

10:58:14

## OCA PAD INITIATION - PROJECT HEADER INFORMATION

12/13/95

Active

Project #: C-36-X74                      Cost share #:                      Rev #: 0  
Center # : 10/24-6-R8763-0A0      Center shr #:                      OCA file #:  
Contract#: FOREIGN AGREEMENT                      Mod #:                      Work type : RES  
Prime # :                      Document : AGR  
Contract entity: GTRC  
  
Subprojects ? : N                      CFDA: NA  
Main project #:                      PE #: NA

Project unit:                      COMPUTING                      Unit code: 02.010.300  
Project director(s):  
HUDSON S E                      COMPUTING                      (404)894-9222

Sponsor/division names: HITACHI LIMITED                      / JAPAN  
Sponsor/division codes: 708                      / 017

Award period:      951120      to      961120      (performance)      961120      (reports)

Sponsor amount	New this change	Total to date
Contract value	8,100.00	8,100.00
Funded	8,100.00	8,100.00
Cost sharing amount		0.00

Does subcontracting plan apply ? : N

Title: COLLABORATIVE RESEARCH WITH TOSHIFUMI ARAI AT HITACHI

## PROJECT ADMINISTRATION DATA

OCA contact: Robert D. Simpkins                      894-4820

Sponsor technical contact                      Sponsor issuing office

MANAGER YOSHIAKI TAKAHASHI                      INTERNATIONAL COORDINATION OFFICE  
(000)000-0000                      (000)000-0000

HITACHI RESEARCH LAB. HITACHI, LTD.                      HITACHI RESEARCH LAB., HITACHI, LTD.  
OMIKA-CHO 7-1-1                      OMIKA-CHO 7-1-1  
HITACHI-SHI, IBARAKI-KEN, 319-12                      HITACHI-SHI, IBARAKI-KEN, 319-12  
JAPAN                      JAPAN-

Security class (U,C,S,TS) : U                      ONR resident rep. is ACO (Y/N): N  
Defense priority rating : NA                      NA supplemental sheet  
Equipment title vests with:      Sponsor X                      GIT  
NONE PROPOSED

Administrative comments -

INITIATION OF "FIXED PRICE" RESEARCH AGREEMENT.

GEORGIA INSTITUTE OF TECHNOLOGY  
OFFICE OF CONTRACT ADMINISTRATION

NOTICE OF PROJECT CLOSEOUT

Closeout Notice Date 01/28/97

Project No. C-36-X74\_\_\_\_\_ Center No. 10/24-6-R8763-0A0\_

Project Director HUDSON S E\_\_\_\_\_ School/Lab COMPUTING\_\_\_\_\_

Sponsor HITACHI LIMITED/JAPAN\_\_\_\_\_

Contract/Grant No. FOREIGN AGREEMENT\_\_\_\_\_ Contract Entity GTRC

Prime Contract No. \_\_\_\_\_

Title COLLABORATIVE RESEARCH WITH TOSHIFUMI ARAI AT HITACHI\_\_\_\_\_

Effective Completion Date 961120 (Performance) 961120 (Reports)

Closeout Actions Required:	Y/N	Date Submitted
Final Invoice or Copy of Final Invoice	Y	951107
Final Report of Inventions and/or Subcontracts	Y	_____
Government Property Inventory & Related Certificate	N	_____
Classified Material Certificate	N	_____
Release and Assignment	N	_____
Other _____	N	_____

Comments\_\_\_\_\_

Subproject Under Main Project No. \_\_\_\_\_

Continues Project No. \_\_\_\_\_

Distribution Required:

Project Director	Y
Administrative Network Representative	Y
GTRI Accounting/Grants and Contracts	Y
Procurement/Supply Services	Y
Research Property Management	Y
Research Security Services	N
Reports Coordinator (OCA)	Y
GTRC	Y
Project File	Y
Other _____	N
_____	N

NOTE: Final Patent Questionnaire sent to PDPI.

C-36-X74  
#1

## Hybrid Paper Electronic Interfaces (Augmented Reality) Project Final Report

Prof. Scott Hudson,  
GVU Center, and College of Computing  
Georgia Institute of Technology

Joint research with Visiting Scholar Toshifumi Arai  
Hitachi Research Labs

Paper is a very convenient mechanism for presenting information. It is familiar, inexpensive, flexible, portable, user modifiable, and offers better readability properties than existing electronic displays. However paper displays are static and do not offer capabilities such as dynamic content, searching, and hyperlinking capabilities that are available in electronic media. This project considered ways to add some of the functionality normally found only in electronic or computational systems to ordinary paper documents or books so that some of the good properties of electronic media could be added to those of paper (this is generally part of the overall research topic often called "augmented reality").

The specific activities of this year long project were to build a hardware and software system (called PaperLink) based on an augmented highlighter pen, and to report on it in the scientific literature. This system allows marks made on paper with a highlighter pen to have meaning and consequences in an electronic world. In particular this system allows highlighter pens marks to be used as hyperlinks from paper documents into electronic content.

