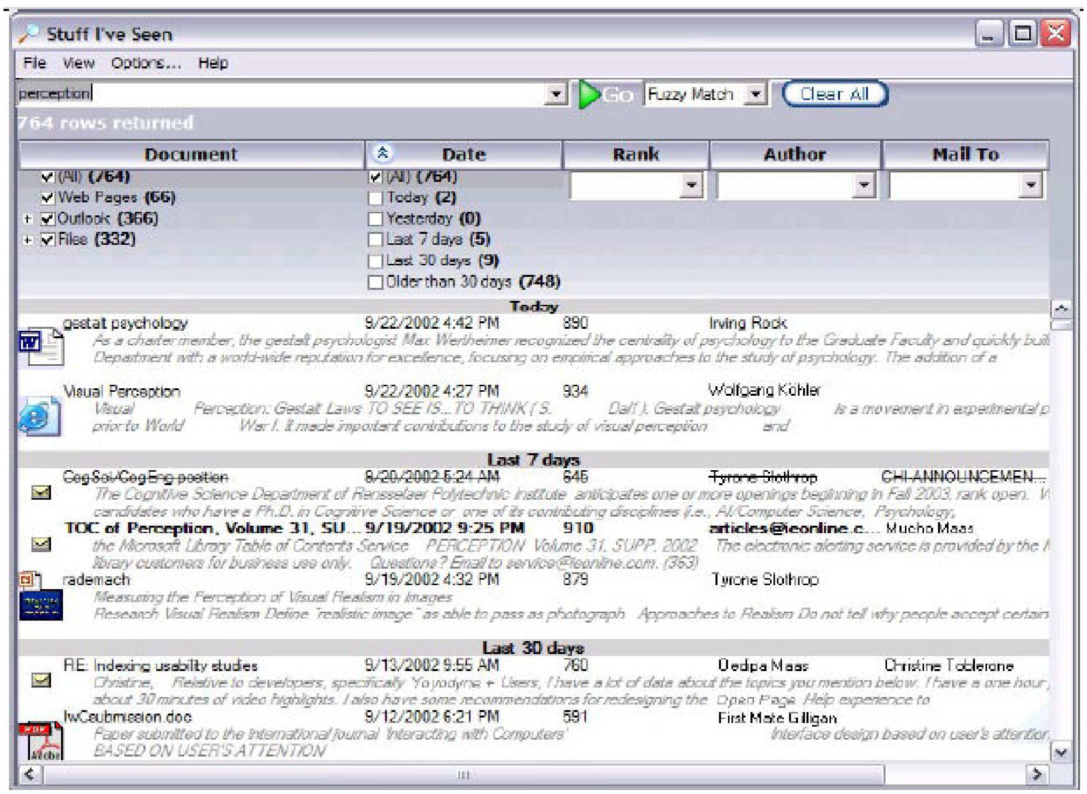


Figure 2.7 A screen shot of the Stuff I've Seen system which makes it easier to find information previously seen by the user [18].

Among the most frustrating issues in PIM is the fragmentation of information sources (email, web-pages, digital documents, paper documents, IM, etc.) which makes it difficult and time consuming to keep relevent information together in one place [18, 8, 9, 11, and 31]. There is also a lack of tools that allow searching over mutliple document sources. In an effort to overcome the fragmentation problem, Dumais [18] created a tool called "Stuff I've Seen" or "SIS" (see Figure 2.7). This tool provides unified access to information items (email, web pages, documents, etc.), that the user has already seen. Rich contextual cues are used to stimulate simple searching of information already seen.

Moreover, information workers face a considerable cognitive overload in the workplace [35]. Too much information supply and demand, constant multi-tasking and interruption, and inadequate support of everyday activities such as planning, monitoring, reminding, etc. are the main reasons for such overload. As the use of hierarchies for document organization is found problematic [43, 10, and 11], the requirement of classifying, naming, and labeling documents in a hierarchical scheme is not suitable for the overloaded environment. We need new schemes that can adapt to the dynamic nature of PIM needs.

2.2 Augmented Reality Systems

Augmented Reality systems are computer systems that combine real world objects with computer generated data that make the objects interactive. Pierre Wellner's digital desk [61] was a pioneering work for augmented reality research. The digital desk introduced the idea of bringing the computer to the user. This idea allows physical objects to gain electronic properties and electronic objects to gain physical properties. Instead of immersing the user in the information environment, the physical environment is complimented by the computer's information.

2.2.1 Augmented Paper Systems

There exists a significant body of research on paper augmented systems that make use of digital technologies to enhance user interaction with paper. The digital desk mentioned above was meant to make use of the advantages of both electronic and paper documents, and to allow interaction with both domains at the same time (see Figure 2.8). The focus of the digital desk, however, was on how to manipulate electronic documents with physical objects. Many experimental systems followed in the footsteps of the digital desk [50 and 3]. An exception is a system that used paper as an interface to a document services system [30]. This system enables the user to specify commands using a sheet of paper as a form. The command form can be scanned along with the document to be processed. The problem with such a system is that scanners are terribly slow and require specially designed forms.

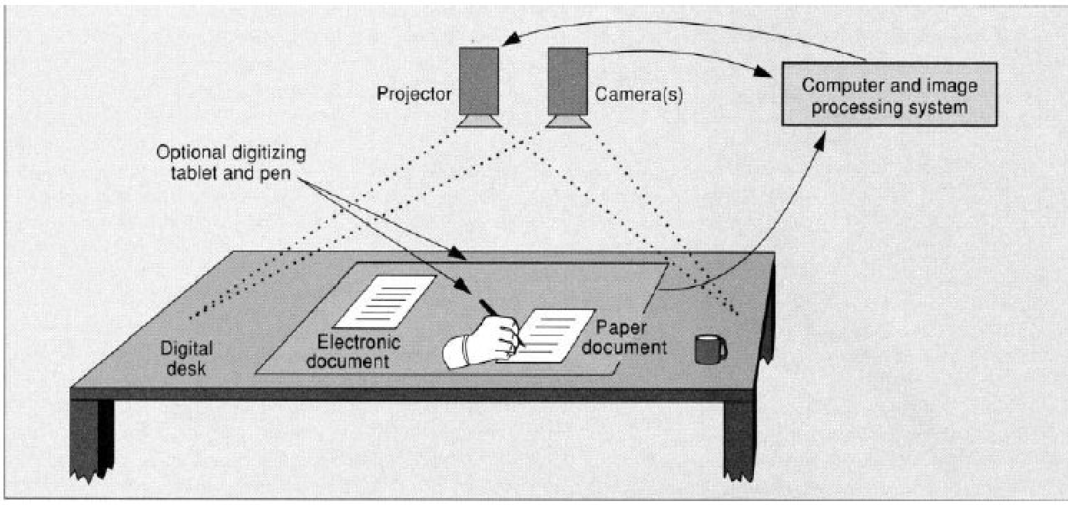


Figure 2.8 Schematic diagram of the DigitalDesk prototype [61]

Several systems have been proposed that combine electronic and paper documents and establish a link between the two domains by special devices or visual patterns printed on paper [6, 19, 23, 26, and 48]. PaperLink [6] allows the hyperlinking from paper to electronic content by using a miniature camera embedded into a highlighter pen that can capture small images around the pen's tip. This system can be useful for extracting annotations and text from paper and associating them to the electronic version of the document. However, the camera's setup is cumbersome (due to the wired connection to the computer) and the system requires the paper to be explicitly linked to the electronic document. In addition, the captured image is too small; to cover the whole paper, the user must discretely position the pen over numerous positions on the paper and manually push the shutter button to capture the image.

Similarly, the work of Guimbretiere [23], paper augmented digital documents (PADDs), couples user's annotations on paper with the digital document associated with it. The idea of PADDs is to allow people to edit and annotate digital documents on real paper and without the need for a computer to be close by. This is done by using a special digital pen that records strokes. The strokes can be transferred and merged with the digital document at any time. Also, the document is printed with a special pattern to allow the system to identify the page and document being processed. The main advantage of this system is that it eliminates the need for a computer while working with the paper document, unlike the PaperLink, which requires the use of a camera wired to a computer.

Figure 2.9 a copy/paste interaction in PapierCraft [41].

PapierCraft and ProofRite were built on the idea of PADDs. PapierCraft [41] makes papers interactive by recognizing gesture-based commands (see Figure 2.9). However, it could be difficult to manipulate text on a paper without immediate visual feedback. Such feed back can be provided using a media projector to display the results on the same or different paper. ProofRite [13] integrates the PADDs architecture with a word processing application. Users can make annotations on the printed document; upon their request, annotations would be merged back with the digital document.

Dymetman and Copperman [19] proposed the architecture for the Intelligent Paper (see Figure 2.10), a system in which physical and digital papers are tightly coupled.

Each page is covered by invisible printed marks that encode a unique global identification code of the page. The codes can be read by a pen-shaped camera that can be used as a pointing device. When the user clicks at some location in the page, an event similar to a mouse-click event is generated and appropriate information is sent to the computer. This allows a conventional paper (that contains the printed pattern) to be used as a touch sensitive surface.

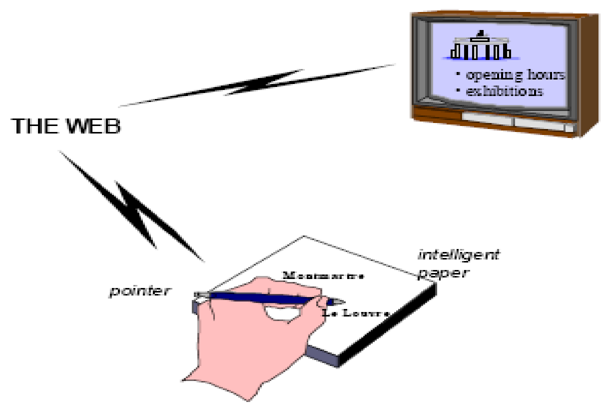


Figure 2.10 Illustration of the Intelligent Paper system [19]

2.2.2 Augmented Desktop Systems

Desktop augmentation systems aim to enrich the user experience by providing a powerful way of interacting with the computer through physical objects. The majority of those systems use a combination of projector and camera to augment the desktop surface, and to track objects. Augmented paper systems focus on augmented papers while desktop systems target the general physical artifacts on the desktop.

The InteractiveDesk system [4] enables linking physical objects to electronic files by using a camera mounted over the desktop. The system can recognize objects visible on the desktop. Multiple files can be linked to any object or paper. The focus of this system is to allow arrangements of electronic files that correspond to the associated physical objects arrangements. Another feature of the system is the easy retrieval of electronic files simply by placing the associated object on the desktop. However, there may be an overhead on the user's memory to maintain a mapping between the physical objects and the electronic files.

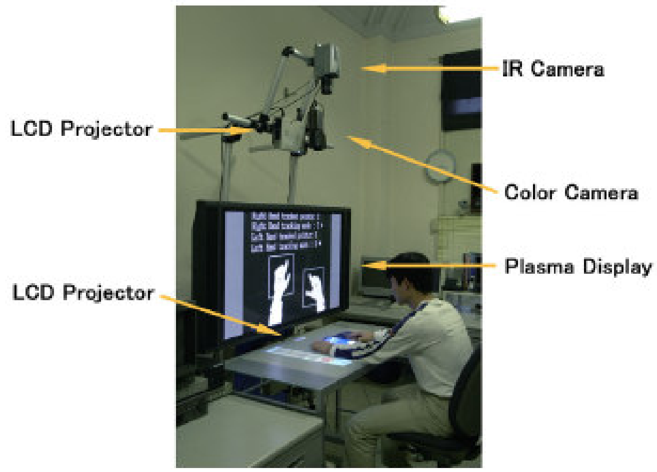


Figure 2.11 Overview of the EnhancedDesk setup [38].

Influenced by the DigitalDesk, Koike et al. [38] designed the EnhancedDesk to serve as an infrastructure for running augmented desktop applications (see Figure 2.11). In addition to the basic features of the DigitalDesk, the system can identify desktop objects using visual marks. They also introduced a novel technique for tracking the position of the user's hand and fingertips; this is done by using an infrared camera to capture the human skin which has higher temperature than other objects. First, the user's hand is located by binarizing the input image with a threshold value (i.e. obtain a binary image). After obtaining the hand region, finger tips are searched for using template matching with a circular template, and the center of palm is determined by estimating the wrist location and using morphological erosion. This technique allows the user to manipulate objects and issue commands to the system.

In an effort to improve on the EnhancedDesk object recognition, Nishi [47] proposed a new technique for object identification that eliminates the need for tags attached to objects. Instead of attaching a tag to an object, the user makes a hand gesture to register the object into the system. The system extracts and stores the color histogram for the object from the color image based on the color indexing method [59]. After object registration, digital information can be associated with that object.

Upon the user pointing to an object the system identifies the object by comparing the color histogram of the object to those stored in the database. In order to overcome the problem of false recognition of objects, the system displays small images of candidate matches from which the user can select the correct object by pointing with the finger. This technique works well for colorful objects, but for certain applications we expect it to fail.

The Self-Organized Desk is a system that keeps track of documents' contents on the desktop autonomously [56]. The system uses a pan-tilt-zoom camera that sweeps over the entire surface of a desktop to find and extract document's content. Extracted document content (by means of OCR) is indexed , organized and stored in a database along with document attributes such as time of arrival, location, and history of which pile the document belong to. The variety of information kept by the system allows the user to make a variety of search queries on the desktop content. The system also provides multiple-view visualization of the data. Although the system can track the contents of the desktop and detect changes, the performance is not real-time (it takes 30 seconds to detect a change). The documents targeted by this system are not linked to digital counterparts on the system. Besides, if we have digital copies of the documents on the desk there is no need to extract the content of the document on the desktop. Instead, we can use the digital content in combination with the location of the paper documents on the desk to organize and structure information.

