PaperUI

Qiong Liu and Chunyuan Liao

FX Palo Alto Laboratory
3400 Hillview Avenue, Bldg.4
Palo Alto, California, U.S.A.
{liu,liao}@fxpal.com

Abstract. PaperUI is a human-information interface concept that advocates using paper as displays and using mobile devices, such as camera phones or camera pens, as traditional computer-mice. When emphasizing technical efforts, some researchers like to refer the PaperUI related underlying work as interactive paper system. We prefer the term PaperUI for emphasizing the final goal, narrowing the discussion focus, and avoiding terminology confusion between interactive paper system and interactive paper computer [40]. PaperUI combines the merits of paper and the mobile devices, in that users can comfortably read and flexibly arrange document content on paper, and access digital functions related to the document via the mobile computing devices. This concept aims at novel interface technology to seamlessly bridge the gap between paper and computers for better user experience in handling documents. Compared with traditional laptops and tablet PCs, devices involved in the PaperUI concept are more light-weight, compact, energy efficient, and widely adopted. Therefore, we believe this interface vision can make computation more convenient to access for general public.

Keywords: paper document, camera phone, human-computer interface, human-computer interaction, document recognition, augmented paper, vision-based paper interface.

1 Introduction

Let's think about an ideal reading device without restricting by state-of-the-art technologies. Is it iPad? In our mind, our ideal reading device should be more interactive than iPad. Moreover, it should be much larger than iPad (for comfortable reading), much smaller than iPad (for portability), much lighter, much cheaper, can be merged or separated easily, and have better readability. We believe that PaperUI is a step toward this goal.

After its invention several thousand years ago, paper has become an essential part of our daily life. Paper production provides a cost efficient way to use wood mill wastes. Additionally, paper can be recycled or decomposed much easier than most plastics or electronic-devices. According to the statistics published by woodconsumption.org [9], the world produces approximately 300 million tons of paper each year and every US office employee generates approximately 9,999 more paper sheets each year. All these

M. Iwamura and F. Shafait (Eds.): CBDAR 2011, LNCS 7139, pp. 83–100, 2012.

facts indicate that paper is one of the most widely used communication media and will continue to be used in our daily life for a very long time.

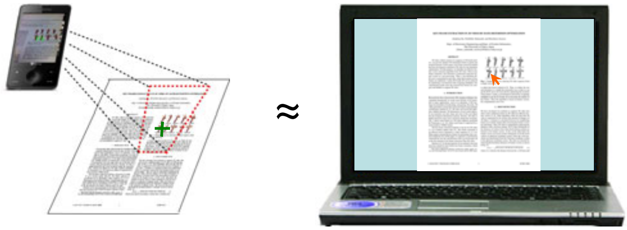


Fig. 1. PaperUI vs. Traditional Laptop UI

On the other hand, technology advances in computer/consumer-electronic industry and chemical industry enable many novel products that can replace or surpass traditional paper products. These new exciting products stimulate people to imagine paperless office from 40 years ago [39]. However, many e-products finally stimulate paper usage and create more e-wastes and plastic wastes beyond paper in the past 40 years.

To make more efficient use of paper products as well as novel electronic devices, new technologies are demanded to combine the merits of both paper and electronic devices. PaperUI is an initial attempt to address this issue. Figure 1 illustrates the PaperUI concept vs. a traditional laptop UI. The PaperUI concept advocates using paper as static content displays for its wide adoption, good readability, flexible display size, flexible display arrangement, robustness, energy efficiency, light-weight, and less long term hazards. It also advocates dynamic human-information interaction via light-weight and energy efficient electronic devices for its portability, fast warm-up time, low energy cost and low long-term e-waste hazards. Through properly balancing information displayed on paper and dynamic mobile display, PaperUI may also help us to reduce hardcopies without sacrificing convenience or changing working habits.

The PaperUI concept is not a concept that comes from nothing. Many digital function designs of this interface come from existing computer interfaces. Therefore, it can provide most digital functions available on state-of-the-art desktop or laptop computers. Different from existing computer interfaces that provides an exclusive working space (paperless working space), the PaperUI interface tries to cooperate with the traditional paper interface to benefit both. This tight cooperation seamlessly merges the physical working space and digital working space. It can save users' efforts for switching back and forth between two different working spaces. It can also save energy by not using power for displaying most static contents.

PaperUI is an add-on interface for existing paper interface widely adopted by people for many centuries. Because it is an add-on interface, anyone can selectively use this interface. When people choose to use the PaperUI interface, they can get state-of-the-art digital functions beyond traditional paper interface. When they choose not to use the PaperUI interface, they still can use the traditional paper interface as usual.

In the following sections, we first present PaperUI overview, then we present emerging technologies that may be helpful to the PaperUI implementation, followed by research prototypes and applications that align with the PaperUI concept. We end the paper with discussions of future work.

2 PaperUI Overview

PaperUI can be situated in the design space of interactive paper systems. Early interactive paper system research such Digital Desk [41], achieves the augmentation of paper by using fixed cameras above a desk to track documents as well as gestures. Because this setup has fixed camera and desk, the mobility of this system is very limited. This fixed camera setup also restricts efficient usage of the camera resolution. Moreover, that system has to use powerful projectors to cover the whole desktop for usable feedback. Different from early interactive paper systems, the PaperUI concept offers more considerations to system mobility, interaction resolution, and energy efficiency.

Different from interactive paper computer [40], which is a film-like electronic device, PaperUI only uses mobile electronic device as a small interactive window, and does not extend electrical wires to the whole paper surface. This setup concept can help the PaperUI interface to reduce energy cost and e-waste. If the interactive paper computer can finally be used as the small interactive window by the PaperUI interface, the energy and e-waste reduction is expected to be more significant.