This is done by augmenting an ordinary highlighter pen with a small video camera (see Figure 1 of attached paper). The video camera allows marks made on paper to be processed by image recognition software so that its unique shape can be recognized when it is encountered again, and so that any text it covers can be extracted and converted to ASCII by an optical character recognition subsystem. This allows a mark to act as a hyperlink. When a mark is made it is recorded and associated with a piece of electronic content (the *target* of the hyperlink). When the mark is encountered again, computer vision and image recognition techniques allow us to recognize the mark and to retrieve associated electronic content (hence following the hyperlink to its target object). Since the electronic content can be quite arbitrary -- an image, a video clip, a sound, a word processor document, or even a command to be executed, this capability is quite general and flexible.

The full details of the PaperLink system have been documented in a paper (attached). This paper has recently been accepted for publication in the proceedings of the annual ACM SIGCHI conference. This is perhaps the most prestigious venue for publication in the HCI community, and is very competitive. The conference employs full peer review, and is quite selective -- accepting only about 20% of the paper submitted to it. As a result, acceptance of this paper represents a significant independent validation of the quality of this work. The paper is also available on the web as:

[http://www.cc.gatech.edu/gvu/people/faculty/hudson/hybrid\\_paper/chi97/seh.htm](http://www.cc.gatech.edu/gvu/people/faculty/hudson/hybrid_paper/chi97/seh.htm)

and

[http://www.cc.gatech.edu/gvu/people/faculty/hudson/hybrid\\_paper/chi97/paperlink.final.ps](http://www.cc.gatech.edu/gvu/people/faculty/hudson/hybrid_paper/chi97/paperlink.final.ps)

In addition to publication of the work in a peer reviewed setting, this project also resulted in a masters thesis by Dietmar Aust who was a visiting student from the University of Dortmund in Germany (graduating from there in the fall of 1996). This thesis is also available on the web as:

[http://www.cc.gatech.edu/gvu/people/faculty/hudson/hybrid\\_paper/aust\\_ms\\_thesis.ps](http://www.cc.gatech.edu/gvu/people/faculty/hudson/hybrid_paper/aust_ms_thesis.ps)

# PaperLink: A Technique for Hyperlinking from Real Paper to Electronic Content

**Toshifumi Arai**

Hitachi Research Laboratory  
Hitachi, Ltd.  
1-1 Omika-cho 7-chome  
Hitachi, Ibaraki 319-12 Japan  
arai@hrl.hitachi.co.jp

**Dietmar Aust**

University of Dortmund  
Department of Computer Graphics  
Otto-Hahn Str. 16  
44221 Dortmund, Germany  
aust@ls7.informatik.uni-dortmund.de

**Scott E. Hudson**

Georgia Institute of Technology  
Graphics, Visualization, and  
Usability Center, and College of  
Computing  
Atlanta, Georgia 30332-0280 USA  
hudson@cc.gatech.edu

## ABSTRACT

Paper is a very convenient medium for presenting information. It is familiar, flexible, portable, inexpensive, user modifiable, and offers better readability properties than existing electronic displays. However, paper displays are static and do not offer capabilities such as dynamic content, and hyperlinking that can be provided with electronic media. PaperLink is a system which augments paper documents with electronic features. PaperLink uses a highlighter pen augmented with a camera, along with simple computer vision and pattern recognition techniques, to allow a user to make marks on paper which can have associations and meaning in an electronic world, and to "pick up" printed material for use as electronic input. This paper will consider the prototype PaperLink hardware and software system, and its application to hyperlinking from paper to electronic content.

## Keywords:

Augmented Reality, Input Devices, Hybrid Paper Electronic Interfaces, Computer Vision Systems, Pattern Recognition, Hyperlinking.

## INTRODUCTION

Conventional paper materials have significant merits. They are ubiquitous, highly portable, and easy to use in a wide range of environments. Paper is inexpensive, can be annotated easily, and provides excellent readability properties [1]. Further, paper documents offer excellent ergonomic properties — no current electronic displays can approach the convenience, "feel", and ease of use of, for example, a paperback book.

On the other hand, electronic presentation technologies allow us to create and publish great quantities of information and can make that information available world-wide in literally a matter of seconds. Further, electronic materials

offer considerably more powerful capabilities such as searching, dynamic multimedia content, hyperlinking, and easy reuse. Finally, going to an electronic medium opens the door to not just information, but to computation — a fundamentally more powerful capability.

If we could add some of the advantages of electronic materials to paper materials, for example if we could use paper materials as interfaces to electronic materials, it might be possible to combine some of the good properties of each into hybrid systems which approach the convenience and ease of use of paper, while allowing access to the power of computational media. This is the underlying goal of our work, and the PaperLink system described here takes one step in that direction.