When using cameras in desktop systems, we have a tradeoff between image resolution and field of view (see Figure 2.12). The problem with the currently available cameras is that they do not provide high resolution for capturing details of desktop objects while covering the entire desktop area at the same time. One solution was introduced in the Augmented Surfaces by Rekimoto and Saitoh [53], which allows users with portable computers to interchange information among their computers and other displays like wall or table, and to interact with physical objects on the desktop. To solve the resolution-view dilemma, they use a combination of two cameras, a fixed camera to detect changes on the table and a motor driven camera with pan-tilt-zoom capability to capture zoomed in images with higher resolution. Depending on the situation, this solution could be suitable, but if real-time tracking is required, this solution is not suitable as it takes significant time to scan the whole desktop surface.

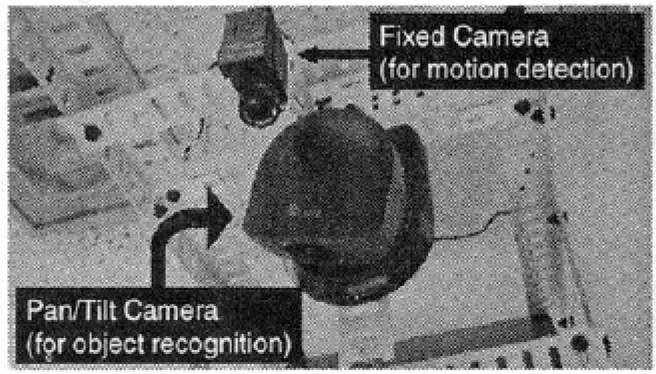


Figure 2.12 the use of two cameras to overcome the resolution vs. field of view tradeoff, as proposed by [53].

There are many more computer augmented desktop systems in the literature. To conclude this section we briefly describe Live Paper [54] and Paper Windows [28]. Live Paper makes sheets of paper of different size-classes interactive through video augmentation. The system assigns new ID's to papers as they appear on the desktop and stores the image of each paper for later comparison. Papers are augmented with video projected annotations and media (see Figure 2.13). The system also extracts writing on paper in real-time.

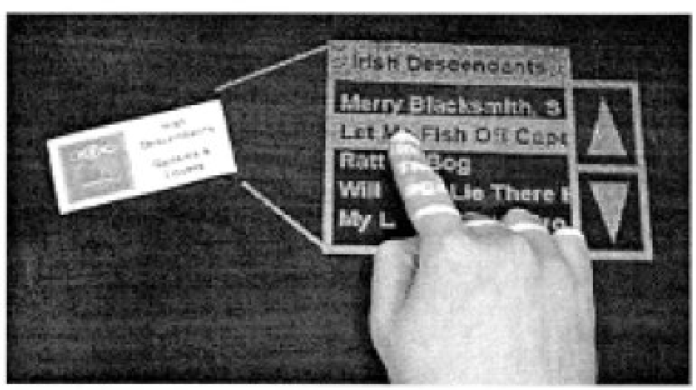


Figure 2.13 An example of live card, an interactive music selection menu is displayed upon recognizing the card [54].

Paper Windows transforms paper into a display surface that captures the paper affordances (see Figure 2.14). A blank paper becomes like a physical window which we can hold in our hands and we can move around the table to some extent. They introduce new interaction techniques to manipulate the paper display using hand gestures.

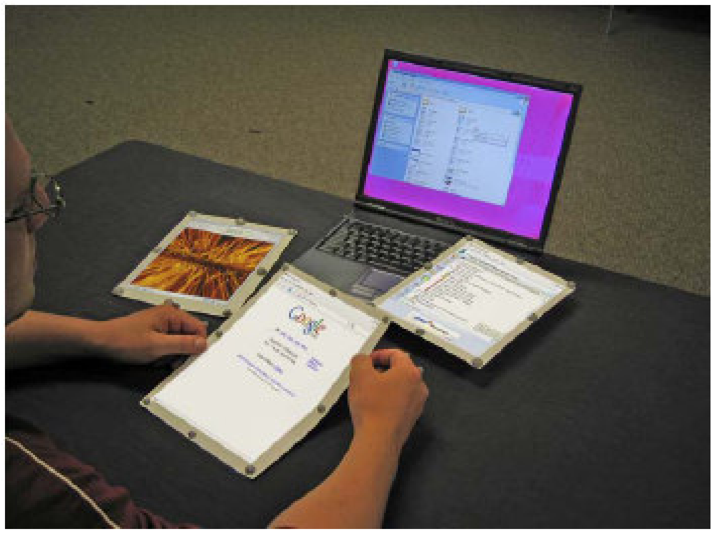


Figure 2.14 A snapshot of PaperWindows prototype in action, three pages are augmented with projected windows [28].

2.3 Tracking Physical Artifacts

Computer augmented systems use various recognition and tracking techniques that allow physical artifacts to be used for input/output. The two most popular methods are electronic sensors and computer vision.

Electronic sensors technology mostly uses RFID tags. A RFID tag is a small encapsulated electronic circuit that generates radio frequency signal communicated to a RFID receiver. These tags are widely used in industrial (tracking parts) and commercial (tracking items in stores) applications. An advantage of RFID tags is that they do not have internal power source, power is transmitted from the reader to the tag. They are also used in many augmented reality systems. Want et al. [60], in their interface system, used RFID tags for bridging the physical and virtual worlds. They attach tags to documents or books to link them to related electronic documents. The problem with using such tags with documents, particularly paper documents with small number of pages, is that their weight is relatively high, which makes it difficult to handle the paper document. In addition, this technology does not allow tracking precise location and rotation of the tagged object.

The other method for tracking objects uses vision-based techniques. Vision-based techniques can be split into two groups, using object features for recognition, and using visual identification tags attached to objects. These techniques use a camera to capture images of the target objects or work space. Then, image processing algorithms are applied to the input images (from the camera) to segment and recognize objects of concern.

Research in the computer vision community introduced several object recognition algorithms that use various types of object features extracted from input images. Mostly geometric features such as lines, points, curves, and corners are extracted from an input image; based on those features an object can be identified. This group of algorithms is suitable for non-real-time applications, and hence can not be used in augmented systems which are interactive in nature.

Another group of object recognition techniques use object appearance; the appearance of an object is directly compared to input images. Unfortunately, these techniques are also expensive in computation time to be used for real-time interaction. However, Murase and Nayar [45] proposed an efficient method based on component analysis of input images and demonstrated recognition in real-time. Instead of using whole image templates, Krumm [36] used binary vector quantization to represent binary features extracted from the input image. The technique was used in a computer-augmented environment to demonstrate its powers [37].

In addition to geometrical object information, color information has also been used for object recognition. The method is based on the histogram model of a color image. Histograms of input images are compared to histograms of registered objects for recognition. Color based techniques have the advantage of being rotation and translation invariant. Swain and Ballard developed the color indexing algorithm, an object recognition method that demonstrates the use of color histograms. This algorithm uses simple representations of object models which makes it suitable for real-time applications. For that reason the SnapLink system [47] uses the color indexing algorithm for registering and recognizing objects.

Several systems make use of hand gestures for directly manipulating objects and initiating commands. Koike et al. [39] developed a method for real-time palm and finger tracking. Using an infrared camera, human skin (with normally higher temperature than other objects) region can effectively be extracted from the image. The palm's center and fingertips are then determined using image processing techniques.

Other vision-based techniques use visual markers (barcode-like tags) attached to real objects for object recognition. The InteractiveDESK uses color coded tags to recognize objects and folders on the desktop, and maintains a link to electronic files. When the system recognizes an object on the desk, it retrieves and displays a list of linked electronic files previously linked to the object. In their first implementation of the EnhancedDesk, [38] used a specially designed two-dimensional matrix code for recognizing objects. Two-dimensional matrix markers allow identification of large number of objects, provide orientation information, can be recognized and tracked in real-time, and are cheap to produce. Rekimoto [52] devised a method for locating and decoding such two-dimensional visual markers. Due to limitation imposed by the video camera resolution, visual markers have relatively large size which makes it difficult to use on small objects.

IR Reflective tags are also used for identification purposes. PaperWindows, a novel system that turns paper into a display screen, uses a number of cameras surrounding the user's work environment to track positions of hands, papers, and pens, all of which are augmented with IR reflective markers [28].

Chapter 3

User Study

The studies reviewed in Chapter 2 primarily focus on understanding document management and organization in the real-world or current digital desktops to inform the design of tools for management of electronic documents. The main goal of our study is to examine the relationship between digital and paper documents and to investigate mechanisms people employ to maintain that relationship. We focused on knowledge workers in an academic environment as they often work with both types of documents (paper and digital).

3.1 Approach

We interviewed 15 researchers (9 students and 6 professors) from different university departments. Interviews took place in their normal work environment, with each interview lasting between 30 to 60 minutes.

The semi-structured interviews started by asking the subject how they organized their computer space, and whether they followed a specific pattern for structuring files and folders or used special document management tools (see Appendix A for the interview questions guide). We asked questions about their file hierarchies and desktop arrangement to understand how users kept the virtual space organized. Following this we asked about the organization of their workspace, in particular on their desk space. We noted the number of piles users had and how they categorized documents into various piles. We observed the participants as they were asked to find the digital version of a paper lying on their desk and vice-versa to understand their organization of information and how they relate the two document types. At the end of the interview we described the notion of an automated link between the two document spaces and asked them to envision positive and negative traits of such a link.

We discuss our findings by first presenting our observations on the organization of the physical and virtual spaces, then presenting our observations on the management of the link between the two spaces.

3.2 The Physical Space Organization

Our focus is on the "active" documents (referred to as "hot" and "warm" documents by other researchers [57, and 12]) that people deal with on a daily basis. Almost all of the participants used "Piles" as the main organizational unit for their active documents on the desktop. This also agrees with the findings of previous studies [43, 44, and 12]. Several users used filing, but mainly for archiving purposes. Also, some participants used folders to keep piles of papers from mixing and others stacked different piles in a criss-cross fashion.

We observed that while some desks had very few papers (almost empty) yet others were full of piles. We identified two types of these piles: coherent and incoherent piles. A coherent pile consists of a coherent set of papers that share a common criteria or purpose. Such piles are created by explicitly categorizing papers into logical groups. An incoherent pile contains papers of mixed content or the grouping criterion is not clear. Such piles are created by collecting items into a stack without consideration for any logical grouping. Also, all of the participants had a pile of "hot" papers which are currently being used and have higher priority than other papers on the desk. Another common type of pile observed in the study is a pile of new information waiting to be categorized.

Depending on the balance between the number of coherent and incoherent piles we could classify the desk organization from *weakly* to *strongly-structured*. The weakly-structured desk consists of all incoherent piles that do not represent a clear classification of desk content while the strongly-structured desk consists of coherent piles that each represents a particular set of papers.

We found that most participants had a mix of coherent and incoherent piles. Three participants had weakly-structured desks, one had a strongly-structured desk and the rest (eleven participants) used a mix of both types of piles with more emphasis on one type or the other.

For example, one participant (an example of strongly-structured desk) had one pile for currently used papers, another for related papers of a certain project, and a third one for personal documents. Each pile grouped documents which shared specific criteria. Another participant with a semi weakly-structured desk (with more incoherent than coherent piles) had several piles scattered around the desk space and only one of these piles actually represented a meaningful group of papers.

People with a weakly-structured desk had more difficulties finding documents on the desk. They relied mainly on spatial memory to find a paper, the mixed content of their piles made it difficult to use other information like category for example. Users with strongly-structured desks faced less difficulty because their piles were better categorized and well separated. The clear distinction between the piles aided their memory in finding a certain paper.

Users also noted that they often started with a clean desk that got cluttered over time. Occasionally, users would perform a major cleanup of their desk space. However, we believe the levels of organization observed in the study reflect their organization style.

3.3 The Virtual Space Organization

In the virtual space, we focused on how people used the file system hierarchy to structure their information. Since the virtual space allows only filing of information, we observed the extent to which people kept their information organized and the methods they used to retrieve it.

Similar to the physical organization of documents, we classify the organization of the virtual space as a continuum from weakly to strongly-structured virtual workspace. The weakly-structured workspace is characterized by a loose structure of the space due to lack of a consistent structuring scheme. Folders and subfolders are created with no specific pattern, documents are poorly categorized, and many out-of-date or unused documents are scattered over the workspace. The strongly-structured workspace can be characterized by a coherent organizational scheme for most files and folders. Usually a person of this type of organization would create a folder structure that makes sense to their own type of work and tasks. Categories and subcategories are carefully created and placed in meaningful locations depending on a variety of factors (e.g. project, course, research etc.) specific to the person [27].

In general, we found participants had a better organized virtual space than the physical one. Also, majority were in the middle of the scale but closer to the strongly-structured. Three users had a strongly-structured workspace, and only one had weakly-structured workspace.

In one case, folder hierarchies were not used effectively; most files were placed in the root directory and the few categories created did not follow a systematic classification pattern. The user in this case relied on search tools (local and internet search) to find documents. In her physical space, information was better organized with papers grouped into coherent piles on the desk or kept in a cabinet for saving space on the desk. On the contrary to that case, a highly experienced user had a clearly defined structure which he used for several years to organize his documents. This user had fewer problems in the digital space and used location-based search for finding information, but his physical space was totally the opposite with many incoherent piles stacked over the desk.

