Discussion and Conclusions

Our work to date has exposed some of the research issues posed by the concept of interacting with paper documents and other passive artifacts. One set of issues concerns the established roles of these artifacts in the work environment, and how these roles are affected by making the artifacts active. For example, usage of paper documents has always tended to rely on the absence of hidden information whereas in an interactive context, hidden information such as translations can be revealed. Another set of issues relates to the problems in making passive objects active: problems such as how to access the information content of the object, and how to provide the necessary feedback to support the user's interaction with the object. A printed document is a relatively easy target compared with, say, a filing cabinet or a bookshelf.

Two central issues in supporting interaction with paper documents are the development of effective styles of interaction, and the provision of tools for the design and implementation of paper-based user interfaces. So far, we have made only a little progress in these two areas. In the applications we have built we have tried to adhere to existing workstation interaction styles so as to avoid surprises for our users. We have experimented with paper "command sheets" as a means of offering the user a range of command modes, but recognize that this puts limits on the interactivity of the application.

Our interest in the digital desk stems in part from its ability to support "tactile manipulation" interfaces. These attempt to let the user interact with physical artifacts and electronic objects in much the same way: by touching them. With this style of interaction, manipulation of electronic documents is similar to manipulation of paper documents, and paper can have computer-based properties that make it possible to interact with it like an electronic object.

Ultimately, we hope to integrate paper documents better into the electronic world of personal workstations, making paper-based information usable in the same ways as electronic information, without requiring the user to abandon paper and use an electronic workstation instead. The ideal system that we envision is one in which users are free to choose either medium as the task requires without being constrained to the limitations of either.

Acknowledgments

We wish to thank the many people who have offered advice and help, including Dan Bloomberg, Ian Daniel, Mike Flynn, Ron Kaplan, Martin Kay, Mik Lamming, Linda Malgeri, Ron Mann, Mike Molloy, Atty Mullins, Peter Robinson, Z Smith, Larry Spitz, and our subjects.

References

[Bloo90] Bloomberg, D., and Maragos, P. "Image Algebra and Morphological Image Processing." *SPIE Conference Proc.*, San Diego, CA, July 1990.

[Gold91] Goldberg, D. and Goodisman, A."Stylus User Interfaces for Manipulating Text." in *Proceedings of the ACM Symposium on User Interface Software and Technology* (UIST '91), November 11-13, Hilton Head 1991.

[Grea92] Greatbatch, D., Heath, C., Luff, P., and Campion, D. "On the use of Paper and Screen-based Documentation in the General Practice Consultation." Rank Xerox EuroPARC Technical Report, Cambridge 1992.

[Hans88] Hansen, W. J., and Haas, C. "Reading and Writing with Computers: A Framework for Explaining Differences in Performance." in *Communications of the ACM*, vol. 31, no. 9, September 1988.

[Ishi91] Ishii, Hiroshi and Miyake, Naomi. "Toward an Open Shared Workspace: Computer and Video Fusion approach of TeamWorkStation." In *Communications of the ACM*, vol. 34, no. 12 December 1991.

[Know77] Knowlton, K. "Computer Displays Optically Superimposed on Input Devices." *Bell System Technical Journal*, vol. 56, no. 3, March 1977.

[Krue91] Krueger, M. *Artificial Reality II*. Addison-Wesley, 1991.

[Lamm92] Lamming, M. and Newman, W. "Activity-based Information Retrieval: Technology in support of Personal Memory." Rank Xerox EuroPARC Technical Report 91-03. February 1992.

[Mill91] Miller, R. "Behold the Humble Barcode." Multimedia and Videodisk Monitor. Arlington VA, Sept. 1991.

[Newm91] Newman, W. "Interacting with Paper Documents" in *Proceedings of NordDATA '91*, June 16-19; Oslo 1991.

[Well91] Wellner, P. "The DigitalDesk Calculator: Tangible Manipulation on a Desk Top Display" in *Proceedings of the ACM Symposium on User Interface Software and Technology* (UIST '91), November 11-13, Hilton Head 1991.

[Weis91] Weiser, M. "The Computer for the 21st Century" *Scientific American*, September 1991.