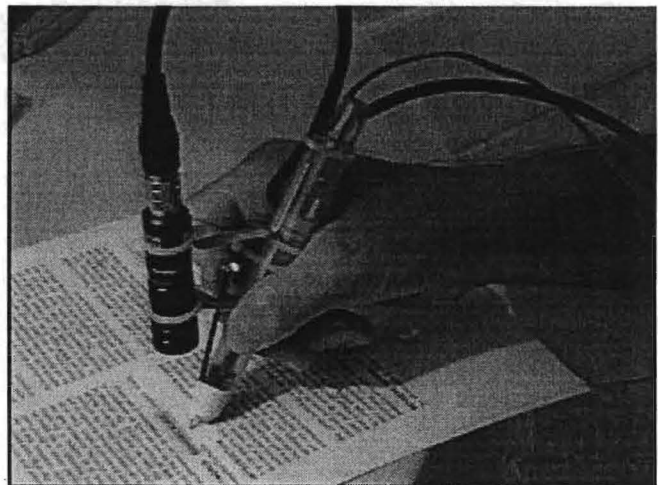


Figure 1. A Photo of the VideoPen Prototype

## FEATURES OF PAPERLINK

The PaperLink system provides two primary capabilities. First, it allows marks made on paper, either with a highlighter or standard pen, to be associated with electronic content and/or assigned a meaning. Second, it allows words on paper to be "picked up" and used as input.

The first of these capabilities is used to provide the basic paper to electronic hyperlinking mechanism that is the subject of this paper. To create a link from paper, the user employs a special augmented

highlighter pen as shown in Figure 1. This pen, which we call the VideoPen, is a retractable highlighter pen which has been augmented with a miniature video camera and a tip switch. The visual properties of marks made with this pen can be recorded and associated with a piece of electronic content (or in general an executable command) — thus establishing a hyperlink. When the mark is returned to later, its associated content can be retrieved and displayed (or executed) — thus following the link.

The second capability, to "pick up" small pieces of text from the paper, allows additional capabilities to be supported. Captured image data which is not recognized as a hyperlink mark can optionally be fed to an OCR routine to extract the corresponding text. This text can then serve as a parameter to a command, or otherwise used as input data. This is useful, for example, to record page numbers. Allowing us to provide an at least rudimentary electronic to paper hyperlinking and indexing mechanism.

Since the system can associate executable commands as well as content with marks, a convenient control mechanism in the form of a simple paper bookmark can be provided. This bookmark initially has marks associated with standard commands such as creating a hyperlink and recording a page number. In addition, the user can extend this set simply by writing down a new command symbol or word on the paper bookmark, and entering or composing the command to be associated with it. This provides a very convenient and natural extension mechanism.

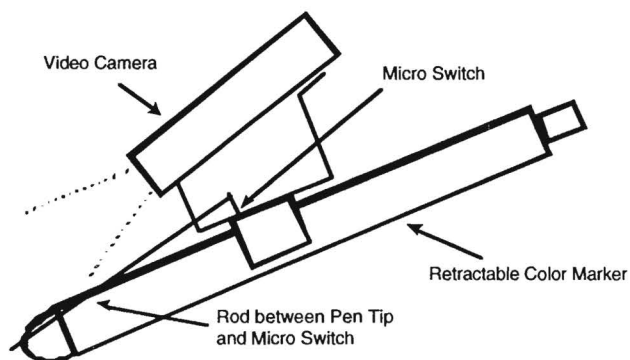


Figure 2. VideoPen Components.

#### THE VIDEOPEN

The VideoPen is the primary input device for the PaperLink system. Figure 1 shows a photo of the current prototype of the VideoPen, and Figure 2 indicates its components. A small video camera is attached to a color highlighter pen with a retractable pen tip. Since the camera is quite close to the paper, a normal resolution (NTSC) camera is sufficient to recognize words or patterns on paper. A micro switch is also attached to the pen and there is a thin rod between the end of the pen and the

switch to provide a tip switch (this switch is designed for use primarily when the highlighter tip is inside of the pen).

In the current prototype, the video camera is attached to a frame capture card in a conventional desktop PC and the tip switch is wired to the switch of an input device button (in our case a trackpad pointing device).

Our preliminary experience with the prototype VideoPen shows that it is quite usable, but a bit more cumbersome than we would like. We envision that in future systems the camera could be made smaller and placed closer to the pen (for example, [2] describes a device with a camera *inside* a ball-point pen) and that a powerful processor and/or a wireless link to a desktop machine could be placed inside the pen to eliminate our current hardwired link.

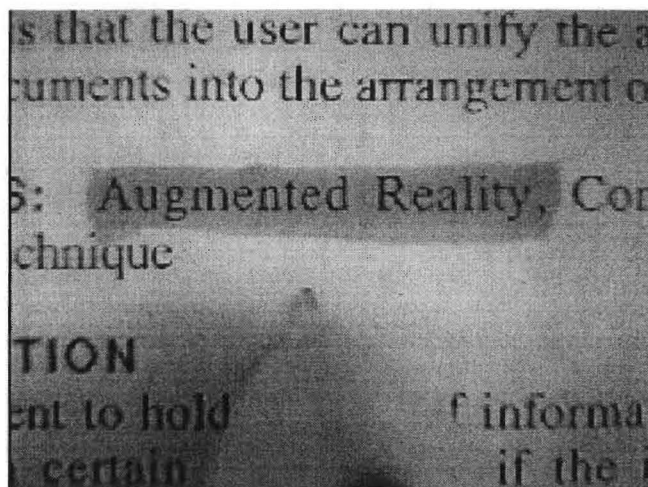


Figure 3. Image from the VideoPen Camera

The PaperLink system captures an image when the tip switch is triggered so that a user can "pick up" marks, words, or patterns on paper. The camera captures a 2 x 1.5 inch image around the pen tip. Figure 3 is an example of an image captured with the VideoPen, in this case a hyperlink mark.