Users with weakly-structured workspaces (both physical and virtual spaces) spent more time retrieving information than those with a stronger structure. As found by other researchers, most users preferred location-based search over using search tools to find information [7].

3.4 Comparing Virtual and Physical Spaces

We classified all users according to their levels of organization in each space. They were assigned subjective scores in the range 1-5, with 1 being weakest and 5 being strongest structure (see Figure 3.1). There is a trend that people with strongly-structured physical space tend to have a weaker structure in the virtual space and vice-versa. Of our sample, 8 people had strong organization in one space and a weak one in the other.

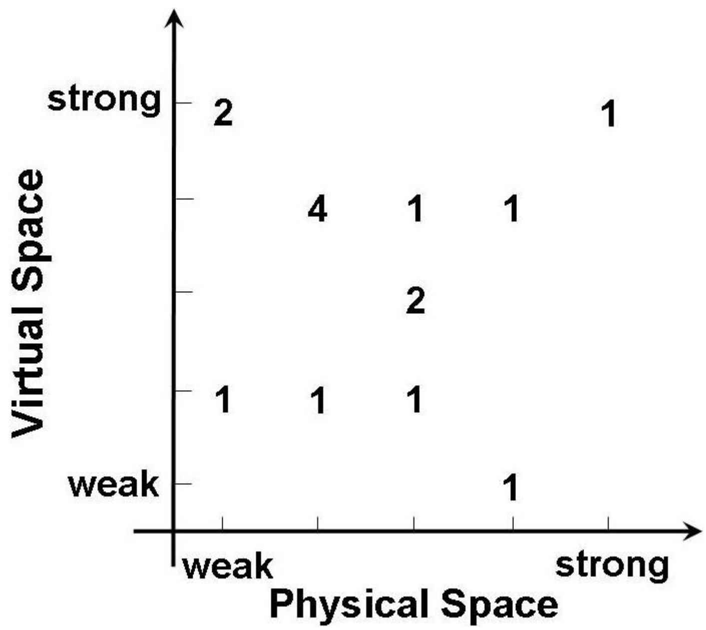


Figure 3.1 The classification of users' levels of organization in the physical and virtual spaces.

3.5 The Virtual-Physical Link

By the Virtual-Physical link we refer to the means through which a person relates a physical document to its digital version and vice-versa. This is a two-way link often juggling between the two spaces.

We found that participants were not conscious of the link; but developed different coping mechanisms to deal with it. While several participants relied mainly on a mental model for the two work spaces allowing them to identify paper documents from the corresponding digital ones or vice-versa; others, in addition to the mental model, followed a certain scheme that automatically represents a connection between the physical and virtual documents. Only one participant created an explicit link by writing comments, pertaining to the category/project the paper belongs to, on the first page of the printed paper.

From these observations we can classify the link between the digital and paper documents based on the strategy employed by users on a scale from weak to strong link. The distinction between weak/strong links is determined by two factors. The first factor is the strategy invested by the user to maintain the link. The other factor is the level of organization that the user employs in both workspaces (weakly/strongly-structured). As mentioned above, we found that people employ three strategies: memory based, using a file naming scheme or attaching marks to the printed document.

The memory based strategy represents a weak link when combined with a weakly-structured workspace, but when combined with a strongly-structured workspace the link is made stronger as the user can easily remember the location of a document in a well structured space. The other two strategies also involve memory usage, but they support it with extra information saved on either the paper or the digital side (some people may add information to both media, but none of the participants in our study did that). These two strategies represent a stronger link depending on the relevance of information attached to the document and the organization of the workspace (both the virtual and physical spaces).

For example, one participant, with a strongly-structured virtual space, used the full title of the document to name his files and had no difficulty in retrieving the file when given the paper version. Another participant, with a less organized virtual space, used the category of the document with the addition of some numbers to name the file; this scheme was not effective and the participant found it difficult to locate the electronic file for a randomly selected paper from a pile on his desk.

We must note, however, that these strategies provided only a one way link, from paper to digital. For the other direction of the link, from digital to paper, all users relied on their memory to find the needed paper. The more recent papers were easier to find on the desk because they were easier to remember.

We found that people with strong structure in at least one of the spaces, have less difficulties than people with a weaker structure of information. A strong structure is one that follows a systematic scheme and allows easy retrieval of items. Importantly, in such an organization scheme the user is dedicated to place items in their appropriate category, otherwise the categories would get out of order and the structure starts to get loose.

As an example, one participant had a simple but effective categorization of his electronic files and folders, but his physical space organization was relatively loose and was having hard time finding a paper. When he couldn't find a paper, he would simply reprint it since it was relatively easy for him to find the electronic version of the paper. Another participant had a coherent paper space but not so for the electronic space, so it was easy for her to find a paper in the physical world, but in the digital space she mostly relied on search tools to find a digital paper.

These observations suggest that people develop work-around strategies (by reprinting papers or resort to working exclusively with one type of media) to cope with the absence of the virtual-physical link. These strategies require cognitive effort to maintain and cause people to spend more time than necessary to find a document from either workspace. They also deprive people from taking advantage of the benefits of both spaces [49].

3.6 Requirements for a Media Bridging Tool

Results from the above study suggest that users in different levels of organization and link maintenance would benefit from a media bridging tool. On one hand, people classified as having a weak structure in one or both of the two spaces can benefit by obtaining a stronger link. On the other hand, it will reduce the effort that strongly-organized users usually have to make to maintain their strong structure.

Current strategies allow the user to retrieve a digital document by looking at the paper version, but can not support the other way, i.e. find a paper in some pile on the desk. A bridging tool should *support both transitions* and should provide accurate information about the *document location* in both spaces.

Most users did not use special document management tools because they felt that the effort they had to make was not worth it and such tools often hinder with their work practices. This implies that the bridging tool should provide *light-weight interaction* and *not interfere with their work practices.*

Participants often shuffled their desk contents around, this caused them to loose context of a previous scene, bringing back that context can help them retrieve valuable information attached to it. There is a need for *supporting history of pile arrangements* on the desk.

Our observations of the variety of organization levels lead us to believe that users need a tool that is flexible enough to allow them to use it in a personalized way. It is important that the bridging tool should be *general and flexible* to accommodate a large spectrum of users.

Chapter 4

PaperSpace Design

PaperSpace is a media bridging tool that maintains the link between paper and digital documents. Its design was influenced by the user study in chapter 3. We built PaperSpace as a proof-of-concept tool to test the feasibility and usability of such tool.

In this prototype we work with a subset of documents; the type of documents a typical researcher collects into a list of references (or bibliography). We chose this type of documents in our prototype for two reasons. Firstly, we needed to evaluate the system in a real implementation where the test subject deals with many papers simultaneously. Since we work in a research lab, it is relatively easy to find researchers working with large amount of papers on their desks. Secondly, we needed to use a domain that we already understand (research papers). While we test PaperSpace in this specific domain, we hypothesize that this system will work for other document types as well.

The main idea of the system is to keep track of paper copies of electronic documents and use their arrangement on the desktop to naturally organize, group, and represent the corresponding digital documents. Document tracking is achieved by using a computer vision based tracking system, described in Section 4.1 (*Document Tracking Component*). This maintains the link between the two document spaces with minimal user effort. We want to explore the usability of such a system and whether it would actually support the integration of paper and digital workspaces.

4.1 System Overview

As the user prints documents and places them on his desk, the system recognizes the documents and constructs an internal structure of paper piles. Papers are recognized using computer vision techniques to read special 2D tags printed in the margins of the paper. The piles of papers are represented in a virtual desk view that shows the piles' top papers. A pile can be moved, split or merged with another pile. When the user is working with a digital document, he can easily find its paper version by asking the system to locate it on the virtual desk view. Likewise, a digital document is easily retrieved by clicking on its paper version displayed in the virtual window.

The system provides multiple views that represent the desk paper arrangement. The main view provided on the screen is the virtual desk view, in which piles of documents are represented by icons of their digital copies. This virtual desk view, allows the user to search the desk contents electronically without having to scramble through their physical papers. As an alternative to the virtual desk view, a pile list view, that displays the title of each document in the list, allows browsing through the contents of each pile. Section 4.4.4 on *Document Views* elaborates on how the views are built and presented to the user.

In order to allow the user to query the system for a specific document, multiple search methods are provided, visually searching the simulated desktop, searching by some criteria, or searching lists of previous piles of papers. On the user interface, the user can search for a document by some criteria, or directly browsing the virtual piles as describe above. Section 4.4.3 (*Finding Documents*) describes the different search methods possible through the system.

The user also needs a way to add documents to the system with relative ease. The system provides a few methods for adding documents. Printing a document through the system is necessary for the document to be trackable. This also allows the system to provide interesting information about documents to the user, such as how many times a document was printed, by whom, and the location of a document (i.e. on which desk is it seen). Adding and printing documents is discussed in Sections 4.4.1 and 4.4.2 respectively.

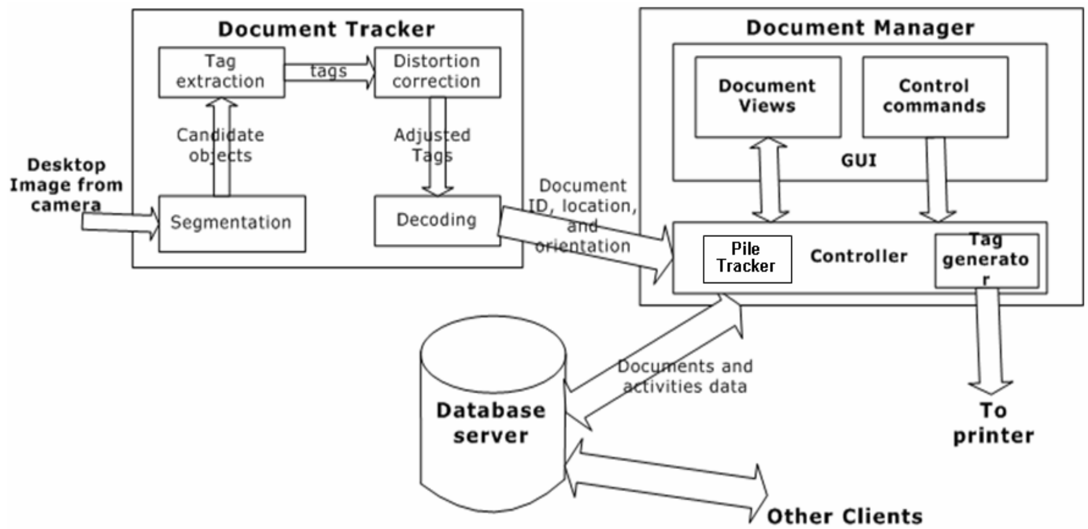


Figure 4.1 System architecture diagram

The system is built for a Microsoft Windows XP platform and consists of four main software modules (see Figure 4.1): a *Document Tracker* to recognize and track documents, a Pile Tracker that assembles papers into piles, a *Document Manager* to enable the user to control the system, and a *Database* for storing document information and user activities. The Document Tracker was implemented using C++ and openCV (Intel's computer vision library), while all other modules were written in Java JDK 1.5. The results of the Document Tracker are sent to the Pile Tracker using memory mapped files (using Java Native Interface technology). In the following sections we describe the system modules in more detail.

4.2 The Document Tracker

The set up consists of a high-resolution PC camera mounted above the desk so that it covers the active document area as defined by the user (see Figure 4.2). Chapter 5 describes our evaluation system setup in more detail. The Document Tracker was implemented using C++ and OpenCV to capture frames from the camera and to perform certain image processing operations. Each frame/image contains the desk contents of papers and other objects. Paper documents are distinguished with special tags printed in the margins of the paper which can be identified by image processing techniques described below.

Figure 4.2 Desk and camera setup

4.2.1 Visual 2-Dimensional ID Tags

These are 2-dimensional visual codes (i.e. encode information in 2 dimensions) that carry the paper ID. They are generated automatically by the system whenever a paper is printed and added to the paper's corners. These tags were initially designed by Jeff Smith [58] and are similar to the recent QR codes (Quick Resonse codes are 2-dimensional bar codes, http://www.denso-wave.com/en/index.html/) used for tracking parts and inventory management in a wide variety of industries. Our code is tolerant to image distortions. It uses region adjacency relationships to extract sub-regions of the code.

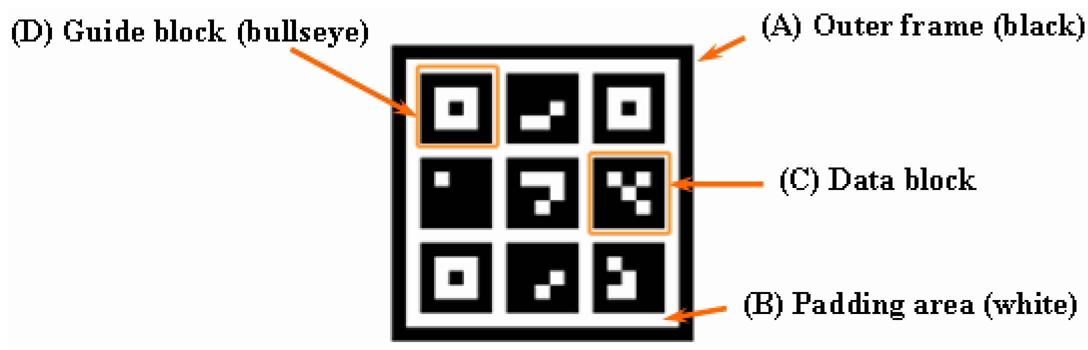


Figure 4.3 A sample of the visual tag used to identify the document in the PaperSpace system.