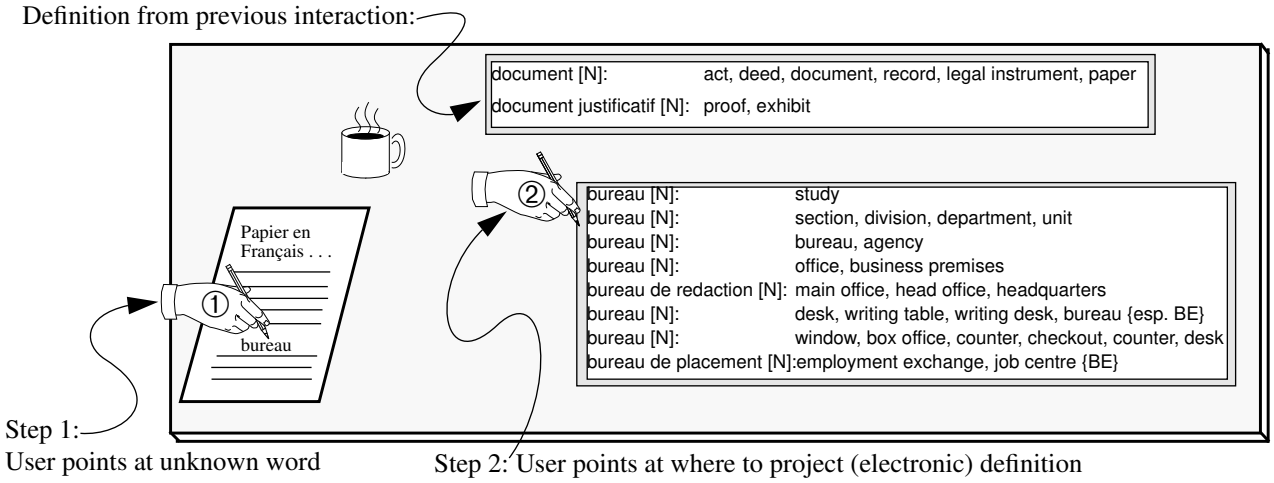

Figure 4. Using translator on surface of desk

Desk top translation

Another example of an application that can benefit from this desk is foreign language translation. Looking up words in a dictionary can take up a substantial amount of a reader's time when reading documents in a foreign language. The time spent looking up words also makes it more difficult to remember the context of the word or the passage. Some readers find this delay so disruptive to their reading that they prefer to read on despite the presence of many unknown words.

We have implemented a system in which French documents can be read at a desk in their paper form. When the user needs to look up a word, he or she can simply point at the word. The system extracts the root of the word, looks it up in a French-to-English dictionary and displays the definitions in an electronic window projected onto the desk, allowing the user to point to the location where the translation should be placed on the desk (see Figure 4). Any number of words can be looked up, and the user can slide the windows around on the desk or remove them when they are no longer needed.

User Reactions

Although we did not conduct any formal experiments with our system, a small number of people have tested the translation application and we recorded their reactions. These related both to the application and to the digital desk technology, both of which were new to the users.

The on-line dictionary provided much faster lookup of unfamiliar words, partly through automatic determination of the root word (e.g., "aller" from "vais") and partly by avoiding the manual lookup process itself. Although on-line lookup look several seconds, users were often able to select a word and continue reading uninterrupted, knowing that the translation window would appear. They liked the fact that translations remained visible, sometimes avoiding the need to look up a word twice. We expect these benefits would also be provided by a screen-based system, but only if the document were available on-line.

For the most part, the digital desk was found comfortable and natural by subjects, all of whom had used traditional workstations or PCs before. None was bothered by the projection system or by shadows cast by hands and other objects; indeed after using the system for about fifteen minutes, one user asked if the desk was made of glass and looked under it to see where the projected display was coming from! They were asked at the end of the session how they felt about using the desk as compared to a traditional workstation. Their overall reaction was that it was quite similar, and they commented specifically that they had "more space," that using it was "more manual," "easier on the eyes," and "much more healthy than a screen." Subjects were unsure whether the paper had to stay in a precise location, and one said that it was "irritating that the paper moved around" relative to the projected images. This could have been avoided by providing different feedback and is being investigated. One of our subjects was left-handed and was inconvenienced by the experimental setup, with paper on the left and space for definitions on the right—her arm obscured the paper while she waited to position the translation window. This observation suggests that there may be interesting differences in how handedness affects the layout of applications on a digital desk as opposed to a workstation. More comprehensive tests are still required, but these initial trials have not uncovered any fundamental problems with our approach, and have encouraged us to pursue the research further.