To improve mobility, researchers use Anoto's digital pen [42] or similar tracking device [43] to interact with paper. Early application explorations of the pen technology use pen path tracking to fill digital forms [44], use pen position on a piece of paper to activate sound corresponding to that position [45], use pen gesture tracking to capture handwritten notes and activate digital functions [48], or even use pen position and pen gesture to manage meeting capture and activate room controls [47]. These explorations may be considered as early prototypes of the PaperUI concept. Even though digital pen and similar tracking devices have better mobility than fixed systems, they do require users to carry a special device for the interaction. This special requirement may hinder wide adoption of this pen based technology. Another major drawback of these early prototypes is that they do not provide much active visual feedback. This drawback severely limits the digital functions that can be offered.

To overcome this extra device barrier and visual feedback barrier, researchers start to develop PaperUI technologies that can enable user-paper interaction via widely adopted smartphones. Beyond the wide adaptability, the smartphone-based interface has an extra display for visual feedback. When the display has a touch input, the PaperUI interface is further enhanced for more accurate user-document interaction.

3 Emerging Technologies for Realizing the PaperUI Concept

Although paper is one of the most widely used viewing devices, it cannot play dynamic media such as video and cannot be used to access the Web. It also lacks digital functions such as copy-paste, or search. Cellphones and other portable electronic devices are commonly used to play video and access Web pages, but do not have the affordances of paper such as high resolution and readability. Paper patch identification technology advances can provide the best of both devices by linking digital media to paper documents. To access digital functions associated with a paper document patch, a mobile device (e.g. camera phone) is used to identify a document patch, and digital functions associated to that patch in the document are enabled on the mobile device. With this approach, paper surface is used as a counter part of a traditional computer display and the display-equipped mobile device is used as counter parts of a traditional computer mouse and the mouse pointed sub-region on a traditional display. By building a PaperUI system like this, most digital functions supported by a traditional PC/laptop interface may be enabled for a PaperUI interface with more portability, less energy cost, and less e-wastes.

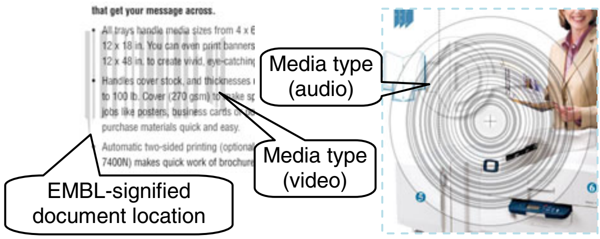
Now, there are mainly seven approaches to identify a document patch. The first approach is to print a barcode or QR code [1] on the image patch for identification. The second approach is to use micro-optical-patterns such as Dataglyph [29] on document patches for identification. The third approach is to modify document content to encode hidden information for identification [5]. The fourth approach is to index underlying paper fingerprint for document identification [4]. The fifth approach is to use OCR or character recognition outputs for identification [6,21]. The sixth approach is to index printed document content features and use these features to identify document [9,10,11,12,13,32,33]. The seventh approach is to put an RFID on the document patch for identification [14,19].

3.1 Barcode

Barcode is an optical marker printed on paper or other object. By changing a series of special patterns' color, shape, thickness and spacing, barcode can encode an ID number or other information associated with its hosting object. Barcode was first invented in 1948 by Bernard Silver [1] and had its first successful commercial use for supermarket checkout in 1974 [8].

Because of its long history, barcode gets used by most people and gets nearly unbeatable robustness and affordance. According to [2], the worst case accuracy for the old UPC code is 1 error in 394 thousand and the worst case accuracy for the new Data Matrix code is one error in 10.5 million. Meanwhile, the cost for providing a barcode is under 0.5 penny. With its unbeatable affordance and robustness, we see more and more barcode printed on document to track document category, price, or even support some basic interaction.

The big disadvantage of using barcode on document is caused by its opaque property. Regular barcodes are visually obtrusive. That makes barcode printing interfere with the document content layout. This fact generates a series of barriers for using barcode to create multiple links to the same document page. First, if we put too many barcodes on the same document page, we will have much less space for real contents. Even though the cost for producing a barcode is cheap, wasting content space may create more cost than producing a barcode. Second, changing original document layout with many obtrusive barcodes also makes the document uglier for many readers. Third, intensive barcode use on document may increase paper waste. Fourth, barcode printing may require changes to a traditional printing process. Besides these disadvantages, traditional barcode is also not very suitable to indicate media link type associated with a barcode. More specifically, if we simply print a barcode on a document page, most users will think it is catalog information or price information. Very few people will think it is a multimedia link. This barcode property may reduce readers' interests to interact with documents.



(a) (b)

Fig. 2. EMBL Examples. Excerpt from Liu et al. [38]

To compensate for these issues, we designed a media awareness-mark, called Embedded Media Barcode Link (EMBL) [38]. Figure2 shows EMBLs printed on paper documents, where the EMBL iconic marks indicate linked video and audio respectively. EMBL is a semi-transparent media-icon-modified barcode overlay on paper document content for linking to associated media. It uses an "EMBL-signified document location" to define the precise location for media association. An EMBL uses semi-transparent form to reduce interference with original document content and get closer to an EMBL signified location. It uses a semitransparent barcode to identify signified document patches, and uses iconic information to reveal associated-media information to a user. EMBL's benefits can hardly be achieved manually. To facilitate EMBL creation, we designed an EMBL authoring tool to arrange EMBL based on barcode blending coefficient optimization in a neighborhood.