The tag (see Figure 4.3) is comprised of the following regions:

- An outer black region (ring).
- A white padding area.
- An odd number of blocks inside the padding area (9 blocks are used currently). Each of these blocks has a black ring that encapsulates the contents of the block. There are two types of these blocks:
    - Guide blocks (bullseye). There are always three blocks of this type in three corners of the tag. These blocks with the outer black ring and the padding white region make it possible to identify a tag from other false objects in the image. Also these blocks allow reorientation and distortion correction of the code.
    - Data blocks. These blocks carry the data necessary for the recognition of documents. Each block carries a maximum of 8 bits of data (one bit is always black to distinguish a data block from a guide block) for a total of 48 bits per tag. Bits of data are represented with on/off (white/black) pixels in the block; each bit is represented by a bit-module which is one pixel in the original tag but can be larger in the camera image. A data block's outer ring can have a maximum height

of 2 (i.e. a depth of 2 in the region adjacency graph); while a guide block always has height 3 (height means the number of child/sub-child regions inside the block).

We print four tags on each paper. The multiple tags on each paper improve the recognition rate and make it possible to recognize partially visible papers. Other systems proposed the use of document image matching (i.e. uses the document's features to match between the image stored in the database and the camera image) to identify the document [34 and 54], but this solution requires more processing and is slower which makes it unsuitable for real-time recognition. Also, their system puts too many constraints on the user's interaction with papers on the desk.

Each tag carries a message that contains the ID of the paper, the corner number (from 0 to 3 starting from the top left corner of the paper and increasing in clock-wise direction), and error correction codes. A document is identified by its ID number which is created by the database. When a document is printed, a new number is generated to identify this specific copy of the document, so there can exist multiple prints of the same document with each being uniquely identified. The tag's corner number (which indicates the corner of the paper the tag is attached to) allows the system to determine the position and orientation of the paper with only a single tag visible. We use RS (Reed-Solomon) codes for error correction. RS error-correction works by over-sampling a polynomial constructed from the data. We use a code word byte size of 4 bits and a maximum of 4 errors can be corrected by the RS decoder. This configuration produces an RS with 15 code word symbols, 8 of these contain parity codes and 7 contain data. RS error-correction codes occupy 32 bits which leaves 16 bits for the actual data bits (3 data code words are not used). Hence we can encode up to 2^16 or 65536 different codes. Data bits start at the beginning of the tag and followed by the error-correction codes.

4.2.2 Finding and Decoding the Tags

The document tracker receives full color images of the desk surface captured by the top mounted camera. Our goal is to locate and read the ID tags to recognize the papers currently on the desk with a reasonable speed. In order to minimize the image processing time we needed simple and efficient algorithms to identify and decode the tags. Since the tags have a square shape, we used that information to locate their initial locations on the image.

This is how the algorithm works for each camera frame:

1. The first step is converting the image to a gray scale image (i.e. convert the image from 3-channels to single channel). This step reduces the size of data to be processed and hence reduces the processing time.
2. Obtain an edge image by applying the canny operator from the openCV library (Intel's computer vision library) (see Figure 4.4). This step results in a binary image that consists of black background and white lines representing the objects' edges.

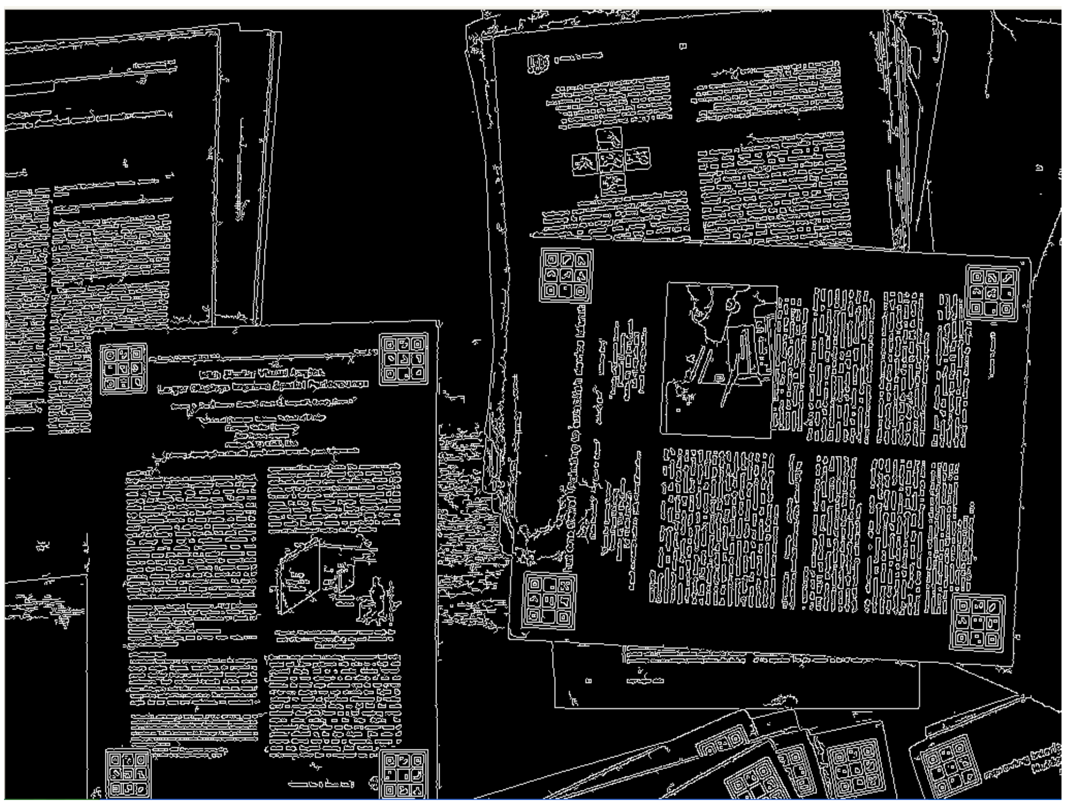


Figure 4.4 Sample Image of the result of applying the canny operator to the original image.

3. Extract all of the connected components from the edge image. Eliminate components based on preset size and geometric shape (height/width ratio) thresholds that are relative to the image size. This results in a list of candidate tags in the image.

4. Extract a sub-image from the original image for each candidate tag that resulted from the previous step. This sub-image is larger than the tag size to accommodate for rotated tags (i.e. tags with non right-angle orientations) (see Figure 4.5).

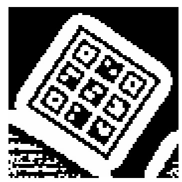


Figure 4.5 Threshold sub-image of a candidate tag. The extracted region is bigger than the tag to compensate for various rotations.

5. Threshold (i.e. convert to black-white image) the sub-image using the adaptive threshold algorithm from openCV library (Intel's computer vision library).

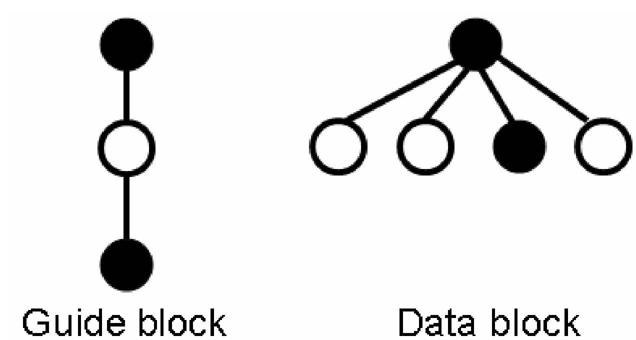


Figure 4.6 Examples of region adjacency graphs for a guide block and for a data block.

6. Build a region adjacency graph (see Figure 4.6) from the binary sub-image using a flood filling algorithm; this step allow us to both eliminate false tags

and extract location and orientation of the real tags. From the region adjacency graph we look for a certain pattern. If the pattern is found then indeed the sub-image contains the tag we seek, otherwise it is discarded.

7. Reaching this step means the sub-image contains a valid tag. Using the region adjacency graph we find the pivot of the bull's eyes and the orientation and size of the tag. Also the bit-module size is estimated by summing the pixels in the whole tag, dividing by the number of pixels of the original tag (when created), and taking the square root of that result.

8. Rotate the sub-image to the standard orientation (as shows in Figure 4.3) and extract an **exact image** of the tag (see Figure 4.7). The tag size is estimated from the outer frame that encircles the tag (region A in Figure 4.3).

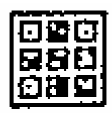


Figure 4.7 An exact image of the tag after rotation and thresholding is applied to the original sub-image.

9. Build a region adjacency graph of the rotated tag and sort the data blocks to be read in the proper order. This step can be avoided by using the rotation information along with the previous region adjacency graph to read the data blocks directly.

10. Read the data blocks in order and group the data bits in a single message. For each data block do the following:

    a. Estimate the center of the data block based on the width, height, and position of the block within the tag.

    b. Using the previously calculated bit-module size, find the center of each bit-module.

c. The value of a bit-module is computed by taking a weighted average of the pixels within a defined neighborhood around the center of the module. The weight of each pixel is estimated from its distance to the bit-module's center as follows:

   - pixel weight = *k* / distance;

   The constant *k* ranges from 0 to 1. We used the value 0.5 which gave the best results in our implementation.

11. Pass the extracted message to the RS decoder to find and correct reading errors. The RS decoder processes each block and attempts to correct errors and recover the original data.

12. The RS decoder will either return the corrected message or return a flag indicating that the message can not be corrected.

13. Finally, the retrieved message is sent to the document manager module which will check the database for a matching document.

The tracking accuracy depends on several factors. Camera resolution, size of the trackable desk surface, lighting intensity and size of the tag should all be considered to achieve acceptable recognition rate. Section 5.1 on Technical Evaluation elaborates on these factors.

4.3 The Pile Tracker

The *pile tracker* processes the set of recognized papers that result from the *document tracker's* processing of each frame. The tracker keeps a list of current piles and current papers on the desk. We defined a paper object and a pile object. Each paper object keeps information about its location, orientation, pile, tags and document related information (e.g. ID, title, file location etc.). The pile object contains its location, orientation and a list of ordered paper objects.

The *pile tracker* allows adding new piles, moving papers from one pile to another, moving the whole pile as one unit, merging two piles into one, splitting a pile and removing a paper from a pile.

We identified three main events that occur on the desk scene: new-paper (a paper that is not currently in the desk scene), paper-moved and paper-removed (a paper that is in the current desk scene but not visible in the new list of papers). There are also other events that result from the main events: paper-surfaced (when a paper deep into a pile is revealed) and paper-covered (when a paper is no longer visible).

The tracker determines these events by comparing the current desk scene with the new list of papers. The desk scene consists of all the piles and papers on the desk while the new list of papers is the list of recognized papers that arrive from the *document tracker*. Only papers that are visible in the desk scene are considered in the processing of the new desk scene. We process these events as follows:

1. New-paper: a new paper is detected when a paper is recognized and can not be found in any of the current piles. Each new paper eventually ends up in a pile as follows:
    a. Paper added to an existing pile. This is determined by computing overlap of the paper with other piles on the desk. An overlap threshold determines whether the paper should be added to a pile. If the overlap ratio is greater than 0.25, the paper is added to the pile.
    b. Create a new pile and add the paper to it. We assume here that the paper is placed in an empty space or on top of other untagged objects. This is done when no overlaps are detected or the overlap is smaller than the threshold value.
2. Paper-moved: this is detected when an existing paper moves over the desk surface. For the paper-moved event there are two cases:
    a. Single paper moved: for a single paper-moved, the processing is similar to that of the new-paper. A single paper move is detected when its pile contains other papers. If the pile has only this paper, it will be processed as in the whole pile moved.
    b. Group of papers moved: this requires detection of whether the whole pile moved or the pile was split. We assume that if any of the lower

papers in the pile become visible then the pile is considered to be split, otherwise, the whole pile has moved. First if the *whole pile moved*: we check its occlusion against the other piles in the desk scene. If no occlusions found, the pile position and orientation are updated. In the case where the pile occludes another pile, the moving pile is added to top of the bottom (occluded) pile. Second is the case of *Pile split*: the stack of papers that moved is removed from the pile. Then we check the overlaps with other piles as in the previous step (whole pile moved)

3. Paper-removed: a paper is removed from the desk if the paper below it becomes visible and the paper itself is no longer seen. Currently, we can only detect removal of a paper if there is another paper below it. Otherwise, paper is not removed from the desk scene. The difficulty is in distinguishing the event of removing a paper from the desk from covering the paper with another object (an untagged object). The same applies to a whole pile being covered by an object. We solved this issue by creating a list of uncertain piles. This list keeps a record of piles that can not be seen by the system (piles the system thought were removed) and acts as a store for recent piles in the system. The user can manually select and remove such piles.

The desk scene is saved to a file periodically to be able to recover in case of a system crash or if the user turns the system off. Papers location, orientation and pile information in the database is updated periodically as well. This allows the system to support viewing of previous desk scenes which may be helpful in reminding the user of recent tasks and context information.