Published in Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '92) May 3 - 7, 1992. pp 587-592.

Skew detection and correction: Skew correction is essential to compensate for the varying angle at which documents are held while reading them or interacting with them. The low-resolution bitmaps lend themselves to rapid skew detection using morphological filtering techniques that search for parallel picture elements [Bloo90].

Feature extraction and encoding: The purpose of the feature extraction is to exploit typographic layout conventions to match the page image with a description of a high resolution image in the database. The features extracted are those that might be said to define the "shape" of the page as perceived by a reader: number of paragraphs, height of individual paragraphs, justification of text, and indentation. These are encoded as a descriptor containing three components: line lengths, paragraph heights and word breaks.

Database lookup: Pages are looked up in the page database by attempting to match the line-length component of the descriptor. If multiple entries are found with same line-length component, the paragraph-height and word-break components are used to discriminate between them. The word-break component encodes the position of words on the first two full lines of text, and is needed in cases where page shapes are otherwise indistinguishable.

Calibration: The system must be able to map each point of the camera image to a point of the projected display. If the digitizing tablet is used, then each point of the digitizing tablet must be mapped to a point in the camera image and to the display. Vibrations and optical distortions make it difficult to maintain accurate calibration over the entire area of interaction, and errors here make the system inaccurate and difficult to use. We have been experimenting with ways to perform calibration both manually and, with the aid of projected test patterns, automatically.

Display: A projected display provides similar capabilities to using a large flat display screen or a rear-projected display. It has the important advantage, however, that computer-generated images can be superimposed onto paper documents. A problem with overhead projection is shadows; for example, the user cannot lean over to look at a projected image too closely. In practice, however, this has not yet proved to be a problem. Another issue with projection is the brightness of the room. The projector used in these experiments works quite well with the room's normal fluorescent lights, but a bright desk lamp makes the display unreadable. The same would be true of direct sunlight, so this limits the desk's usability in some settings. One last problem is that not all materials make good projection surfaces. The projection area should be white and untextured in order to show the projected images most clearly. An ideal system would have projection both from above and from below.

Performance: A considerable amount of work has gone into ensuring adequate performance to support fast, reliable interaction. In its current implementation the system takes about 8 seconds to capture, threshold and skew-correct a full image, and then extract features and search the database. Since this is too slow for normal interaction, partial images are used whenever possible, and with this approach selection can be performed in less than a second. Error rates vary depending on the layout and quality of the printed text, and in order to support the applications described below we have used relatively good-quality printed material. A major focus of the research is to achieve lower error rates in normal use.

Example Applications

We have chosen two simple applications to illustrate how this system might be used: a calculator and a translation application.

Desk top calculator

People using calculators often enter numbers that are already printed on a piece of paper lying on the desk. They must copy the numbers manually into the calculator in order to perform arithmetic on them. Transcribing these numbers slows down the user of a calculator and contributes to errors.

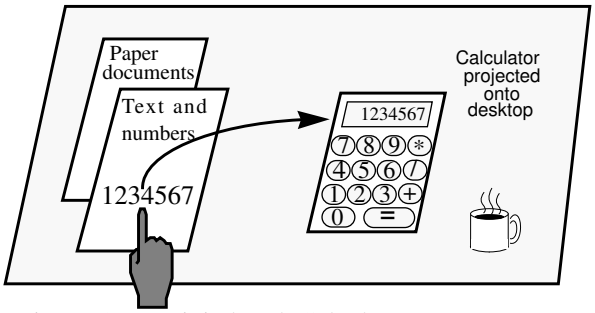

Figure 3. DigitalDesk Calculator

To address these issues, we project a calculator onto the desktop (see Figure 3), in this instance using DigitalDesk rather than Marcel. This enables finger-tip operation much like a regular electronic calculator. The projected cursor follows the user's finger (or pen) as it moves around on the desktop and the user can tap on the projected buttons to enter numbers. The advantage of this calculator over an ordinary calculator, however, is that it has an additional way to enter numbers. If the number to be entered is already printed on a piece of paper lying on the desk, the user can simply point at it and it appears in the calculator's display as though it had been typed in directly. This calculator can save many keystrokes and errors when making calculations with printed data.