3.2 Micro Optical Patterns

To reduce the intrusiveness of barcodes, researchers invented micro optical patterns such as Dataglyph [29], and Anoto dot pattern [42]. Since the encoding mechanisms of these micro optical patterns are very similar to barcodes, they are sometime considered as barcode variations. Because of similar encoding mechanisms, the identification accuracy of these micro patterns are considered comparable to traditional barcodes. Different from traditional barcodes, these micro optical patterns are much smaller in size. According to literature, each Dataglyph pattern consists of a 45-degree diagonal line as short as one over one-hundredth of an inch or less. On the other hand, the Anoto dot pattern divides paper into a grid with a spacing of about 0.3mm. These significant size reductions make these micro patterns much less intrusive than traditional barcodes. The size reduction also makes it possible to print code densely for better "virtual mouse cursor" localization. Additionally, they can encode much more data than traditional barcode on the same size document patch.

There are mainly four disadvantages of using these micro optical patterns technologies for the PaperUI design. First, these patterns need to be printed by high resolution printers. To avoid intrusiveness to printed contents, micro pattern designers have to make the basic pattern units smaller than basic content units (e.g. a stroke). This design strategy demands much better printer resolution. According to literatures, the Anoto pattern printing process needs at least 600-dpi resolution (some claim a required resolution of 1000-dpi). Since the Dataglyph line segments are very small, it also needs high resolution printing. Second, these high resolution patterns require document scanners or very high resolution cameras for image capture. This requirement makes it difficult to use existing camera phones for Dataglyph or Anoto pattern capture. To achieve this requirement, some companies make specific equipment, such as Anoto pens for PaperUI interaction. Third, printing a large amount of dots or dataglyph on paper will make the paper background look grey and reduce the content image contrast. Fourth, this technique also requires printing procedure change that may become a barrier for using the technique in existing printing industry.

3.3 Encode Hidden Information

There are many different approaches to encode hidden information in documents. These approaches are frequently considered as watermarking techniques. Documents include many objects such as figures, lines, words, paragraphs etc. People may change these objects position, size, and contour etc. to encode information in documents [5]. For example, people can shift line upwards or downwards by very small amount to encode information. They may also shift words horizontally to modify the spaces between words for information hiding [3]. There is also a large number of image water marking techniques. Even though these techniques are mainly discussed in the watermarking research field, they do have potential to be used in the PaperUI implementation. One advantage of using these techniques in PaperUI is that most of them are much less intrusive than barcode.

There are also several disadvantages of using these techniques. First, data hiding techniques frequently use content specific knowledge in algorithm design. That makes them less adaptive to a big variety of document contents. For example, it is difficult to use the line shift technology for figures or images in a document. Second, when the host signal for image watermarking is not known, crosstalk between the watermark signal and host signal is a common problem [5]. To suppress crosstalk, many algorithms require original image available for hidden information extraction. This requirement is conflict with the PaperUI procedure, which needs to identify a document patch before getting the digital version of the document patch. Third, these techniques require document providers to change the hardcopy printing process, and this change cannot be separated from content printing.

3.4 Paper Fingerprint

Paper is composed of fine fibers entangled with each other. These entangled patterns are very durable and have very low probability to be identical. Therefore, these patterns can be considered as paper fingerprints [4]. Paper patch identification can be achieved by comparing captured fiber-pattern images with indexed fiber-pattern images. This technique has nearly no disturbance to printed contents. However, since wood fibers are normally much smaller than basic units of micro-optical-patterns, this approach demands a special camera or very high resolution camera for the patch identification task. Moreover, because the fiber pattern does not follow human made rule as the micro-optical-pattern does, performing fiber pattern matching is expected to be slower than micro-optical-pattern decoding.

3.5 Character/Word Recognition

Characters and words are much easier to capture than entangled wood fibers. It is also very intuitive to think about using characters and words for text document patch identification. Commercial optical character recognition (OCR) software is widely used to convert printed books and documents into text for web publication, text-to-speech, text-mining etc. According to [6], most OCR software claims 99% accuracy rates on new good quality clean images. With 45 pages from a digitized newspaper collection 1803-1954, the author found that raw OCR accuracy varied from 71% to 98.02%. These results are collected with scanned documents as inputs. The reported accuracies will decrease further when phone camera instead of scanner is used.

There are several advantages of using existing OCR software for PaperUI implementation. First, because OCR software is considered as an existing module, application development may become much easier than using other self-developed approaches. Second, this approach does not require printing process change. Third, it is not intrusive to original document.

There are also several disadvantages of using the existing OCR software. First, OCR software is language dependent most of the time. If language independent task is demanded, all language models have to be installed on the recognition machine. Second, OCR software normally requires high resolution cameras for image capture. This requirement is only supported by very limited cameras. Third, most OCR software cannot handle angled document capture, or low lighting capture. Fourth, OCR software cannot work on photos or figures that frequently appear in documents. Fifth, most OCR software still cannot work on characters printed with a complex layout.

To overcome some OCR software limitations, researchers proposed methods for layout free and language independent character recognition [21]. We believe research in this direction is a promising direction for document patch identification. On the other hand, we are still not sure if it is possible to improve this method and make it working on figures and natural images that exist in many documents. If document patch identification algorithm cannot work on figures and natural images, we will face difficulties to achieve our final goal - using cell phone cross hair as a mouse cursor over any portion of a document.

3.6 Local Image Features

Using local image features is another promising approach for document patch identification. Recently, researchers invented many different features for cell phone and paper interaction. HotPaper [9] and Mobile Retriever [10] use features based on document text such as the spatial layout of words. Other systems such Bookmarkr [11], MapSnapper [12], EMM [32] use pixel level image features, such as the SIFT [13] and FIT [33] algorithms, to recognize generic document content such as pictures and graphic elements.