There are two limitations on the user's interaction with papers on the desk. First, a paper can not be pulled out of a pile (only the top paper can be moved, either alone or with other papers below). However, there is no restriction on the number of papers moving at the same time. Second, the system can not detect removal of a whole pile (that also applies to a single paper pile) from the desk. To solve this issue, the system can use information of objects movements in the image to figure out whether a paper has been removed or covered by some object.

4.4 The Document Manager

This component includes a controller for processing user commands, updating document views, and communicating with the pile tracker, and a GUI component (see Figure 4.8). The GUI provides several views and actions that allow the user to flexibly control and manipulate their documents. There are a large number of features that can be included in the system. However, for the purpose of this project, the prototype we developed is for evaluating the system design, and hence will include only the more important aspects considered essential for the evaluation.

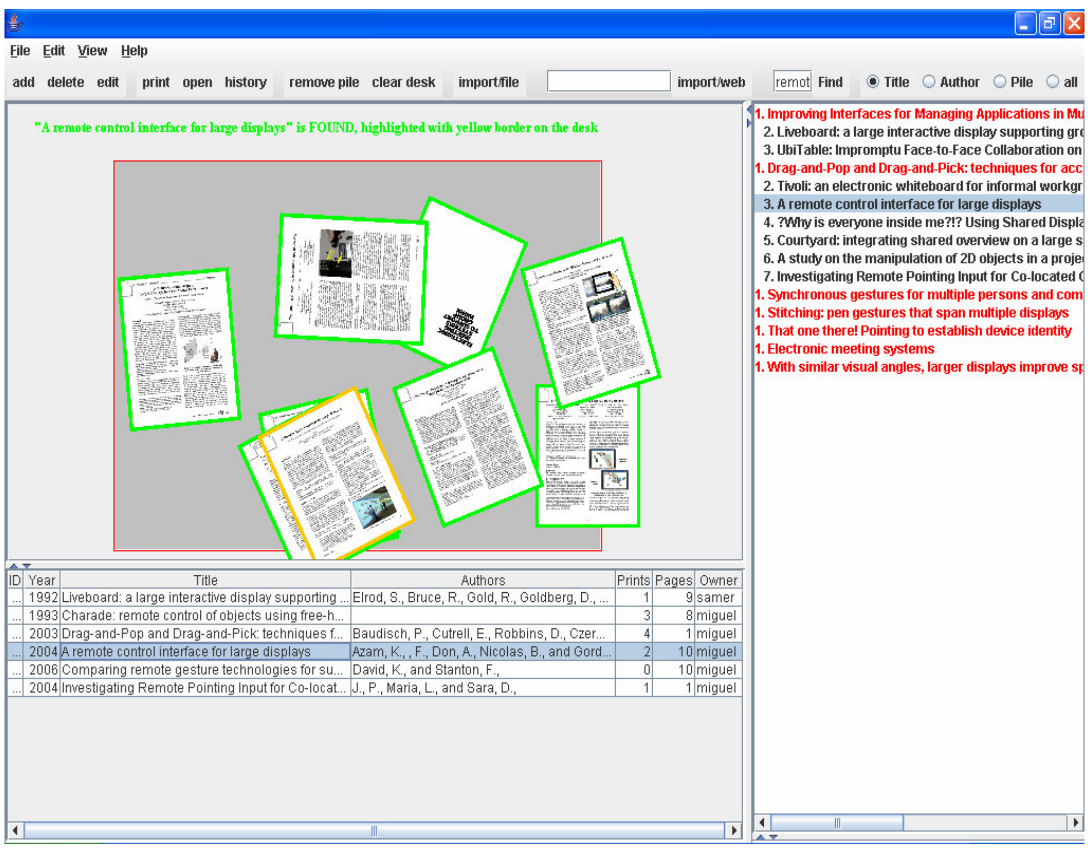


Figure 4.8 A snapshot of the document manager's GUI showing the main 3 views: Pile List View (right pane), Query List View (bottom left), and the Desk View (top left). This snapshot shows a paper highlighted in all views.

In this section we describe how the controller works, and the main features of the user interface.

4.4.1 Adding Documents

This is a critical issue. If the process of adding documents to the system is not user friendly, the user may be inclined to not use the system. The system allows the user to add documents using several methods:

- Manually add the pieces of information about the document. This is necessary in case there is no electronic information available about the document.
- Copy/paste the web address of an ACM digital library's page (http://portal.acm.org/portal.cfm/) of the document. The ideal system should support most of the digital libraries that are available on the internet. Also, in order to reduce the effort on the user, there should be a link added automatically to the digital library's page, where all the user has to do is click that button to add it to the system.
- Copy/paste the document information in the EndNote format (http://www.**endnote**.com/). An alternative is to allow the user to import such information from a text file. Also, the system should be more general by supporting most of the available bibliography formats (BibTeX, RIS, etc.) and allowing the creation of user defined formats.
- Automatically add documents listed in a text file in the EndNote format. This option is important for importing long lists of documents from other document tools.

4.4.2 Printing Documents

For any document to be part of the system (i.e. to be identifiable) it must be printed via the system so that the visual tags are printed on the document. Whenever a print job is requested, the system tells the user if any previous copies have been printed along with their last seen locations (see Figure 4.9). This has the potential of reducing unnecessary reprinting of documents already on the user's desk.

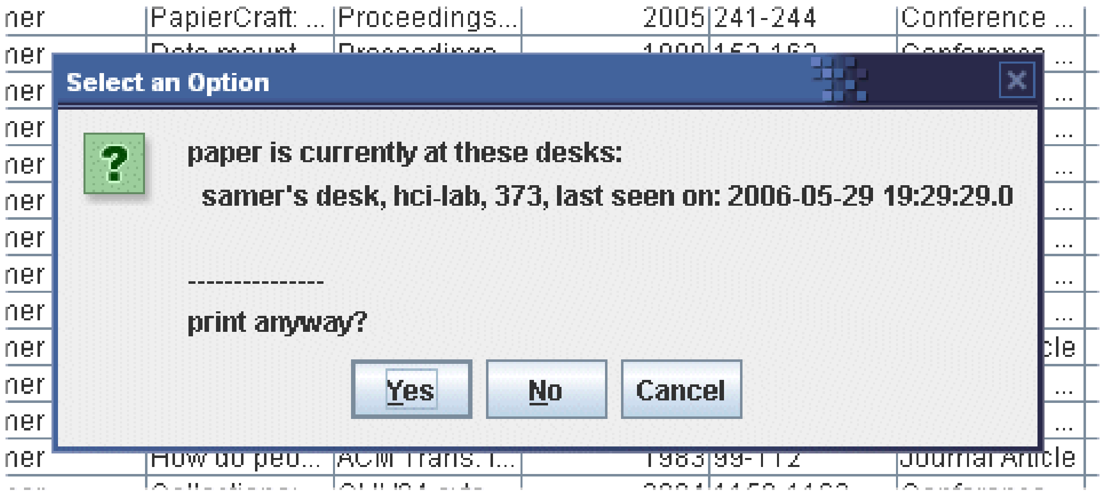


Figure 4.9 Sample of the print confirmation dialogue; this shows on which desk a copy of the paper can be found and when it was last seen.

4.4.3 Finding Documents (search methods)

An important functionality of any document management system is to make it easy for the user to find and retrieve documents. Our system provides interesting alternatives for finding a document:

- Typical text search on some criteria of the document of concern. We implemented search by title, author, and pile. The user types part of or all of the title, author name, or pile attribute. The system searches for partial matches of the entered text. The pile search option searches the database for previously created pile attributes/properties (see Section 4.4.4.2) which allows the user to retrieve previously grouped documents.
- Browsing through groups of documents that are grouped according to their arrangement on the real desk.
- Retrieve an electronic document by showing the paper copy to the camera.

- Locate a paper document on the desk by selecting its representation in one of the document view provided by the user interface.

4.4.4 Documents Views

4.4.4.1 Document Query List View

This view displays a list of document records (see Figure 4.8). This list is populated when the user executes a search query. The result could contain multiple documents ordered by relevance to the search criteria. Each document record includes the following fields: publication year, title, authors, number of prints, number of pages, and the document's owner. This list can be ordered by any of these fields. The user can open a document by double clicking on its record. This list is also linked to the Pile List View described in the next section.

4.4.4.2 Pile List View

All piles on the desk are displayed in this view (see Figure 4.8). Each pile is represented with a vertical list of its documents ordered from top according to their location in the pile. A pile list has two levels, the top document and the documents below. The top document is displayed in red color to mark it as a new pile. Other documents in the pile are displayed in black color and are indented to make the list of piles more readable. Each document is represented by its title. The purpose of this list is to show all of the documents in the pile at the same time, which is not possible with the current implementation of the Desktop View described below.

This view is linked and synchronized with the Desktop View. When the user selects a document from this list, the corresponding paper in the Desktop View is highlighted and displayed on top of all other papers.

To find a paper on the desk, the user can perform a search query which may return a list of documents in the Document Query List View. The user can then select one of these documents and the system displays a message indicating whether that document is on the desk or not. If the document is on the desk, its representation will be highlighted (in both the Desktop View and the Pile List View).

Alternatively, the user can search for papers that were in the same pile with other papers. This method is useful when the user does not remember the title of a paper, but the category it may belong to or other papers it may appeared with in the same pile. This is possible by assigning attributes/names to piles, either manually or automatically. The user can create different piles and assign an attribute to each. Later, the user can search these attributes to retrieve a recent group of papers. This feature allows quick tangible way of creating documents' categories.

4.4.4.3 Desktop View

The desktop view is a virtual window representation of the real desktop (see Figure 4.10). Thumbnails of documents are displayed in relative locations and orientations to those of the paper documents on the desktop. When a user searches for a document using text search, the document can be highlighted in this view to show its location on the desktop. The user can also use this view to directly open a document in the default application. The view is re-sizeable and can be hidden as well.

There are other alternatives for showing the location of papers on the desk. A fancy, but costly, alternative is to install a media projector that will show paper's position directly on the desk. More on this issue in the Future Work section (see Chapter 7).

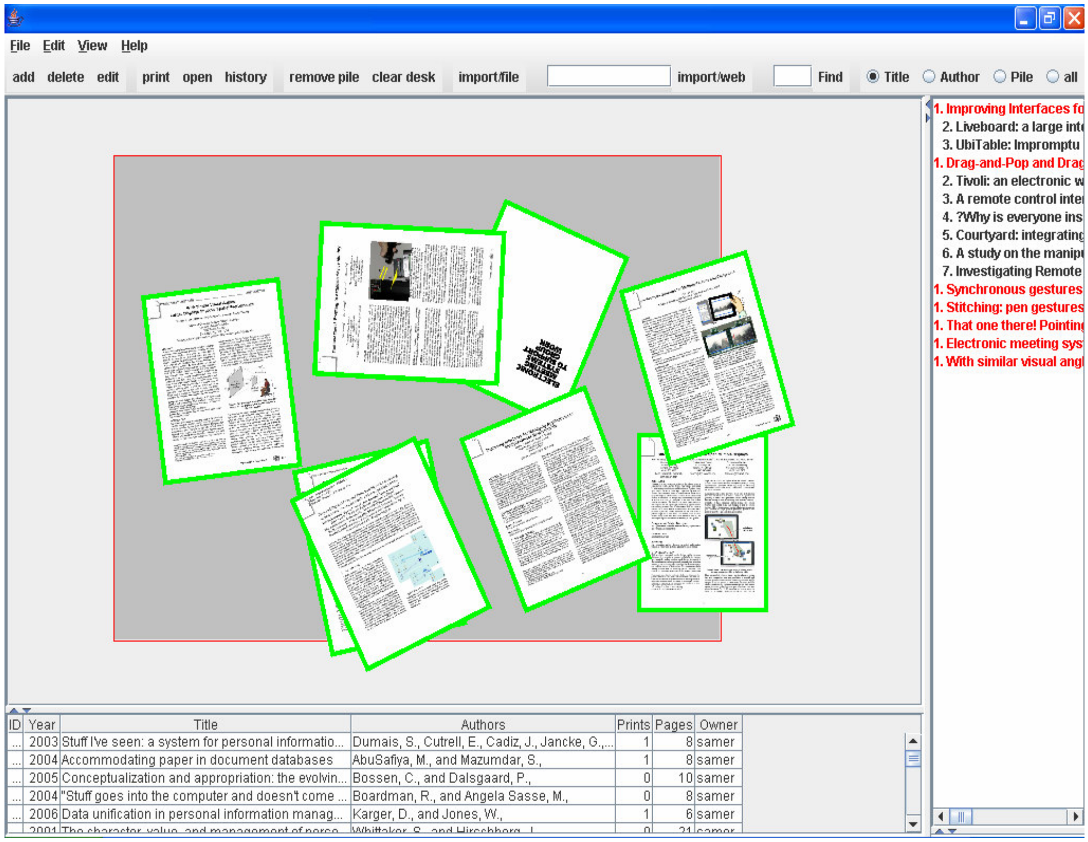


Figure 4.10 The Desktop View (top pane) displays thumbnail images of documents identified on the real desktop.

4.5 Database design

The purpose of the database is to store the required information about documents and their references and to keep record of paper and pile activities on the desk surface. Information stored in the database includes documents' reference information (authors, type, publication date, proceedings, year, etc.), the document's location (in the computer), printing information, and document's location on the desk, document piles, and user actions (see Appendix B for a diagram of the database schema).

We implemented the database using a MySql engine (an open source relational database system, http://www.mysql.com). The main tables in the database are: documents, users, desks, and authors. These tables are interrelated with a set of relations: prints, prints-locations, views, and documents-authors.