After capturing an image, the pattern which appears at the center of the image is extracted. If a contiguous colored region above a threshold size is found, this is recognized as a highlighter mark and is extracted as shown in Figure 4.

For pattern recognition purposes, this image will be treated primarily as a *shape* (considering both the outer boundary of the mark and the inner boundaries of the text it covers). As described below, a feature based classification system is employed to store the characteristics of this shape (when establishing a hyperlink) and recognize it again later (when following a hyperlink). Since each highlighting mark is drawn by hand, and the extracted image includes interior boundary data corresponding to the covered text, the shape of the



mark will normally be quite unique, and in practice has been quite easy to recognize accurately.

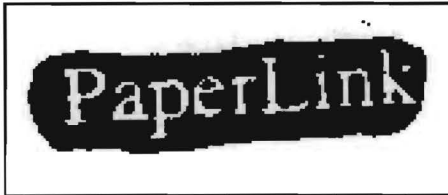
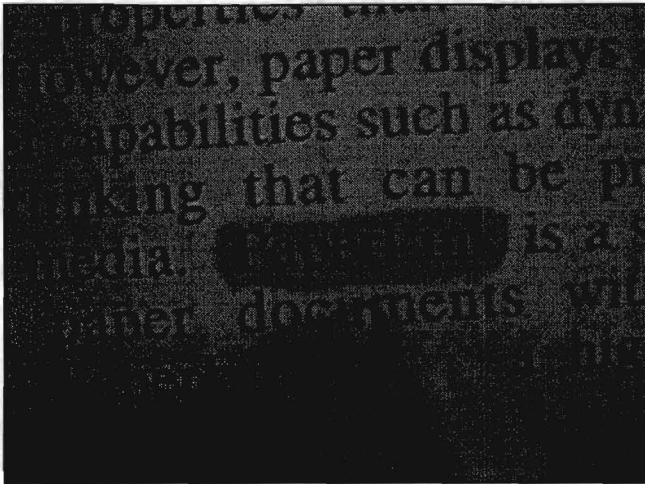


Figure 4. Original and Extracted Color Region.

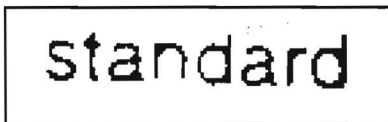
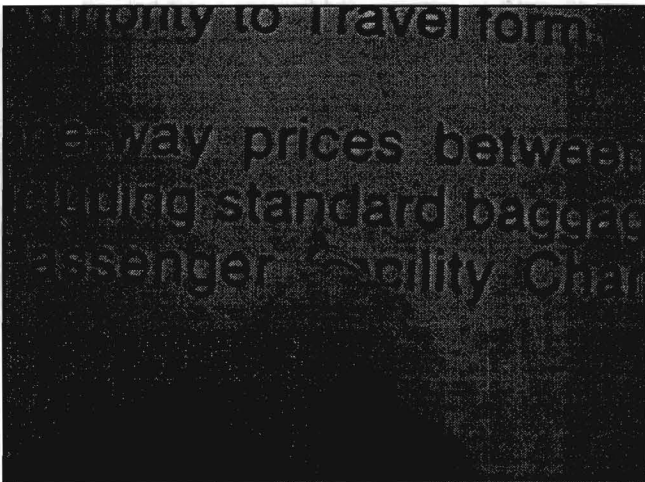


Figure 5. Original and Extracted Text Image

If no colored region is found then the system assumes the user is interested in the text or pattern itself. In this case, the word near the center of the image is extracted (using a simple heuristic that considers the relative sizes of gaps between connected regions) and a thresholded image like that shown in Figure 5 is extracted for processing.

This image can be used directly as a pattern or later passed to an OCR routine to attempt to extract the corresponding ASCII text.

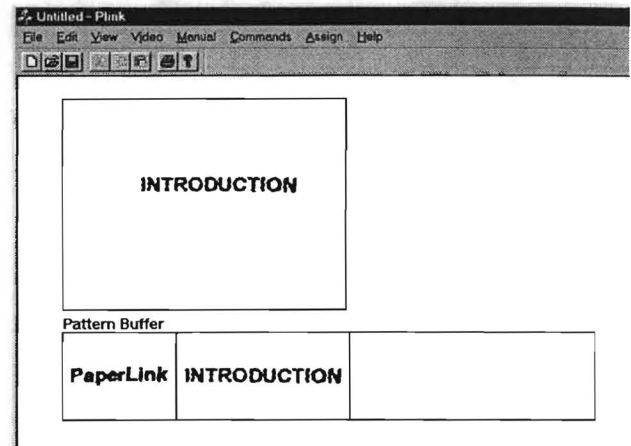


Figure 6. The Pattern Buffer

#### INTERACTION

As each image is captured and processed it is classified as either a command or as data (hyperlinks are treated internally simply as a command for displaying the associated electronic content). Data images are placed in a pattern buffer as shown in Figure 6 (this also shows a large window with the current image which would not appear in many applications). When a command is encountered the current contents of the pattern buffer is used as its parameters. Particular commands can also request that the OCR routine be applied to one or more images and receive the resulting text instead of, or in addition to, the image data.