Using image local features have many advantages. First, these systems do not require exclusive spaces on paper for marker printing. Second, they do not change document printing procedure. Third, most of these features can work on camera phone captured images. Fourth, some of these algorithms allow us to accurately locate a cellphone crosshair corresponding location on paper so that we can use the cellphone crosshair as a mouse cursor for human-document interaction.

Because HotPaper and Mobile Retriever use word level layout features, they are limited to western text regions in a document. Method described in [31] can work on Japanese characters or other eastern Asia characters via parameter adjustment. However, it is still limited to text regions in a document. FIT and SIFT are tested on document mixtures including, western text regions, eastern Asia characters, figures, and natural images with reasonable recognition results [30] (99+% recognition rate). Encouraged by those test results, several recent PaperUI systems are developed based on these features. On the other hand, because FIT and SIFT work on pixel level, their constructions are normally slower than word level features. This is an issue we should pay more attention for future PaperUI research. Beyond recognition speed, another issue of this local feature based approach is the interaction indication. More specifically, because this patch identification approach does not put any marker on paper, it is difficult for a user to figure out where and how to interact with document contents. To overcome this issue, some products use dedicated text paragraphs to explain the interaction. Researchers also try to add some less intrusive markers to facilitate user's interactions with paper [32].

3.7 RFID-Based Document Recognition

Recent advances in RFID technology make RFID chips small enough to be embedded in sheets of paper [14]. These advances reveal potential of using RFID technology for paper patch identification. RFID technology may allow users to interact with paper at a distance. It also has fast response speed. However, we still need to overcome some issues before we can finally use this technology for the PaperUI development. First, special printer need to be developed for accurately "printing" these RFID devices on paper. Second, we need to develop portable RFID identification devices that can be easily carried by users. Third, we also need to develop technology that can avoid RFID interferences from other pages. Fourth, technology is needed to allow RFID devices to identify the user selected RFID from proximate RFIDs on the same page. Fifth, to facilitate interactions with the whole paper surface, it is better to develop technologies that can accurately estimate the RFID device's pointing direction based on nearby RFIDs.

4 PaperUI Applications

Based on emerging technologies, many research demos have been developed in the PaperUI direction. In this section, we will present applications based on their underlying emerging technologies and interaction resolution.

4.1 Digital Pen Based Applications

After Anoto™ commercialized its digital paper technology, companies such as Logitech™, Maxell™, Nokia™, Leapfrog™, and Livescribe Connect™ developed digital pen hardware for feeding users' paper inputs to computers. The technology advances on digital pen enabled many interesting applications. Early applications in this direction include medical and bank form filling.

The Leapfrog FLY Fusion Pentop Computer™ [48] is sold in many supermarkets. This device can read things a user writes on Fly Paper and automatically upload the user's notes to a computer. With this UI system, a user can convert some text in notes to document and back up the document; it can also perform basic math tutorials, translations, spell check, games, and Trivia. We believe this kind of computer form is much easier for kids to master than laptops. Compared with desktops and laptops, this pen computer also has less impact to kids' traditional activity.

Recently Livescribe Connect™'s smart pen enables its users to record audio while taking lecture notes and retrieve audio with pen and notes. It also connects its user to Email, Google Docs, Facebook, and Evernote®. Additionally, it enables users to play music with pen and paper. Compared with heavy and bulky laptops, this light-weight interface is more convenient for students to take notes in classes [45].

For field workers, Capturx™ provides pen-paper interfaces for Excel forms, Microsoft® Office OneNote, ArcGIS, and PDF. Moreover, its disaster response kit is designed to help emergence response teams map, collect, and share data from the scene of wildfires, hurricanes, and floods etc [49]. In a different application scenario, Infomax™ provides a digital pen and paper solution for environmental compliance technicians to fill out various complicated inspection forms, such as restaurant inspection forms, methane producing well maintenance forms, traffic violation forms, and sales forms [50]. Compared with traditional laptop interfaces, this pen interface is much easier to be adopted by field workers.

In research fields, researchers also invented many novel applications with the digital pen technology. Yeh and Liao's ButterflyNet [46] enables biologists to integrate paper notes with information explicitly captured in field sites: digital photographs, sensor data, and GPS etc. With this system, biologist can directly transfer their collected contents to spreadsheet, browse all synchronously created media, and share their work with other colleagues.

Song et al. developed PenLight [52] and MouseLight [51] systems that can visually augment paper documents and give the user immediate access to additional information and computation tools. More specifically, PenLight allows an architect to use a projector and a digital pen on or above paper surface to retrieve additional information and project the information on the printout. It supports copy-paste, section view, 3D model navigation, related computation, and coordination of different versions and different collaborators.

To support meetings in conference rooms, researcher developed a system that allows meeting participants to control services in a meeting environment through a digital pen and an environment photo on digital paper [47]. Figure 3 shows the paper interface and deployment environment of this system. Unlike state-of-the-art device control interfaces that require interaction with text commands, buttons, or other artificial symbols, the photo-on-paper enabled service access is more intuitive. Compared with PC and PDA supported control, this approach is more flexible and cheap. With this system, a meeting participant can initiate a whiteboard on a selected public display by tapping the display image in the photo, or print out a display by drawing a line from the display image to a printer image in the photo. The user can also control video or other active applications on a display by drawing a link between a printed controller and the image of the display. Beyond meeting room device control, Liao et al. also developed a PaperCP system that can facilitate student-instructor communication for active learning. With the PaperCP system, users can enjoy the inherent advantages of paper. Moreover, students can electronically submit their handwritten notes to the instructor, thereby maintaining two-way communication with the instructor [53].