All database logic is implemented on the server side in stored procedures for improved performance. Currently the database contains 22 stored procedures that handle adding, updating, deleting, searching and retrieving data.

Information of paper and pile locations must be kept in a consistent state. In this system –where the information depends on a recognition process that has a degree of uncertainty- we can use confidence measure with each piece of paper location information. This confidence value reflects the level of certainty of that information.

Chapter 5

User Evaluation and Technical Performance

The purpose of building the system described in Chapter 4 is to test the success/failure of a media bridging tool for paper and digital documents. The success of this tool relies on the user's approval of it. If users do not find it practical, there is no point of using it. As such, the tool's performance and user's interaction with it are important properties to evaluate.

Therefore, we ran two studies to evaluate these properties. The first study evaluates the technical aspects of the system such as the recognition accuracy and pile tracking. The second study evaluates user's interaction with the system by deploying the system for a two-week period in the user's own office.

5.1 Technical Evaluation

We aim to find performance measures for the efficiency and accuracy of recognizing the visual tags, and tracking of piles on the desk. First, we describe the system setups followed with the results of tag recognition and pile tracking.

5.1.1 System setup

We used a high resolution USB 2.0 camera (DNV3001) from Digital Network Vision, LLC. The camera produces 1600x1200 color images at 10 frames per second. The system was run on a Dell Precision Workstation 650 with a 3.2GHz Intel Xeon processor and 1.5GB memory. The document tracking module was coded in C++ and openCV (Intel's computer vision library). The camera was mounted 68 inches (172 cm) above the desk. The tracked area of the desk was approximately 30x40 inches (76x101 cm) which accommodates up to 14 A4-size papers (see Figure 5.1).

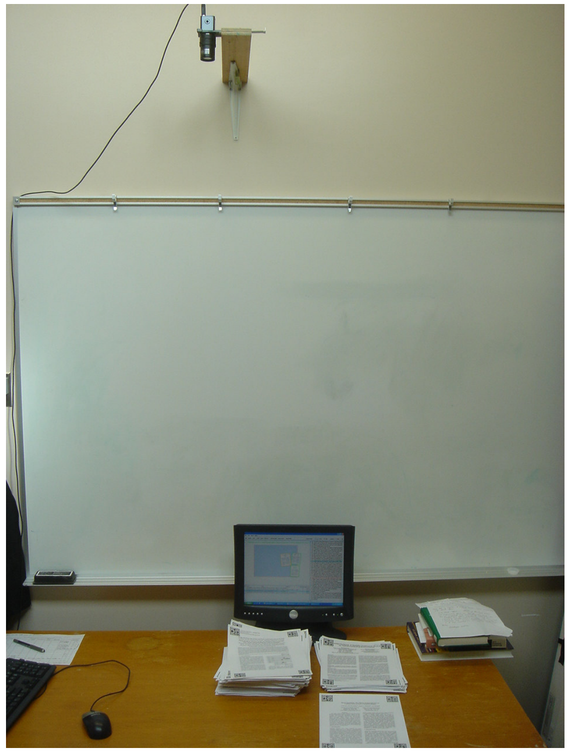


Figure 5.1 PaperSpace setup for the technical evaluation study

For both test-1 and test-2 the software automatically recorded data over 80 frames. The data included the pixel-size of each recognized tag and the following numbers: initially detected tags (may include random square-shaped objects), false/rejected tags (these include distorted tags and other square-shaped objects), true tags (identified as valid tags), true but unrecognized tags and recognized tags.

5.1.2 Test-1

This test was designed to find the minimum size of the tag that can be recognized accurately. For this we printed 16 same-code tags of different sizes on a sheet of paper. 10 of these papers, each having a different tag code, were placed on the table such that all the tags were visible to the camera and all had the same orientation. In total there were 160 tags visible to the camera in all of the frames (see Figure 5.2).

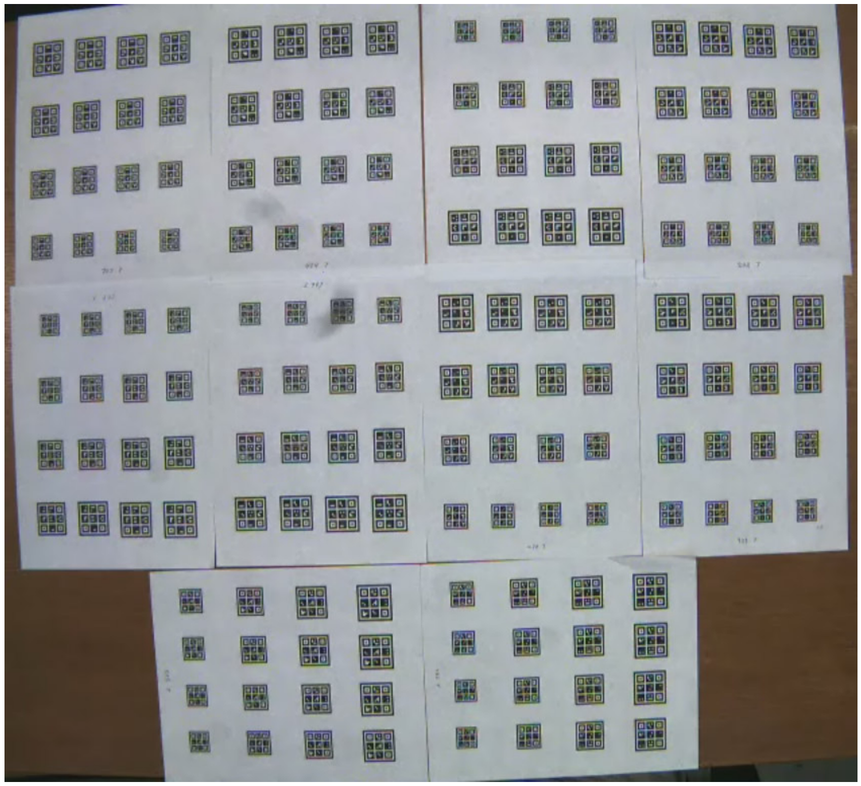


Figure 5.2 Papers used in Test-1. A total of 10 papers each having 16 tags of different sizes were used in the test.

On average we had 159.8 tags detected. Of these on average 105 were recognized, 6.6 rejected, 12 false tags and 36.4 tags could not be decoded. The minimum size of a recognized tag was 47x47 pixels on average. Each frame took an average of 8.6 seconds to process. We must note that this frame processing time is not realistic, in the real situation there can be a maximum of 4 tags per paper while in this test we had 16 tags per paper to push the system to the limit.

5.1.3 Test-2

We used the minimum size of tag found in test-1 to produce all tags for this test. The goal of this test is to find the true recognition rate of our recognition method. A total of 10 papers with 4 tags printed in the corners were placed on the desk similarly to test-1 (see Figure 5.3).

On average we had 49.8 tags detected. Of these on average 38.9 were recognized, 0.025 rejected, 10.85 were false tags and 0.0625 tags could not be decoded. Frame processing took 2.9 seconds on average. Finally, the successful recognition rate for individual tags was 97.2% (true positives) while there was no false positives in the data.

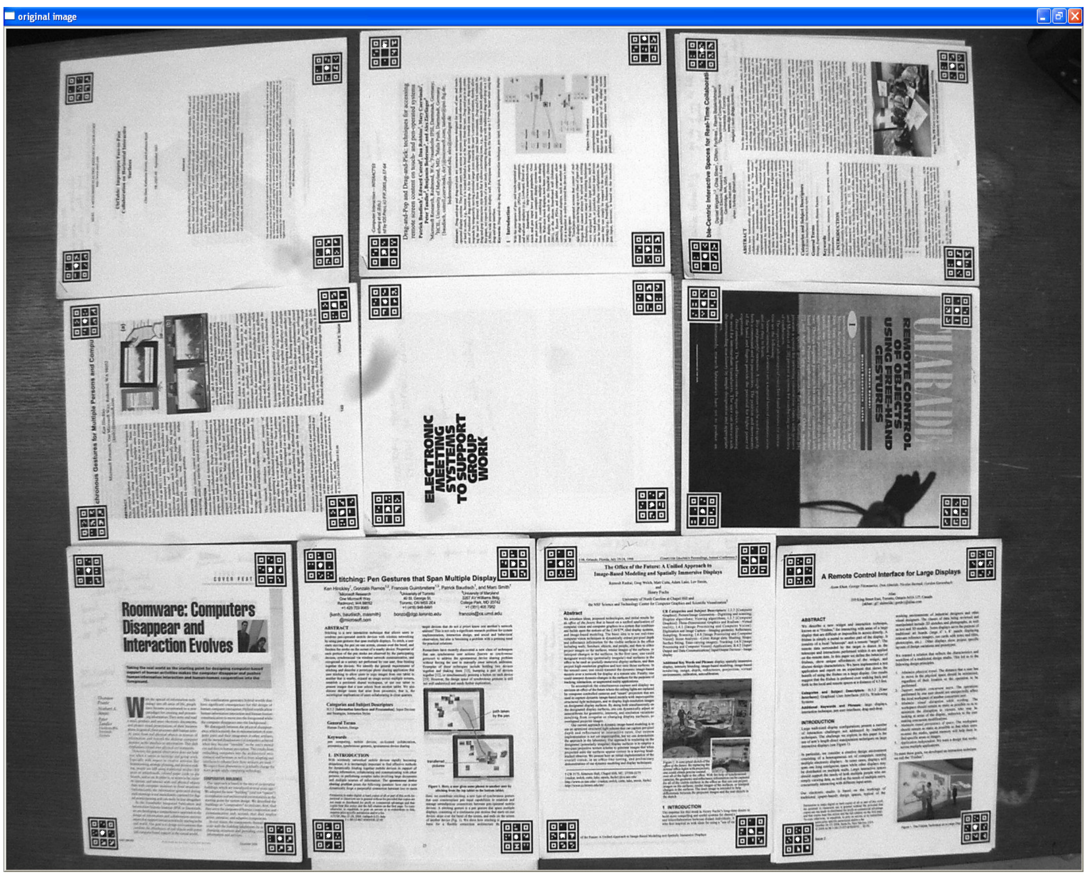


Figure 5.3 Papers used in Test-2. A total of 10 papers with 4 tags each.

5.1.4 Pile Tracking

This test was designed to estimate the accuracy of our pile tracking method described in Chapter 4. To achieve this, we prepared 50 papers with their ID tags printed in the corners of the paper. We used all of these papers to create and shuffle piles on the desk. After each paper added to the desk we checked the Desk View and the Pile List View to see the result. All misrecognitions, recognition failures, and consistencies were recorded immediately. Tasks were carried out in the following order (see Appendix C for sequence of pile manipulation):

- Populate Desk: create 5 piles by placing 5 papers on the desk consecutively in different spots. The desk space can fit more piles, but we needed empty space to split and move piles around the space.
    - Stack more papers on top of the 5 piles until all the 50 papers were in piles on the desk.
- Paper-moved: move piles to different spots on the desk. Piles can be moved simultaneously as long as they are not moved into the same spot.
- Merge piles: group all piles into a single pile. This was done by moving one pile at a time and placing it on top of the fixed pile.
- Split Piles: create new piles by removing sub-stacks of papers from the single pile and placing them in empty space on the desk, thus creating 5 piles again.
- Split-merge: remove sub-stack of papers from 3 piles and add them on top of other piles.

All papers were recognized correctly. Piles were also tracked accurately except for one paper that was placed in its own pile. This was caused by adding papers to the piles faster than the system can process. Improving the efficiency of the recognition method would solve this problem. We anticipate that a frame processing time of one second or less would be appropriate for this task.

Another problem detected is that when the user is looking through a pile of papers, papers inside the pile are revealed to the system; this messed up the order of the papers inside the pile. The problem is that the paper on top of the pile is covered with the user's hand, while the tag of an inside paper becomes visible to the camera and hence the system thinks the top paper was removed from the pile. A possible solution is to enable the system to detect such an activity. When the user is looking through a pile the system can ignore the changes to that pile until the top paper is back on top.

Accuracy of the pile tracking is critical for providing consistent information to the user. If the piles contents are not accurate, the user can not trust the system, and hence the system would be unusable.

5.2 User Evaluation

Because this system is interactive in nature, technical evaluation and performance measures are not sufficient for the evaluation process. Therefore, we conducted a user study to evaluate the system in action. The goal of this study is to confirm the feasibility of using the tool and to highlight points of failure that need to be addressed in future implementations.

5.2.1 Method

We have setup the prototype system on three different desks and computers in separate offices (see Figure 5.4). All three users were computer scientists, two professors and one graduate student. Each setup used a different camera (resolution, frame rate and image quality differed). As a result, the size of the tracked area of the desk was also different for each setup.

We collected qualitative data based on the experience that each user had while using the system. Each user recorded incidents, events and interactions on a special logging form that included events of failure, success, system flaws, and general behavior effects on the user's routine work practices. However, this data was not meant for a statistical analysis of using the prototype. Several interviews were held at the end of the users' working day to further obtain deeper understanding of their experiences and concerns about the system.

Figure 5.4 The three PaperSpace setups used in the user evaluation study.

5.2.2 Results

All users were excited about the idea of automatically keeping track of their desk papers. The system made it easier for them to find papers on the desk, simply by showing which pile the paper is in and its order within the pile. One user pointed out that they would spend few minutes trying to find a paper in their stacks. If he does not find the paper within a reasonable time, he just reprints it again.