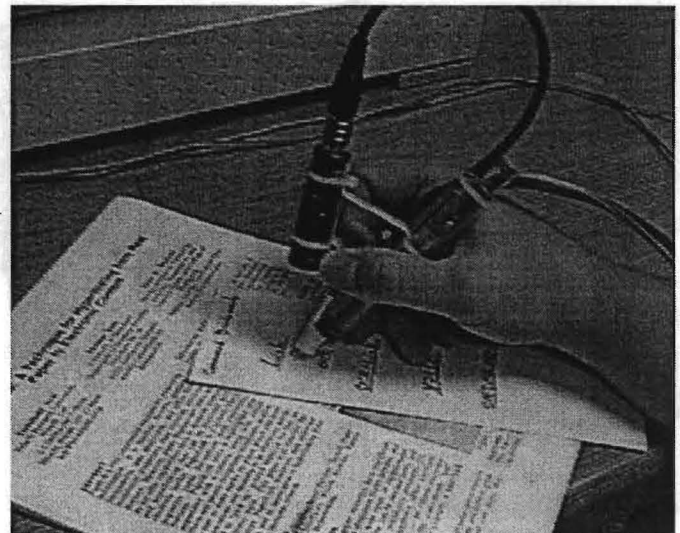


Figure 7. A Command Bookmark

As indicated above, the system uses a paper bookmark to provide standard and user supplied commands. As shown in Figure 7, this bookmark is simply a strip of paper with standard commands either printed and highlighted, or hand written on it.

Each of the standard command images is registered in advance and associated with the appropriate action.

Central among the predefined standard commands is the one used to create a hyperlink. This command expects to find a data image in the pattern buffer. As a result a typical usage scenario for creating a hyperlink would include: making a highlight mark on the paper, recording that mark in the pattern buffer by pressing the tip switch on the paper just under the mark, and selection of the command to make a link (again by depressing the tip switch just under the appropriate command pattern on the command bookmark). Electronic content to be associated with the mark would then be prompted for in an application specific way (in one of our prototypes a simple file picking dialog box is used, in other cases more sophisticated actions like making an audio recording could be performed). Finally, the features of the hyperlink mark image are stored in a database, along with an appropriate display command, and a reference to the associated electronic content.

If this hyperlink mark is then entered again later (by depressing the tip switch directly under its highlight mark), it will be located in the command database and executed to retrieve the associated content.

Multiple bookmarks with different common command sets, or command sets for different applications, can easily be prepared in advance. Even bookmarks with small keypad or keyboard images could be used to eliminate the need for a traditional keyboard in non-text intensive applications.

In addition, since the command bookmark is simply a piece of paper, the end-user is free to write in new commands or make hyperlink marks to commonly used information at any time. We believe that this free-form end-user extension mechanism has considerable potential because of its flexibility and informality. We are currently beginning work to look at new techniques for specifying commands (perhaps by demonstration [3] or based on simple composition of existing commands) which are suitable for use by non-programming end-users in this context.

## IMPLEMENTATION DETAILS

### Structural Overview

As mentioned above, the desktop machine hosting the PaperLink system has a video capture board to input video images from the VideoPen, and a track pad for pointing (this turns out to be more convenient to use than a mouse since it can be placed anywhere and can be employed while still holding the VideoPen). These are the only additions made to our standard desktop PC for this system.

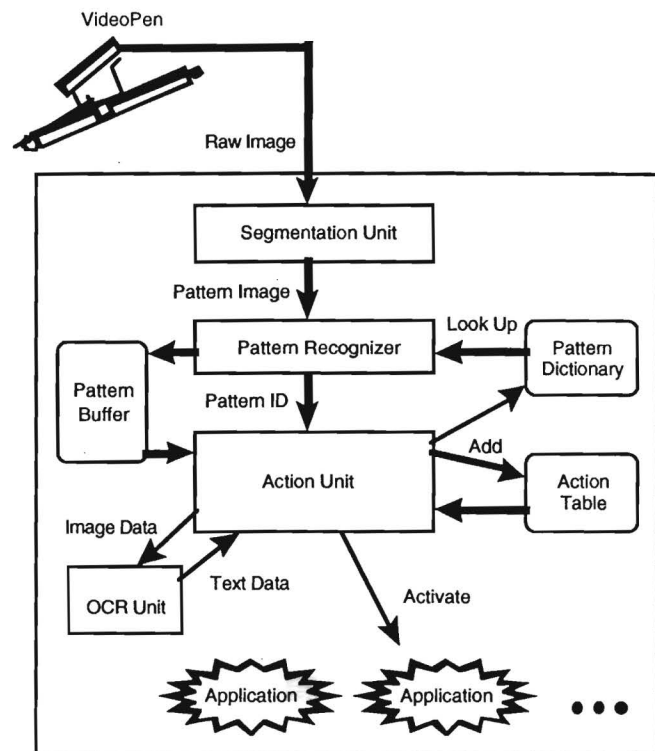


Figure 8. Software Architecture of the System.