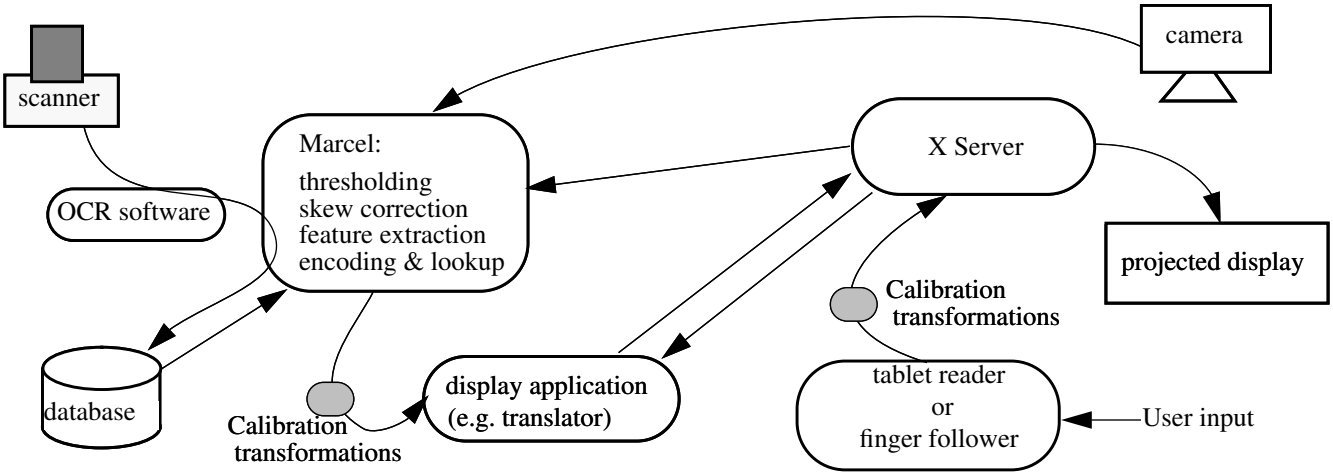

Figure 2. System Architecture

A currently active area of research and development is pen-based computer interaction (for examples see [Carr91] and [Gold91]). Much of this work is focused on note-pad sized computers instead of desk-sized displays, but the techniques developed for pen-based interfaces, such as handwriting and gesture recognition, will greatly enhance the style of interaction we describe in this paper. Our focus is on the unique aspects of this type of desk system, which allow paper documents to be used on the desk in similar ways as the electronic objects.

The Design of Marcel

The main difficulty in implementing a system to support computer-based interaction with paper documents is to relate a selected location on the paper to the text at that position. The ideal solution would be a high-resolution camera, capable of capturing a selected region of the desk surface and applying optical character recognition (OCR) to it. While there are hand-held scanning devices capable of capturing high-resolution images, they are inconvenient; OCR also poses problems of slow response and sensitivity to skew angle. We are interested in solutions that would allow the user to point with a stylus or even with a finger, at documents lying at various orientations. Devising a means for providing fast response to this form of text input is a major component of the system design. We have tested two solution strategies, both using a video camera mounted over the desk: The DigitalDesk and Marcel.

The DigitalDesk [Well91] uses supplementary cameras zoomed into narrow fields of view to obtain images of paper documents that are suitable for limited OCR, and it uses a finger-following camera for pointing. Although a desk calculator was implemented as an example to illustrate the concept, this means of pointing and OCR is not yet precise enough for real user testing.

Marcel, on the other hand, uses low-resolution video images, corrects them for skew, and correlates them with document images obtained from a printer or a high-resolution scanner. Commercial OCR software is applied to the scanned images ahead of time, and the digitizing tablet enables very precise pointing. The result is a prototype that is reliable enough for user testing.