We were surprised to see that simply informing the users whether a document has a printed version or not actually helps save them from searching through their piles of papers trying to find something that may or may not exist. In one example the user queried the system for a paper that he thought was among his papers on the desk. But he got negative result, the paper is not on the desk, then he looked at the document information and found that the document was never printed. After that, the user always looked at that information first.

The pile list view, although mainly created to avoid having to implement direct interaction with the virtual desk view, turned out to have interesting benefits as an alternative view of the paper space. For example, users found that the title of the document is easier to identify a paper when the virtual desk view was relatively small.

Users stressed the advantage of having the system embedded into a full document management tool -in this case for researchers- a reference manager system is a good example. We believe PaperSpace or a similar tool would be very useful for researchers if integrated into a commercial bibliography manager.

An interesting finding is that accurate pile tracking was particularly important. One user complained about the inaccurate order of a paper in the pile as the system showed in the pile list view. A media bridging tool must provide accurate and reliable information about the environment, otherwise, the user will loose trust in the system.

System delay, although small, was not easy to accommodate by the users behavior. Users usually work with documents at their own pace, they do not have to think about how fast/slow they move or place documents on the desk. Because of the delay in processing each frame, the user had to pause between adding papers to the desk (the delay averaged 2-3 seconds for each frame). This could undermine the lightweight characteristic of the system depending on how severe its effect on the user is.

Although users in this study found problems with some aspects of the system, these did not hinder their ability to use the system successfully to manage their research related document management. Most of the problems were implementation specific, which could be avoided in future versions. System crash occurred several times throughout the study period. However, the effects of this failure were avoided through a recovery mechanism. When the system is restarted it recovers the previous paper piles from a saved file.

Another problem that was evident is inserting new documents into the system. The system provided 2 methods for adding a document: manually typing the information into a data form or automatically extracting the data from an online digital library (the ACM library was used). Users stated that the manual option was not convenient. The automatic option was limited to a single source; it would be more useful to support most available libraries. For the purpose of this study, we added all of the users' documents into the system.

Overall, participants realized the benefits of this system in keeping track of their paper documents on the desk. Nevertheless, the problems identified above need to be resolved in a fully working system. In particular, the addition of documents and their information to the system must be as light as possible (drag/drop the document's icon for instance).

Chapter 6

Discussion

The user evaluation study showed that users appreciated the benefits of tracking papers. One of the users had a paper submission deadline, towards the end of the deadline he was working on the references list. He found the system very helpful in quickly finding a paper. Even for the few papers that were not in the system's database, it saved him from searching through the stacks of papers on the desk. This clearly shows that the system meets our first requirement. Technology should be used to meet users' needs and not merely introducing a fancy tool that serves no defined purpose.

Participants in the user evaluation worked on their documents naturally, while the system was running. They liked the fact that they can get this support for paper tracking without having to make extra effort to make it work. Simply printing the paper through the system makes it trackable. Seamless support for the users' work environment makes the bridging tool more convincing to use.

Users also stressed the need for visualization of piles history. Our system already supports the pile history by saving the history of paper and pile locations on the desk. Although the system does not have a visual representation of the piles history, it notifies the user of previous locations of papers when the user searches for a paper. Also, the system allows the user to assign attributes to groups of documents which can be later retrieved by these attributes. This can be extended by automatically creating context groups, in which all papers in a pile are assigned an attribute based on the title of the top document.

Although other researchers built paper tracking systems, these systems were not tested in a practical situation, and there has been no user study of using such systems in real work settings. Our method for tracking papers using visual tags outperforms previous attempts of tracking papers [34]. Our system provides more accurate recognition and faster processing of the desk scene. Test-2 showed a recognition rate of 97.2% for individual tags. For the recognition of papers with 4 tags each, recognition rate is close to 100%. None of the users in the user evaluation study complained of a paper not being recognized by the system.

There was a challenge in keeping the virtual view in sync with the real desk content. In the case of partially overlapping papers, recognizing the tags alone was not sufficient for finding which paper is on top. This can be solved by using the paper edges to sort out their order. Also, using information from the paper edges in the image could improve the overall processing of piles, in particular, it may ease some of the constraints on the user's interaction with the system.

Despite the small frame-processing delay, users felt that the system should have faster response time to allow natural tracking of piles. Our test showed an average response time of 2.9 seconds, the deployment study showed that this speed is still not enough for uninterrupted interaction. We expect a response time of one second or less to be acceptable. Our recognition method has room for more optimization and it is possible to achieve a one second response time. A simple solution was added to the system to overcome this delay by revealing to the user that the current frame has finished processing. This actually improved their interaction with the system. Previous research also confirms this improvement [24].

It might be interesting to combine the idea of link management with other augmented reality dreams. For example, the system can be extended to extract marks/highlights made on paper which can be merged or associated with the digital version. A projector can be used to show the location of a paper on the desk directly.

Computer vision seems to be the most viable solution to the problem of tracking piles of papers on the desk. It also fits well with other paper augmentation systems proposed by several researchers [6, 5, 53, 54]. Other technologies like RFID tags can only provide information about the presence of the object and not the exact location and orientation of the object. There are also other problems with using these tags; they cost more than printing, the user has to explicitly stick the tag on the paper and link that tag to a digital document in the system.

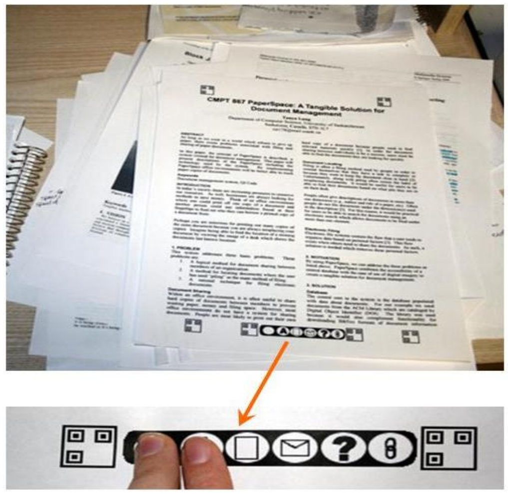


Figure 6.1 Top: command bar printed on bottom of paper. Bottom: the user issues a command to the system by making a simple gesture.

To allow rich and convenient user interaction, simple hand gestures can be recognized to issue system commands. A command bar can be printed on the paper to allow the user to issue system commands related to that paper (e.g. open, print, annotate or cite the digital version of the paper). Initiating a command can be as simple as covering two spots in the command bar (see Figure 6.1).

We intend to test the system exhaustively for users that belong to different categories of link management. This would allow us to confirm our definition of link management categories as well as finding the significance of the PaperSpace system for users in different categories.

Our user study focused exclusively on paper piles, it maybe of interest to conduct a similar study with more focus on the archived or filed papers in the office desk.

Chapter 7

Conclusion

The problem this thesis addressed was that people have difficulty in keeping the link between the paper and digital document spaces. Our main motivation is that current technology has the potential to support both document spaces, and can assist people in keeping track of their documents. This support will also reduce the user's cognitive load for keeping their document spaces organized.

Our solution was to build a system that tracks paper documents on the desk and link them to their digital versions on the computer. First, this thesis explored the problem by running a semi-structured user study of their paper and digital space organization of information. This study yielded a set of guidelines for building our prototype system.

User evaluation of the prototype showed that tracking of papers and their links to digital documents reduces the user's dependency on memory for finding and keeping track of their documents.

7.1 Summary

We conducted a user study to find the way people organize their paper and electronic document spaces and how they keep track of the link between the two. This study yielded a set of requirements for a document management tool that fills the gap between paper and digital documents.

Based on the requirements from the user study, we built and tested a prototype system that tracks paper documents and provides the user with an interactive interface to visualize and organize their documents. The system tracks papers and piles on the desk with acceptable accuracy and processing speed.

A two-week case study of three users trying the system showed our prototype to be useful not only for the single user document needs, but also for sharing documents and collaborating among peer workers. Results also showed the importance of the link between paper and digital documents to make it easier to juggle between the two media.

7.2 Contributions

The major contribution of this research is the expansion of the design space of personal document management systems by using the physical desk space in organizing and finding documents. The tracking of paper documents on the desktop opens a new paradigm upon which electronic documents can be viewed and organized. Minor contributions are:

- Identification of the requirements for supporting and integrating the paper and digital work spaces. There has been extensive work on supporting the digital work space, which is not the case for the paper work space and combining the two work spaces into an integrated system.
- Demonstrating through a working prototype the feasibility of building a system that manages the link between paper and digital document spaces.
- Identification of the benefits, tradeoffs and limitations of using such a system. A working paper and digital document bridging system has not been tested previously.
- A paper tracking method based on computer vision techniques. Previous computer vision methods for tracking papers on the desk top suffered from slow response time.

7.3 Future Work

This research introduced the first study of a paper-digital document bridging tool. The user study conducted was a preliminary one, a larger and more diverse sample of participants may yield a more thorough set of requirements for building a bridging tool. The prototype system that we built and evaluated was limited by time constraints and has room for improvements. Also this prototype can be extended with many interesting features that would make it more user-friendly.

The tracking method used in the current prototype can be optimized for even faster processing of the desktop's images. Improving the speed of recognizing and tracking the papers would make the system more robust and tolerant to user's manipulation of paper documents on the desk. Another issue related to the tracking method is the size of the visual tags. Current tags are considered too large to include on all pages of a document. One possible solution is to use two or more cameras to increase the total resolution of the image and hence the tag size can be reduced significantly.

The user interface can be extended in several ways. Papers' thumbnail representations in the virtual desk view can be given a physical feel in which the user can directly manipulate them with the mouse [2]. These thumbnails can then be shuffled and moved within the virtual desk space. Another aspect of the user interface that can be enhanced is the addition of documents to the system. Commercial document tools –bibliography tools for instance- already have integrated the importing of documents and their information from internet pages with the click of a button.

Our evaluation of the prototype was limited in the number of subjects and the period of evaluation. A future enhancement of the system should be evaluated with a larger number of subjects and for an extended period of time. Such an evaluation would allow us to better understand the users' needs for managing their document spaces.

This system also has a potential for supporting sharing of document spaces among peer researchers and knowledge workers. Especially when a group of people collaborate on some project, the system can reduce the effort spent on synchronizing their project related documents.

This thesis presented a promising tool for supporting and integrating the paper and digital documents simultaneously.

References

[1] AbuSafiya, M., and Mazumdar, S., (2004). Accommodating paper in document databases. In Proceedings of the 2004 ACM symposium on Document engineering, ACM Press, 155-162.

[2] Agarawala, A. and Balakrishnan, R., (2006). Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. Proc. CHI '06. 1283-1292.

[3] Aliakseyeu D., Martens, J.-B., Subramanian, S., Vroubel, M., Wesselink, W. (2001). Visual Interaction Platform. In Proceedings. of Interact 2001, July 2001, Tokyo, Japan (in press).

[4] Arai, T., Machii, K., Kuzunuki, S., and Shojima, H., (1995). InteractiveDESK: a computer-augmented desk which responds to operations on real objects. In Conference companion on Human factors in computing systems, ACM Press, 141-142.

[5] Arai, T., Machii, K., and Kuzunuki, S., (1995). Retrieving electronic documents with real-world objects on InteractiveDESK. In Proceedings of the 8th annual ACM symposium on User interface and software technology, ACM Press, 37-38.

[6] Arai, T., Aust, D., and E. Hudson, S., (1997). PaperLink: a technique for hyperlinking from real paper to electronic content. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 327-334.

[7] Barreau, D. and Nardi, B. A. 1995. Finding and reminding: file organization from the desktop. SIGCHI Bull. 27, 3 (Jul. 1995), 39-43.

[8] Bellotti, V., Dalal, B., Good, N., Flynn, P., G. Bobrow, D., and Ducheneaut, N., (2004). What a to-do: studies of task management towards the design of a personal task list manager. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 735-742.

[9] Bergman, O., Boardman, R., Gwizdka, J., and Jones, W. (2004). Personal information management. In CHI '04 Extended Abstracts on Human Factors in Computing Systems. CHI '04. ACM Press, 1598-1599.

[10] Boardman, R., Spence, R., and Sasse, M. A. (2003). Too many hierarchies? the daily struggle for control of the workspace. In Proceedings of Human-Computer Interaction International 2003.

[11] Boardman, R., and Angela Sasse, M., (2004). "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 583-590.

[12] Bondarenko, O., and Janssen, R., (2005). Documents at Hand: Learning from Paper to Improve Digital Technologies. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 121-130.

[13] Conroy, K., Levin, D., Guimbretiere, F. (2005). ProofRite: A Paper-Augmented Word Processor. ACM Transactions on Computer-Human Interaction. Submitted Nov.

[14] Curiel, J.M. & Radvansky, G.A. (1998). Mental organization of maps. In Journal of Experimental Psychology: Learning, Memory and Cognition, 24, 202-214.

[15] Darken, R. & Sibert, J. L., (1996). Navigating Large Virtual Spaces. International Journal of HCI, 8, 49-72.

[16] Dourish, P., Keith Edwards, W., LaMarca, A., and Salisbury, M., (1999). Presto: an experimental architecture for fluid interactive document spaces. In ACM Trans. Comput.-Hum. Interact., ACM Press, 133-161.

[17] Dourish, P., Keith Edwards, W., LaMarca, A., and Salisbury, M., (1999). Using properties for uniform interaction in the Presto document system. In Proceedings of the 12th annual ACM symposium on User interface software and technology, ACM Press, 55-64.