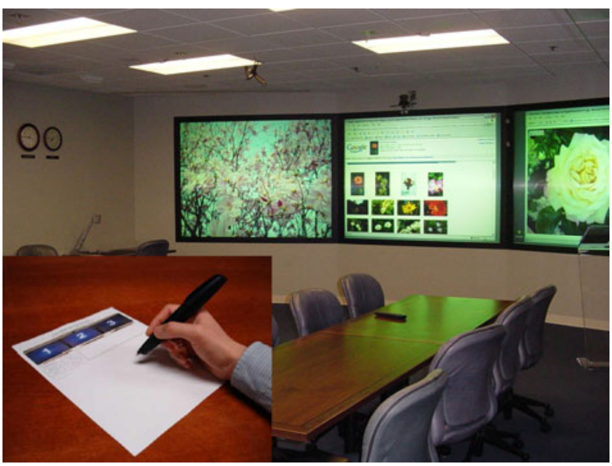


Fig. 3. The deployment environment and interface of POEMS (Paper Offered Environment Management Service) for meeting room control. Excerpt from Hu et al. [47].

4.2 Barcode Based Applications

Because of paper's good physical property, Barcode have been used on paper surface for a long time. Due to reading device limitations, the use of barcode was only limited to business operators in the past. Recently, many barcode readers were released for the fast growing cameraphone market. This trend makes it much easier for general public to use barcode. For example, consumers can use cameraphone captured barcode to compare product price [15], read reviews, acquire coupons, shopping, input a business card, navigate a city guide or map [16], get athletes' videos, pictures and fan data from a poster [17], get updates of a news story, get weather information from a map location [18], or read additional contents linked to an IEEE article. Ricoh’s iCandy [27] also allows kids to select a movie on TV based on a barcode capture. These systems are much more convenient than desktop or laptop systems that require manual input of product information or web addresses. They also overcome typing difficulties encountered by many cell phone users.

4.3 RFID Based Applications

Even though RFID has been invented for a long time, its application for PaperUI is still new. Derek et al. developed Marked-up Maps by setting a RFID grid under a paper map. With this RFID underlying grid, a user can wave a handheld computer equipped with an RFID reader above the region of interest on a paper map and get digital information. In their example, they assume a tourist can use their system to get extra information from a Marked-up Map of the Montreal subway, Nottingham, and Greater Vancouver. This extra information may be restaurant and hotel information near a subway station, shopping centers linked to store directories, theatres to current and upcoming shows, historical sights, or spatially accurate transit, district, and landmark data for navigation. One issue of that system is that a tag placed beneath a landmark that sits in relative isolation compared to other tagged landmarks will have a greater read range than landmarks that are more closely spaced [19]. This inconsistency issue makes it a little difficult for beginners to interact with a map.

4.4 Character/Word Recognition Based Applications

Camera phone based text recognition also has some interesting applications for PaperUI. For example, ScanR® and rivals can convert a camera phone image of a hardcopy into PDF for search, editing, email, text to audio translation etc. Abbyy® [7] and Google Goggles allow users to take photographs of business cards, translate the card into digital information and add that information into address book. With proper OCR software, cameraphone can also be used as a translator for foreign restaurant menus, posters etc. [20, 21]. Google Goggles’ recent release already can translate text restaurant menus from a language to other languages [23].

4.5 Encoding Hidden Information Based Applications

DigiMarcTM and its rivals try to embed digital information in images and use the embedded information to initiate many different applications. These applications include bring videos, interaction widgets, and other multimedia information to a cameraphone via capturing a DigiMarc® encoded paper page. It also claims finding product coupons, comparing product price, and finding product stores via capturing a product package [22]. In theory, DigiMarc® has the same function as barcode and therefore can be used for nearly all applications that a barcode can do. Better than a traditional barcode, DigiMarc® has much less impact to original pictures. This feature makes a DigiMarc® encoded page look nicer than a barcode attached page. On the other hand, because the encoded data has less contrast than barcode in the signal space, it will be less reliable than barcode. Moreover, since embedded data can be invisible to human. That makes it difficult for a user to find out the DigiMarc® encoded page. To solve this problem, DigiMarc® uses a small visible icon to remind users about the extra data existence.

4.6 Original Document Feature Based Applications

Recently, many companies, research labs and universities developed some interesting PaperUI applications based on original document features. For example, Google Goggles can do book search by capturing the book cover; it can recognize some artwork and bring back related information of the artwork; it can also perform product search based on wine marks and spencer [23]. Kooaba's Paperboy can provide interactive storytelling about print ads; it can navigate a consumer to a nearby store based on ads capture; it also allows food makers to provide product ingredients and origins to users via the phone capture [24]. Ricoh's technology allows people to get an updated guidebook via capture an old guidebook; it can automatically associate a http link to its surrounding text arrangement features so that users can be directed to the link by capturing the surrounding features of the http link [25]; it also supports a user to voice-annotate a real-estate brochure[26]. Amazon SnapTell can use the cameraphone photo of any CD, DVD, book, or video game to retrieve the product and find ratings and pricing information online [28]. Mobile Retriever [10] suggests using document identification technology to help visually impaired persons [10]. MapSnapper enables users to query a remote information system based on photos of a paper map taken with a mobile device [12]. Bookmarkr allows users to share photos with friends by taking an image of a photo in a photobook with the mobile phone's camera [11]. Rohs [18] augments pre-defined regions in a printed map with dynamic weather information.

Fig. 4. (left) An EMM in a cartridge installation manual (right) The associated step-by-step video tutorial. Excerpt from Liu et al. [32].