The system architecture of Marcel is shown in Figure 2. Images are captured from the over-the-desk camera, and are fed through a number of stages: thresholding, skew-detection and correction, feature extraction, encoding and database lookup. Meanwhile, coordinates are read through X Windows from a hand-held stylus or other pointing device. These are used to select contents from the text found in a database, and the contents are passed as arguments to the appropriate computer function. Results are displayed on the desk surface using a projected computer display. The following sections discuss some of the issues this system needed to address; more details can be found in [Newm91].

Image capture and thresholding: Each gray-scale image is converted to binary by means of an adaptive thresholding routine that compensates for variations in lighting across the desk surface. With the camera adjusted to cover a 25 x 18 inch area, the effective scanning resolution is about 30 dots per inch. High Definition TV would of course improve image resolution, and was considered as an alternative but ruled out because it would also increase processing time without producing the quality needed for direct OCR.

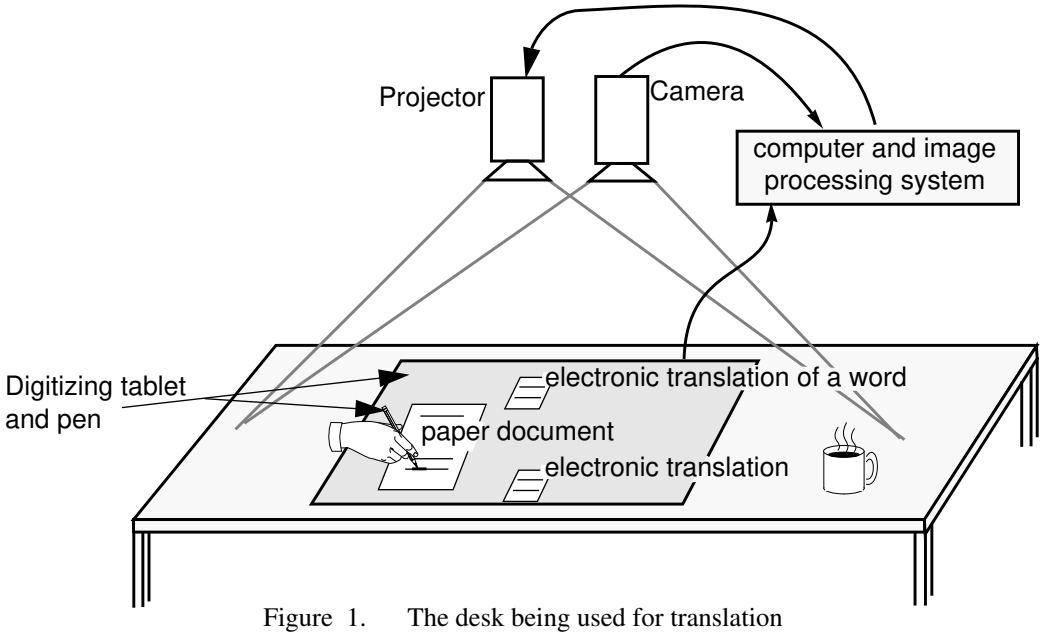

Figure 1. The desk being used for translation

only with electronic documents. We then explain the way we have gone about designing and implementing Marcel. To explore the characteristics of this technique, we have implemented two simple applications. The first is a desk calculator that allows the user to select numbers printed on paper for input into the calculator. The second is a translation program that allows the user to select French words and display the English translations. We comment on the effectiveness of the approach, which has undergone some user testing, and on the general implications of interacting with computers in this way.

System Description

The system is built around an ordinary physical desk, and can be used as such. It has additional capabilities, however. A video camera is mounted above the desk pointing down at the work surface. Also, a computer-driven projector is mounted above the desk, allowing the system to superimpose electronic objects onto paper documents and the user's work surface (see Figure 1). The camera's output is fed into an image processing system that can see what the user is doing on the surface of the desk. The image processing system is in fact capable of detecting where the user's finger is pointing, but in Marcel a digitizing tablet and cordless pen are used to allow more precise pointing.

Although the system uses unconventional means of input and output, users can still interact with documents in much the same way as on a workstation. Ordinary X Window applications run exactly the same way as on a standard workstation without distinguishing pen or finger-based input from mouse-based input. The system enables new types of interaction techniques, however, and new applications that exploit the user's direct access to the printed word.