Figure 8 shows the software structure of the PaperLink system. An image from the VideoPen is first sent to the segmentation unit, and the unit extracts the pattern that a user is pointing with the VideoPen. If a colored region is found around the center of the image, the segmentation unit extracts the region. Otherwise, the unit extracts the black and white pattern corresponding to the text itself.

Then the pattern is sent to the pattern recognizer. If the recognizer can not identify the pattern as being associated with a previously entered command, it is stored in the pattern buffer. On the other hand, if the pattern is identified in the existing pattern dictionary, the result, a pattern id will be sent to the action unit.

The Action unit extracts the appropriate command from an action table, and activates the command with the pattern(s) in the pattern buffer. The pattern(s) from the pattern buffer can also optionally be processed with the OCR unit before being passed to the command.

### Segmentation and Pattern Recognition

The recognition of patterns that is central to the operation of the PaperLink system proceeds in three steps: segmentation, feature extraction, and classification. The complete details of these steps is described in [4], only a brief outline will be provided here.

Segmentation for images with highlighted areas is done with a color filtering algorithm. To improve robustness the raw red green and blue color values

are normalized into chromaticity coordinates (see [5]) and a Euclidean distance in the two dimensional chromaticity space is used for color similarity. An adaptive thresholding technique is used to identify an initial colored region and a 4-neighbor connected component labeling algorithm (the "blob coloring" algorithm described in [6]) is used to refine this region and extract the largest colored region. An additional filtering pass based on greyscale thresholding is used to improve rejection of interior black pixels that are part of printed text within the highlight (these locations pick up ink from the highlighter pen and hence contain small amounts of the highlight color which causes them to be occasionally classified improperly). Once the largest colored region has been identified, a boundary following algorithm creates a chain-coded representation of the outer boundary of the region. Finally, interior connected regions within this boundary are extracted — these represent the interior highlighted text.

For non-highlighted areas a simpler thresholding process is used to initially identify text versus non-text areas. Connected regions (typically corresponding to letters) are extracted, and a simple gap size threshold algorithm is used to select a set of regions approximating a word or tightly grouped pattern near the center of the image.

Once segmentation has been completed a series of features are computed for the selected image components. A feature is a numerical quantity that represents some image property useful for distinguishing it from other candidate images during pattern recognition. For the PaperLink system, all features need to be invariant to translation, scale, and rotation since the exact camera position and orientation cannot be precisely controlled.

Features computed include: an approximation of eccentricity, a compactness metric, the number of extracted character regions, the ratio of colored region to text region pixels, three boundary signature features [7], and three features based on shape moments [8] (again see [4] for the complete details).

Classification of images is done using a simple *nearest neighbor* algorithm. Once a set of features have been computed for an image (and normalized to account for metric range differences), the feature set can be treated as a vector representing the coordinates of a point in a high dimensional feature space. The similarity of two images can then be measured in terms of a Euclidean distance within that space. Pattern recognition is then performed on the basis of this similarity. When a new image is encountered, its feature vector is computed. This vector is then compared against the vectors associated with previously stored images. If the closest vector (representing the nearest neighbor

image) is within a threshold distance, the image is classified as a match to the stored image.

Taken together the procedure outlined above provides robust pattern recognition results for our application.

#### Code

The image capture, pattern recognition, and other functional core components of the PaperLink system are coded in C++, with a commercial OCR software package being used to recognize words picked up from paper. Finally, Visual BASIC™ is used as "glue" between the core modules and various application programs.

#### APPLICATIONS

Materials used with the PaperLink system can come in three forms: preexisting *arbitrary* content, *known* content, and *controlled* content. In the first case, the system is capable of operating with any existing printed material. This "walk up and use" capability provides a significant advantage since the world is full of printed documents which we would like to integrate with electronic content, but which has not been prepared in advance.

If more information is available about the paper content in advance, additional features can be provided. In particular, if the full content of the paper document is known in advance (for example from having extracted its text via OCR), then indexing and cross reference capabilities can be provided. In addition, it may be possible to automatically determine location within the document by examining the set of words contained in the several line fragments within view of the camera. These words can then be used to query the corresponding electronic text for a location. In most types of documents this will produce a unique match and determine a location (in the worst case we would expect a very small number of non-unique locations, and then only rarely).

If paper content is controlled — that is we can print it to our own specifications — then even more features can be included. For example, preexisting hyperlink symbols can be included in the printed content (perhaps to material accompanying a book on a CD ROM) and/or symbols for standard commands could be printed in the margins.

Overall the three types of content provide a convenient scale of tradeoff between the ability to deal with arbitrary printed content and the additional features that can be provided if advance preparation and integration with the system is possible.

#### Prototypes

For our initial investigations we began looking at arbitrary content applications first. To explore possible interaction techniques before the PaperLink system was complete we acquired a



commercially available OCR pen scanner [9]. This device places a small optical scanner in a pen sized package which is connected to OCR software on a PC. This device is capable of extracting ASCII from preprinted text, but does not appear appropriate for dealing with user created marks. Thus it allows text to be "picked up" from paper and used as input, but does not directly support the marking and pattern recognition needed for hyperlinking (with some modifications, however, it might be suitable for following prepared hyperlinks in controlled content applications).