When content-based feature are used, there is no on-paper indication at all to the user that there is media linked to the document. As a result, a HotPaper [26] user has to pan a camera phone over the paper document to look for hotspots until feedback such as a red dot or vibration is presented on the cell phone. Moreover, there is no media type indication either. Additionally, because a user does not know where to capture and how to capture, digital links may be missed or large amount of resources has to be used to index all possible captures. This is awkward when digital links are sparsely distributed through many pages.

To solve this problem, Liu et al. augment paper with meaningful awareness-marks, called Embedded Media Markers (EMMs) that indicate the existence, type, and capture guidance of media links. Figure 4 shows an EMM and its application scenario. On seeing an EMM, the user knows where to capture an image of the EMM-signified document patch with a cell phone in order to view associated digital media. This is analogous to Web pages that use underlines, font differences, or image tags to indicate the existence of links that users then click for additional information. Unlike barcodes, EMMs are nearly transparent and thus do not interfere with the document appearance. Unlike Embedded Data Glyphs [29] or Anoto patterns [42], EMMs can be printed with a regular low-resolution printer and identified from an image captured by a normal cell phone camera. Unlike other appearance-based approaches, EMMs clearly indicate signified document patches and locations. The design of EMMs also indicates what type of media (e.g. audio, video, or image) is associated with the EMM-signified document location. Furthermore, by requiring the captured image to cover the whole mark, we can improve feature construction accuracy, matching accuracy, and efficient resource usage [32]. Currently, the EMM system supports links to 5 types of marks including audio, video, web, image, and text.

4.7 Fine-Grained Phone-Paper Interactions

Ideally, we want to use cellphone or pen as a mouse or even better device in a PaperUI system. Applications in the previous session focus on creating digital links to a large paper patch. Operations in these applications are similar to very rough point-and-click mouse-operations. There is still a big gap between these coarse operations and fine-grained mouse operations. For example, these applications do not use gestures such as the marquee, bracket and lasso selectors, which offer more flexibility to manipulate document content and have been widely deployed in GUIs [34]. They also lack resolution for selecting a small region, such as a math symbol region, in a set of adjacent regions.

To overcome these problems, Liao et al. developed PACER [34] that features a camera-touch hybrid interface. Figure 5 illustrates a PACER application scenario. The system recognizes documents based on natural document visual features instead of any special markers on paper or specific end user hardware. More importantly, it allows users to manipulate fine-grained document content with various gestures beyond point-and-click. With this system, a user first aims a camera phone roughly at the region and captures a picture. PACER recognizes the picture, and presents on the screen the corresponding high quality digital version of the document, rather than the raw video frames (We call this design loose registration). The user then operates the phone as an embodied see-through "Magic Lens" [35], with its crosshair center treated like a mouse pointer. To fine tune the starting/ending point of the gesture, the user can also switch from the embodied interaction to the touch interaction by directly touching the screen and moving the pointer in a zoomed and scrollable view.

With this camera-touch hybrid interaction, a PACER user can select, copy and email an interesting region from a paper document; pick the title of a reference from a journal, and then search for its full text on Google Scholar; specify a word or math symbol for text search on paper; check dictionary of a foreign words in the document; snap to a sightseeing drive on a paper map, and browse the street views on the phone while sweeping it along the route; play a music score by moving the phone over the intended sections; using voice or multimedia data to annotate a very small region in the document; perform free-form gestures for document annotation; or discuss a document with a remote user via pointing, drawing, copy-paste, and speaking. Through enabling fine-grained gesture to the paper-cameraphone system, PACER greatly extended the possible application scenarios of the PaperUI concept.

Different from the PACER usage scenario, a user may also want a portable system support for reading with papers on a desk. This kind of scenario is less mobile than the PACER scenario. However, reading with the paper on a desk and a light pen in hand is a more relaxed setup for long time reading. FACT [36] is designed for this need. Figure 6 illustrate the FACT system and one application scenario. FACT consists of a small camera-projector unit, a laptop, and ordinary paper documents. With the camera-projector unit pointing to a paper document, the system allows a user to issue pen gestures on the paper document for selecting fine-grained content and applying various digital functions a traditional computer can issue. FACT can also support various applications in information transfer, association, sharing and synchronous navigation across the paper-computer boundary. For example, a FACT user can select a picture on paper and then copy the digital version of this picture into a Word document on the computer; this operation can also be reversed so that Multimedia annotations created on the laptop can be attached to a specific word, phrase, paragraph or arbitrary shaped region for pen-based retrieval. FACT may also be used as a platform for sharing paper/web information with remote friends. For example, a paper user can select a map region and ask a remote friend for tour suggestions, and the remote friend can attach interesting web contents to the document and project back on the real paper map.

The PaperUI concept is not proposed to completely replace existing computer interfaces. Actually, the PaperUI interface can be used with existing computer interfaces to take both advantages. The recently developed MixPad [37] allows a user drag a picture on a paper page to a nearby laptop with a finger on paper or type text via the laptop keyboard to annotate an illustration in a printout. Figure 7 illustrates the MixPad idea.