[18] Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., and C. Robbins, D., (2003). Stuff I've seen: a system for personal information retrieval and re-use. In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, ACM Press, 72-79.

[19] Dymetman, M., and Copperman, M., (1998). Intelligent Paper. In Proceedings of the 7th International Conference on Electronic Publishing, Held Jointly with the 4th International Conference on Raster Imaging and Digital Typography: Electronic Publishing, Artistic Imaging, and Digital Typography, Springer-Verlag, 392-406.

[20] Fertig, S., Freeman, E., and Gelernter, D. (1996). Lifestreams: an alternative to the desktop metaphor. In Conference Companion on Human Factors in Computing Systems: Common Ground, M. J. Tauber, Ed. CHI '96. ACM Press, 410-411.

[21] Gemmell, J., Bell, G., and Lueder,n R., (2006). MyLifeBits: a personal database for everything. In , , 88-95.

[22] Gemmell, J., Lueder, R., and Bell, G., (2003). The MyLifeBits lifetime store. In Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence, ACM Press, 80-83.

[23] Guimbretiere, F., (2003). Paper augmented digital documents. In Proceedings of the 16th annual ACM symposium on User interface software and technology, ACM Press, 51-60.

[24] Gutwin, C., Benford, S., Dyck, J., Fraser, M., Vaghi, I., and Greenhalgh, C. (2004) Revealing Delay in Collaborative Environments, Proc. CHI '04, 503-510.

[25] Harper, R. and Sellen, A., (1995). Paper supported collaborative work. Technical report, Xerox Research Centre Europe.

[26] Heiner, J., E. Hudson, S., and Tanaka, K., (1999). Linking and messaging from real paper in the Paper PDA. In Proceedings of the 12th annual ACM symposium on User interface software and technology, ACM Press, 179-186.

[27] Henderson, S., (2005). Genre, task, topic and time: facets of personal digital document management. Proc. CHINZ '05, 75-82.

[28] Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns, D., (2005). Paper windows: interaction techniques for digital paper. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 591-599.

[29] Janssen, W. C. and Popat, K., (2003). UpLib: a universal personal digital library system. Proc. ACM Symp. on DocEng '03, 234-242.

[30] Johnson, W., Jellinek, H., Klotz, Jr., L., Rao, R., and K. Card, S., (1993). Bridging the paper and electronic worlds: the paper user interface. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 507-512.

[31] Kaptelinin, V., (2003). UMEA: translating interaction histories into project contexts. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 353-360.

[32] Kidd, A., (1994). The marks are on the knowledge worker. In Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, ACM Press, 186-191.

[33] Kim, J., M. Seitz, S., Agrawala, M., (2004). The Office of the Past: Document Discovery and Tracking from Video. IEEE Workshop on Real-Time Vision for HCI, 2004.

[34] Kim, J., M. Seitz, S., and Agrawala, M., (2004). Video-based document tracking: unifying your physical and electronic desktops. In Proceedings of the 17th annual ACM symposium on User interface software and technology, ACM Press, 99-107.

[35] Kirsh, D. (2000). A Few Thoughts on Cognitive Overload. Intellectica, 19-51.

[36] Krumm J. (1997). Object detection with vector quantized binary features. cvpr, IEEE 179, 1997.

[37] Krumm, J., Harris, S., Meyers, B., Brumitt, B., Hale, M., Shafer, S. (2000). Multi-Camera Multi-Person Tracking for EasyLiving, vs , IEEE, 3.

[38] Koike, H., Sato, Y., Kobayashi, Y., Tobita, H., and Kobayashi, M., (2000). Interactive textbook and interactive Venn diagram: natural and intuitive interfaces on augmented desk system. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 121-128.

[39] Koike, H., Sato, Y., and Kobayashi, Y., (2001). Integrating paper and digital information on EnhancedDesk: a method for real-time finger tracking on an augmented desk system. In ACM Trans. Comput.-Hum. Interact., ACM Press, 307-322.

[40] Lawrie, D. and Rus, D. (1999). A self-organized file cabinet. In Proceedings of the Eighth international Conference on information and Knowledge Management. S. Gauch, Ed. CIKM '99. ACM Press, 499-506.

[41] Liao, C., Guimbretiere, F., and Hinckley, K., (2005). PapierCraft: a command system for interactive paper. In Proceedings of the 18th annual ACM symposium on User interface software and technology, ACM Press, 241-244.

[42] Mackay, W.E., (1988). Diversity in the Use of Electronic Mail: A Preliminary Inquiry. TOIS: ACM Trans. on Office Information Systems, Vol. 6(4).

[43] Malone, T., (1983). How do people organize their desks? Implications for the design of office information systems. In ACM Trans. Inf. Syst., ACM Press, 99-112.

[44] Mander, R., Salomon, G., and Yin Wong, Y., (1992). A "pile" metaphor for supporting casual organization of information. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 627-634.

[45] Murase, H. and Nayar, S. k. (1995). Visual Learning and Recognition of 3D Objects from Appearance. International Journal of Computer Vision, 14(1):5-24.

[46] Newman, W., and Wellner, P. (1992). A desk supporting computer-based interaction with paper documents. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, 587-592.

[47] Nishi, T., Sato, Y., and Koike, H. (2001). Interactive object registration and recognition for augmented desk interface. In CHI '01 extended abstracts on Human factors in computing systems, ACM Press, 371-372.

[48] Norrie, M.C., Signer, B. (2003). Switching over to paper: a new Web channel. In Proceedings of Web Information Systems Engineering '03, IEEE, 209- 218.

[49] O'Hara, K. and Sellen, A., (1997). A comparison of reading paper and on-line documents. Proc. CHI '97, 335-342.

[50] Rauterberg, M., Fjeld, M., Lauche, K., Dierssen, S., & Bichsel, M., (1998): BUILD-IT: A Brick-based integral Solution Supporting Multidisciplinary Design Tasks. In A. Sutcliffe, J. Ziegler & P. Johnson (eds.) Designing Effective and Usable Multimedia Systems, pp. 122-133.

[51] Ravasio, P., Schar, S., and Krueger, H. (2004). In pursuit of desktop evolution: User problems and practices with modern desktop systems. 156-180.

[52] Rekimoto, J. (1998). Matrix: A Realtime Object Identification and Registration Method for Augmented Reality. In Proceedings of Asia Pacific Computer Human Interaction (APCHI '98).

[53] Rekimoto, J., and Saitoh, M. (1999). Augmented surfaces: a spatially continuous work space for hybrid computing environments. In Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, ACM Press, 378-385.

[54] Robertson, C., and Robinson, J., (1999). Live paper: video augmentation to simulate interactive paper. In Proceedings of the seventh ACM international conference on Multimedia (Part 2), ACM Press, 167-170.

[55] Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D., and van Dantzich, M., (1998). Data mountain: using spatial memory for document management. Proc. UIST '98, 153-162.

[56] Rus, D. and de Santis, P., (1997). *The Self-Organized Desk*. In Proceedings of the 1997 International Joint Conference on Artificial Intelligence, Nagoyo, Japan.

[57] Sellen, A. and Harper, R., (2001). The Myth of the Paperless Office. 1st ed. 2001: MIT press.

[58] Smith, J., Long, J., Lung, T., Anwar, M. M., and Subramanian, S., (2006). PaperSpace: a system for managing digital and paper documents. CHI '06 Ext. Abstracts, 1343-1348.

[59] Swain M., Ballard D., (1991) *Color Indexing*. Intern. J. of Computer Vision, 7, 1, pp. 11-32.

[60] Want, R., P. Fishkin, K., Gujar, A., and L. Harrison, B., (1999). Bridging physical and virtual worlds with electronic tags. In Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, ACM Press, 370-377.

[61] Wellner, P. (1993). Interacting with paper on the DigitalDesk. Commun. ACM 36, 7 (Jul. 1993), 87-96.

[62] Whittaker, S., and Hirschberg, J., (2001). The character, value, and management of personal paper archives. In ACM Trans. Comput.-Hum. Interact., ACM Press, 150-170.

Appendix A - Interview Structure

The following list of questions guided the flow of the interviews.

1. How many documents that you have authored do you have in your file system (10, 100, 1000 …)?

2. How many documents that others have authored do you have in your file system (10, 100, 1000 …)?

3. Please describe the way you organize documents on the computer (is there a system).
   Any tools used?
   Is there a difference between own documents and other's documents?
   Are there any criteria for **grouping** documents (type, author, project, length, properties …)?
   Any difficulties/problems?

4. Do you ever find it difficult to find a document on the desk? How often?
   What made it difficult?

5. How many documents do you have on your desk?
   What are the sources of these (ref. papers, forms, letters, emails, notes …)?

6. Please describe the way you organize documents on the desk (any mechanism).
   Do you use **piles**, binders …?
   Are there any criteria for **grouping** documents?
   Are there any difficulties/problems?

7. What is the source of your documents on the desk?
   Do they have a digital version in the computer?
   Do you prefer working with physical or virtual papers?

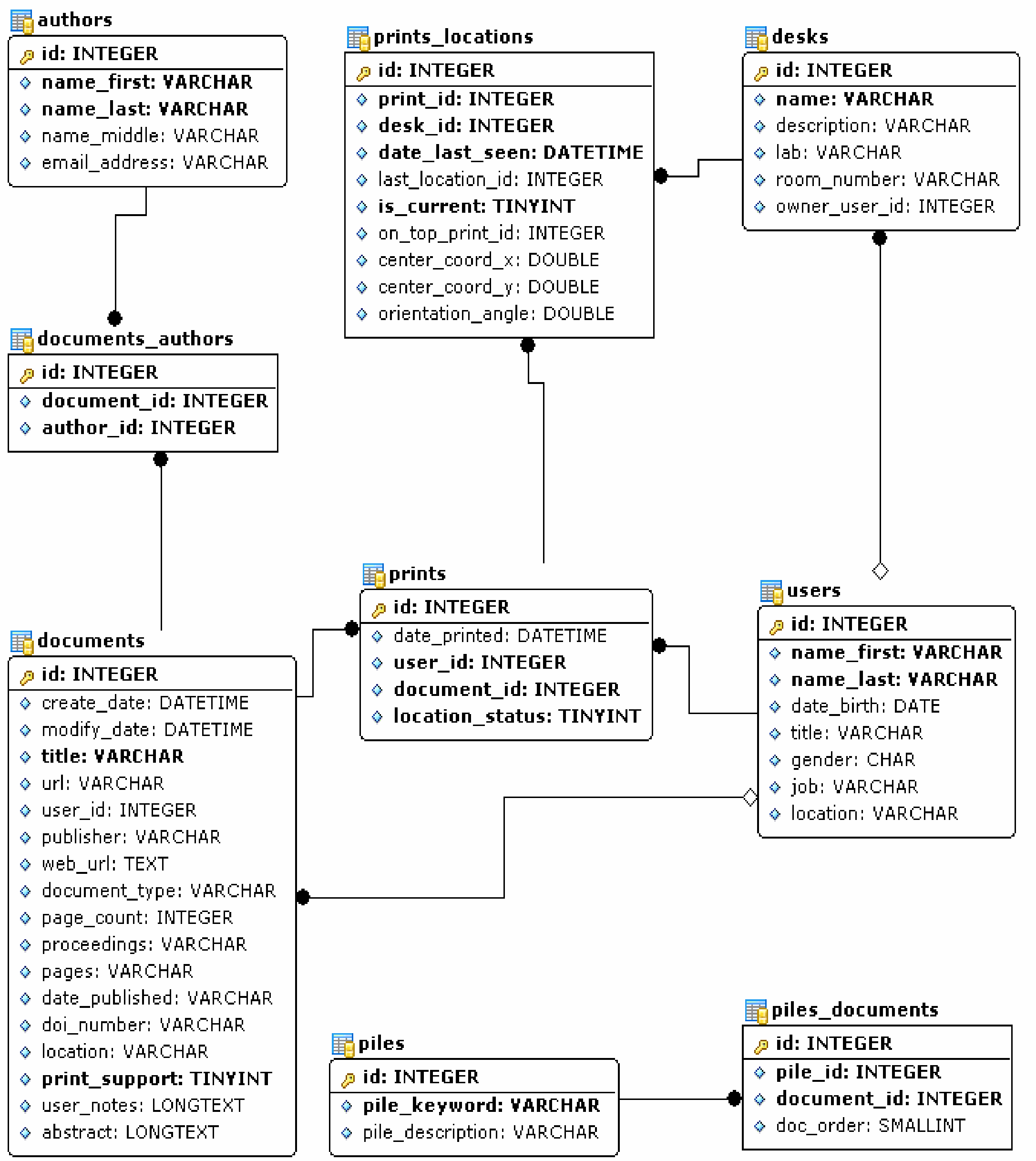
8. When you print documents, do you keep a link to the digital source?
   Why? How? Is it important? Is the link mostly mental?

9. Do you ever find it difficult to find a document on the desk? How often?
   What made it difficult?

10. Are there any tools or techniques (bibliography tools for example) that you use to maintain (establish/manage) connections between physical and digital documents? Please describe it if any. What difficulties/problems are associated with such tools?

11. Is there a case where you are reading some paper, and at the same time need to open or get its digital version? Or vise versa? Please provide description of such cases.

Appendix B - PaperSpace Database Schema



Appendix C - Pile Tracking Test

The following sequence of images demonstrates the creation and manipulation of paper piles during the pile tracking test described in Chapter 5. The sequence progresses from left to right and top to bottom.