Using this device we built an English-German dictionary application in order to explore alternative interaction techniques based on the command bookmark concept. This application allows German words to be "picked up" and used for dictionary and thesaurus lookup by students reading German. In addition a simple annotation capability was provided which allows comments to be entered and associated with various words or phrases.

Based on our experience with bookmark interactions using the more limited pen OCR device, we constructed a prototype paper to electronic hyperlinking application with the completed PaperLink system. This prototype application supports a wider range of hyperlinking and annotation activities connecting arbitrary paper content to user created annotations and richer electronic content. As described in previous sections this prototype allows the user to make marks with the highlighter pen and to record associations between those marks and commands to be executed in the electronic world. In addition new images, and optionally the result from OCR of these images, are saved in a buffer and can be applied as the parameters to commands associated with recognized images. This control setting is very similar to control framework used in the early MIKE UIMS [10]. Because any executable commands can be linked to traversal of a paper hyperlink, this basic framework is very flexible and can be employed in a number of more specific applications. For example, we have employed this capability in conjunction with a commercial "street finding" program for retrieving maps, and an electronic encyclopedia for concept lookup.

Although at the time of this writing the PaperLink system has not been stable and robust long enough to conduct realistic user studies, our early and informal experience with the system has been positive.

#### RELATED WORK

There are several augmented reality systems which enable users to retrieve electronic materials through manipulating real objects. As with PaperLink, several of these systems feature video cameras to monitor and recognize objects in the real world.

One of the first works in this area, DigitalDesk [11, 12], was a computer-augmented desk which has a video camera above the user to monitor the real desktop. This system can detect the movement of a user's hands and the contents of documents on the desk. However, it is extremely hard to capture enough detail of the contents with typical cameras in that setting, and the size of the augmented area is limited by the resolution of the camera. This system first presented the concept of using a piece of paper placed within view of the camera as a button to invoke a command — a concept similar to the command bookmark in PaperLink.

Ariel [13] is a system to augment engineering documents, and also has a video camera to monitor the documents. In this case the camera is not used to detect the contents of the documents, but is used to detect x-y positions of red-colored pointers. Therefore, the system requires some other method to know what document is placed in its view and to align or *register* the document with the camera.

The InteractiveDESK system [14, 15] is also a computer-augmented desk which responds to operations on real objects to assist users. This system also has a camera above the desk monitoring its real desktop. In this case, the system can identify the objects on its desktop by reading special color coded tags attached to the objects. This allows the system to, for example, associate a document with the file folder it is stored in, and to make associations between documents commonly used together. However, the system is limited to a coarse resolution. It can make links and associations with whole objects, but can not make links from the contents of the objects, such as words or marks on pages in documents.

NaviCam [16] is a portable video see through system. The idea behind this system is to point a hand held camera at objects and view augmentations of those objects through the camera image. NaviCam also uses color coded strips to recognize objects and hence operates at a similar granularity to the InteractiveDESK system.

Compared with those systems, the major advantage of PaperLink is that it can operate on the actual content of paper objects. It can capture enough details of the words or marks a user is pointing to (with a normal resolution camera, and without a registration process) to allow them to be used reliably as input. On the other hand, because VideoPen only captures a very small area at a time the PaperLink system cannot make use of document context, and it is hard to know what is happening in a large area such as a desktop. Consequently, this system is complementary to coarser grained approaches described above.

The MEMO-PEN [2] uses a related hardware system for a different task. The MEMO-PEN

places a single chip video camera, a strain gauge, and a small processor inside a pen. The camera is placed above a very short ball-point pen cartridge and views the paper through the transparent lower end of the pen. A series of images along with directional information from the strain gauge is captured in order to record and later recognize the user's handwriting. Since this information is stored completely within the pen only relatively low resolution images are captured and recorded.

In the paper-based audio notebook described in [17], a user's handwritten notes are synchronized with an audio recording, and the user can retrieve the part of audio data by specifying a part of memo. The hardware setting and the interaction techniques are different from PaperLink, but the basic idea of providing access to electronic materials with something written on paper is the same. The system does not recognize what is written on paper, so it does require registration methods. In this case it uses optical sensors to read special marks prepared in advance to detect a page number, and uses a digitizing tablet to detect positions within the page.

Finally, the XAX [18] and Protofoil [19] systems provide retrieval of electronic documents with requests written or printed on a paper form and FAXed to and from a server. These systems use pencil and paper along with a conventional office equipment for linking paper to computational media. However, because of the use of hardware involved, they are not very interactive in nature and most suitable for applications based on paper forms.

#### CONCLUSION AND FUTURE WORK

In this paper we have described the PaperLink hybrid paper electronic system. This system allows marks made on paper to have associations and meaning in an accompanying electronic world. The goal behind this work has been to bring some of the capabilities of computational media to the conventional and convenient every day world of paper documents. The PaperLink system, with its use of the familiar and metaphorically useful highlighter pen, along with its informal user extensible bookmark command menus is one step in that direction.