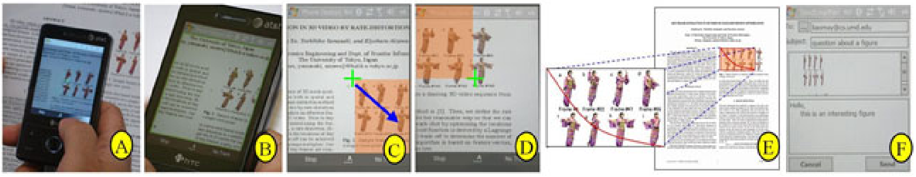


Fig. 5. Copy&Email via PACER. (A) Point the phone crosshair to an interesting area on paper and take a snapshot. (B) Once the snapshot is recognized, the corresponding high-quality version is displayed. (C)-(D) Move the phone (in the arrow direction) over the paper to select a region (highlighted in orange) with a marquee gesture. (E) Overview of the gesture/region within the document. (F) The selected region is sent via email, together with the hyperlink pointing to the original page and document. Excerpt from Liao et al.[34].

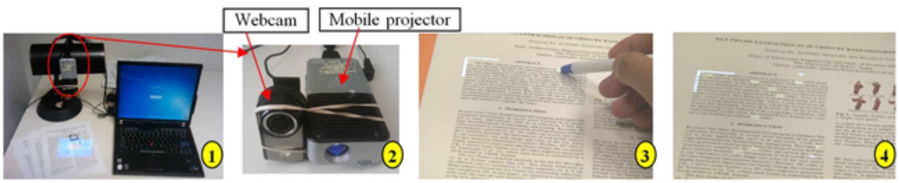


Fig. 6. (1) Interface prototype, (2) Close-up of the camera-projector unit, (3) A word (highlighted by the projector) selected by a pen tip for full-text search, (4) The resulting occurrences of the word highlighted by the projector. Excerpt from Liao et al.[36]

Fig. 7. (left) MixPad interface prototype, (middle) Close-up of the camera-projector unit, and (right) steps to select a picture portion: (1) roughly pointing a finger to a region, (2) mouse cursor being projected to where the fingertip is, and (3) Drawing a mouse marquee gesture to select a region at fine granularity. Excerpt from Liao et al.[37].

5 Concluding Remarks

Emerging technologies open a door for us to explore the PaperUI concept and practices. With many research prototypes and products developed in this field, we believe the prime time for this new interface is coming. Through migrating existing PC applications to the more portable PaperUI interface and developing new applications in this field, we strongly believe that we can find more promising research topics on novel document patch identification methods, mobile device pointing direction estimation, and user interactions etc.

References

1. Wikipedia, Barcode, http://en.wikipedia.org/wiki/Barcode
2. IDAUTOMATION, About Barcode Accuracy and Misreads, http://www.barcodefaq.com/barcode-encoding.html
3. Maxemchuk, N.F., Low, S.: Marking text documents. In: Proc. IEEE Int. Conf. Image Processing 1997 (ICIP 1997), Santa Barbara, CA, vol. 3, pp. 13–16 (October 1997)
4. Fuji Xerox, Paper Fingerprint Recognition Technology (August 3, 2011), http://www.fujixerox.com/eng/company/technology/xaya/
5. Hartung, F., Kutter, M.: Multimedia watermarking techniques. Proceedings of the IEEE 87, 1079–1107 (1999)
6. Holley, R.: How Good Can It Get? D-Lib Magazine (March/April 2009), http://www.dlib.org/dlib/march09/holley/03holley.html
7. Jacobi, J.L.: ABBYY FineReader 10 Professional Edition review (August 3, 2011), http://www.pcadvisor.co.uk/reviews/software/3258293/abbyy-finereader-10-professional-edition-review/
8. Fox, M.: Alan Haberman, Who Ushered In the Bar Code, Dies at 81. The New York Times (June 15, 2011)
9. Resource Conservation Alliance, Focus On Paper Consumption, http://www.woodconsumption.org/products/paper.pdf
10. Liu, X., Doermann, D.: Mobile Retriever: access to digital documents from their physical source. Int. J. Doc. Anal. Recognit. 11(1), 19–27 (2008)
11. Henze, N., Boll, S.: Snap and share your photobooks. In: Proceedings of ACM Multimedia 2008, pp. 409–418 (2008)
12. Hare, J., Lewis, P., Gordon, L., Hart, G.: MapSnapper: Engineering an Efficient Algorithm for Matching Images of Maps from Mobile Phones. In: Proceedings of Multimedia Content Access: Algorithms and Systems II (2008)
13. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)
14. FujiSankel, Hitachi Develops RFID Powder, http://healthfreedoms.org/2009/10/23/hitachi-develops-rfid-powder/
15. Asaravala, A.: Camera Phones Help Buyers Beware, Tech Biz: IT, http://www.wired.com/techbiz/it/news/2004/01/61936
16. Hanlon, M.: ScanBuy - barcode software on your camera phone creates the Physical World Hyperlink, Mobile Technology, http://www.gizmag.com/go/6701/
17. Byron Acohido, New '2D barcodes' puts info at the tip of your camera phone, USA TODAY (May 20, 2009)
18. Rohs, M.: Real-World Interaction with Camera Phones. In: Murakami, H., Nakashima, H., Tokuda, H., Yasumura, M. (eds.) UCS 2004. LNCS, vol. 3598, pp. 74–89. Springer, Heidelberg (2005)
19. Reilly, D., et al.: Marked-Up Maps: Combining Paper Maps and Electronic Information Resources. Personal Ubiquitous Computing 10(4), 215–226 (2006)
20. Ellison, J.: How To Make The Cell Phone Into A Portable Scanner. News Week Magazine (October 20, 2007)
21. Iwamura, M., Tsuji, T., Kise, K.: Memory-based recognition of camera-captured characters. In: Proceedings of the 9th IAPR International Workshop on Document Analysis Systems (DAS 2010), pp. 89–96. ACM, New York (2010)