Related work

Electronic and physical objects have been merged together using a half-silvered mirror in an early system for placing virtual labels on physical buttons [Know77]. Like our desk, VIDEODESK [Krue91] uses an over-the-desk camera, but tracks only silhouettes against a light table. The TeamWorkStation [Ishi90] also shares many of our goals in that it attempts to integrate traditional paper media with electronic media. It uses video mixing techniques to merge the images of paper documents on a desk with computer-generated images and the result is viewed on a monitor. In our system the result is viewed on the desk itself instead of a monitor. The main difference, however, is that our system not only provides a merged view of the two media, but also allows the user to perform computer-based interaction on the text of paper documents.

Another example of enhancing paper documents with computer functionality is the technique of barcoding text books [Mill91]. This allows a reader to scan a barcode with a laser wand and automatically play a particular track in a videodisk recording. Our approach presents the computer's response differently, but more importantly, it is much more general in that it does not require specially prepared documents. It can work with any paper document and the interaction is designed to be compatible with users' current work practices: with both electronic and paper media.

Published in Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '92) May 3 - 7, 1992. pp 587-592.

A Desk Supporting Computer-based Interaction with Paper Documents

William Newman and **Pierre Wellner***

Rank Xerox EuroPARC
61 Regent Street
Cambridge CB2 1AB (United Kingdom)
Newman@EuroPARC.Xerox.Com, Wellner@EuroPARC.Xerox.Com
*also University of Cambridge Computer Laboratory

Abstract

Before the advent of the personal workstation, office work practice revolved around the paper document. Today the electronic medium offers a number of advantages over paper, but it has not eradicated paper from the office. A growing problem for those who work primarily with paper is lack of direct access to the wide variety of interactive functions available on personal workstations. This paper describes a desk with a computer-controlled projector and camera above it. The result is a system that enables people to interact with ordinary paper documents in ways normally possible only with electronic documents on workstation screens. After discussing the motivation for this work, this paper describes the system and two sample applications that can benefit from this style of interaction: a desk calculator and a French to English translation system. We describe the design and implementation of the system, report on some user tests, and conclude with some general reflections on interacting with computers in this way.

Keywords: user interface, interaction technique, display, input device, workstation, desk, desktop, document recognition.

Introduction and Motivation

Today's office, with its dual system of paper and electronic documents, presents users with a range of problems. Electronic and paper documents have very different properties, are not well integrated, and working with one often requires giving up the advantages of the other. Electronic documents lack the portability, tangibility and universal acceptance of paper. Working with paper, on the other hand, denies direct access to the wide variety of interactive functions available on personal workstations—functions that include spelling-correction, electronic mail, keyword searching, numerical calculation and foreign-language translation.

The standard solution to problems of dealing with information on paper is to devise a screen-based alternative. These solutions are often based on a real-world metaphor. Users interact with computer-generated synthetic objects that simulate objects such as paper documents and a desktop. However, the replacement of paper by computer is not always appropriate. For example, reading from a screen is generally slower than from paper [Hans88], and this poses a problem with longer documents. Keeping medical records on-line means that doctors must consult a desktop computer rather than glance through a written record, and this changes the interaction between doctor and patient, not always for the better [Grea92].

An alternative to screen-based designs is offered by the opportunity to *augment* real-world objects with computers instead of simulating them. This is the concept of "embodied virtuality" or "computerized reality" [Weis91, Well91]. We apply this idea to the work environment and specifically to the desk. Instead of using the "desktop metaphor" and electronic documents, we use the real desk top and real paper documents. Instead of "direct manipulation" with a mouse, one of our goals is to explore "tactile manipulation" of real artifacts that we augment to have electronic properties. Our ultimate goal is to reduce the incompatibilities between the paper and electronic domains so that interchange is cost-free and users no longer pay a penalty for choosing the "wrong" medium at the outset. The research reported here is a first step towards attaining this goal.

In this paper, we describe a system, Marcel, developed originally for experiments on tracking paperwork (see [Lamm92]), and since extended to allow direct interaction with paper documents in ways that are normally possible

Published in Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '92) May 3 - 7, 1992. pp 587-592.