Based on our experience with the PaperLink system, we see several areas for future work. First, in addition to exploring ways to package the VideoPen more compactly and to eliminate a wired link to a desktop machine, we are interested in exploring the use of audio instead of video feedback in interfaces constructed with the system. These steps would bring us closer to a stand-alone device which could eventually be more like a familiar and ubiquitous highlighter pen and less like an exotic computer input device.

Another area of particular immediate interest is the use of the system in a collaborative setting. We envision that in, for example, a collaborative writing setting, a user could comment on a paper document and then share those comments with others via their highlighting marks on paper. Using controlled content techniques, special images reminiscent of the "change bars" used in some documents to indicate areas of revision could be printed in the margins to preserve links to an accumulated electronic commentary or history of the document.

Finally, we would in the future like to look at employing continuous video rather than still frames captured only on request. This would allow for more fluid interaction (although it also presents the problem that there would be no clearly defined "point of activation"). This could also help overcome the limited field of view of our current prototype by allowing larger image areas to be built up as a mosaic of smaller images (hence allowing more multiple word input, and in general more context). To support continuous video input faster and more sophisticated computer vision techniques will be needed which can dynamically acquire and segment highlighted areas at (near) video rates.

#### ACKNOWLEDGMENTS

The authors wish to thank Jen Mankoff for her work on the first German to English dictionary application. This work was conducted under cooperative research funding from the Hitachi Research Laboratory, Hitachi Ltd. Additional funding was provided by the National Science Foundation under grants IRI-9500942 and CDA-9501637.

#### REFERENCES

1. Hansen, W., HAAS, C. "Reading and Writing with Computers: A Framework for Explaining Differences in Performance" CACM Vol.31 No.9, 1080-1089, 1988.
2. Nabeshima, S., Yamamoto, S., Agusa, K., Taguchi, T. "MEMO-PEN: A New Input Device" CHI'95 Conference Companion, 256-257, 1995
3. Cypher, A.,(ed), "Watch What I Do: Programming by Demonstration", MIT Press, 1993.
4. Aust, D., "Augmenting Paper Documents with Digital Information in a Mobile Environment" MS Thesis, University of Dortmund, Department of Computer Graphics, 1996. Also available electronically as [http://www.cc.gatech.edu/gvu/people/Faculty/hudson/hybrid\\_paper/aust\\_ms\\_thesis.ps](http://www.cc.gatech.edu/gvu/people/Faculty/hudson/hybrid_paper/aust_ms_thesis.ps).
5. Foley, J., van Dam, A., Feiner, S., and Hughes, J., "Computer Graphics, Principles and Practice", Addison Wesley, 1990.

6. Ballard, D. and Brown, C., "Computer Vision", Prentice-Hall, 1982.
7. Bryant, A., and Bryant, J., "Recognizing Shapes in Planar Binary Images", *Pattern Recognition*, 22(2), 155-164, 1989.
8. Dai, M., Baylou, P., and Najim, M., "An Efficient Algorithm for the Computation of Shape Moments from Run-Length Codes or Chain Codes", *Pattern Recognition*, 25(10), 119-1128, Feb. 1992.
9. Primax Electronics Ltd., "DataPen™ User's Guide", 1995.
10. Olsen, D., "MIKE: The Menu Interaction and Kontrol Environment", *ACM Transactions on Graphics*, 5(4), 318-344, 1986.
11. Wellner, P. "The DigitalDesk Calculator: Tangible Manipulation on a Desk Top Display" *Proceedings of UIST'91*, 27-33, 1991.
12. Newman, W., Wellner, P. "A Desk Supporting Computer-based Interaction with Paper Documents" *Proceedings of CHI'92*, 587-592, 1992.
13. Mackay, W.E., Pagani, D.S., Faber, L., Inwood, B., Launiainen, P., Brenta, L., Pouzol, V. "Ariel: Augmenting Paper Engineering Drawings" *CHI'95 Conference Companion*, 421-422, 1995.
14. Arai, T., Machii, K., Kuzunuki, S., Shojima, H. "InteractiveDESK: A Computer-Augmented Desk Which Responds to Operations on Real Objects" *CHI'95 Conference Companion*, 141-142, 1995.
15. Arai, T., Machii, K., Kuzunuki, S. "Retrieving Electronic Documents with Real-World Objects on InteractiveDESK" *Proceedings of UIST'95*, 37-38, 1995.
16. Rekimoto, J., Nagao, K. "The World through the Computer: Computer Augmented Interaction with Real World Environments" *Proceedings of UIST'95*, 29-36, 1995.
17. Stifelman, L. "Augmenting Real-World Objects: A Paper-Based Audio Notebook" *CHI'96 Conference Companion*, 199-200, 1996.
18. Johnson, W., Jellinek, H., Klotz, L., Rao, R., Card, S. "Bridging the Paper and Electronic Worlds: The Paper User Interface" *Proceedings of INTERCHI'93*, 507-512, 1993.
19. Rao, R., Card, S., Johnson, W., Klotz, L., Trigg, R. "Protofoil: Storing and Finding the Information Worker's Paper Documents in an Electronic File Cabinet" *Proceedings of CHI'94*, 180-185, 1994.