22. DigiMarc, Enhanced, Interactive Experiences from Media Using Smartphones (August 3, 2011), http://www.digimarc.com/
23. Google Goggles, Google Goggles in action (August 3, 2011), `http://www.google.com/mobile/goggles/#label,`
24. Kooaba, Paperboy: Deliver digital extras for print (August 3, 2011), `http://www.kooaba.com/`
25. Hull, J.J., Erol, B., Graham, J., Ke, Q., Kishi, H., Moraleda, J., Van Olst, D.G.: Paper-based augmented reality. In: Proc. 17th Int. Conf. Artificial Reality and Telexistence
26. Erol, B., Antunez, E., Hull, J.J.: HOTPAPER: multimedia interaction with paper using mobile phones. In: Proceedings of ACM Multimedia 2008, pp. 399–408 (2008)
27. Graham, J., Hull, J.: Icandy: A tangible user interface for itunes. In: Proc. CHI 2008: Extended Abstracts on Human Factors in Computing Systems, Florence, Italy (2008)
28. Toto, S.: SnapTell: Instant Product Lookup From The iPhone. You Want This. Tech Crunch (November 19, 2008), `http://techcrunch.com/2008/11/19/snaptell-instant-product-lookup-from-the-iphone-you-want-this/`
29. Hecht, D.L.: Embedded Data Glyph Technology for Hardcopy Digital Documents. In: SPIE -Color Hard Copy and Graphics Arts III, vol. 2171, pp. 341–352 (February 1994)
30. Nakai, T., Kise, K., Iwamura, M.: Camera-Based Document Image Retrieval as Voting for Partial Signatures of Projective Invariants. In: Proc. 8th International Conference on Document Analysis and Recognition (ICDAR 2005), pp. 379–383 (September 2005)
31. Nakai, T., Kise, K., Iwamura, M.: Real-Time Retrieval for Images of Documents in Various Languages using a Web Camera. In: Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR 2009), pp. 146–150 (July 2009)
32. Liu, Q., Liao, C., Wilcox, L., Dunnigan, A., Liew, B.: Embedded Media Markers: Marks on Paper that Signify Associated Media. In: Proceedings of ACM IUI 2010, pp. 149–158 (2010)
33. Liu, Q., Yano, H., Kimber, D., Liao, C., Wilcox, L.: High Accuracy and Language Independent Document Retrieval With A Fast Invariant Transform. In: Proceedings of IEEE ICME 2009, pp. 386–389 (2009)
34. Liao, C., Liu, Q., Liew, B., Wilcox, L.: Pacer: fine-grained interactive paper via camera-touch hybrid gestures on a cell phone. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI 2010). ACM, New York (2010)
35. Rohs, M., Oulasvirta, A.: Target Acquisition with Camera Phones when userd as Magic Lenses. In: Proceedings of CHI 2008, pp. 1409–1418 (2008)
36. Liao, C., Tang, H., Liu, Q., Chiu, P., Chen, F.: 2010. FACT: fine-grained cross-media interaction with documents via a portable hybrid paper-laptop interface. In: Proceedings of the International Conference on Multimedia (MM 2010), pp. 361–370. ACM, New York (2010)
37. Liao, C., Liu, Q.: MixPad: Augmenting Interactive Paper with Mice & Keyboards for Cross-media Interaction with Documents. To be published Proc. ACM UBICOMP 2011, Beijing, China (2011)
38. Liu, Q., Liao, C., Wilcox, L., Dunnigan, A.: Embedded media barcode links: optimally blended barcode overlay on paper for linking to associated media. In: International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI 2010). ACM, New York (2010)
39. The Office of the Future, Business Week (2387): 48–70 (June 30, 1975)

40. Lahey, B., Girouard, A., Burleson, W., Vertegaal, R.: Paperphone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays. In: Proceedings of CHI 2011. ACM Press (2011)
41. Wellner, P.: Interacting with paper on the DigitalDesk. Communications of the ACM 36(7), 87–96 (1993)
42. Anoto, THE DIGITAL PEN, (August 3, 2011), http://www.anoto.com/?id=19146
43. Mackay, W.E., Fayard, A.-L.: Designing Interactive Paper: Lessons from three Augmented Reality Projects. In: Proc. IWAR 1998, pp. 81–90 (1998)
44. easeMD, Digital Pen for Off-site Use by Physicians (August 3, 2011), http://www.easemd.com/doctors/digital-pen.php
45. Livescribe, Record and Play Back (August 3, 2011), http://www.livescribe.com/en-us/smartpen/
46. Yeh, R.B., Liao, C., Klemmer, S., Guimbretière, F., Lee, B., Kakaradov, B., Stamberger, J., Paepcke, A.: ButterflyNet: A Mobile Capture and Access System for Field Biology Research. In: Proceedings of CHI 2006, pp. 571–580 (2006)
47. Hu, C., Liu, Q., Liu, X., Liao, C., McEvoy, P.: Poems: A Paper Based Meeting Service Management Tool. In: ICME 2007, February 7 (2007)
48. Leapfrog, Pentop Computer User Manual, http://www.leapfrog.com/etc/medialib/leapfrog/fly/ffls_37735.Par.40053.File.dat//Users/na/Desktop/Parent%20Guides/FFLS_37735_PG.pdf
49. Capturx, Automate Any Paperwork, http://www.adapx.com/
50. Infomax, Digital Writing Slutions Made Simple, http://www.e-infomax.com/case.html
51. Song, H., Guimbretiere, F., Grossman, T., Fitzmaurice, G.: MouseLight: bimanual interactions on digital paper using a pen and a spatially-aware mobile projector. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI 2010). ACM, New York (2010)
52. Song, H., Grossman, T., Fitzmaurice, G., Guimbretiere, F., Khan, A., Attar, R., Kurtenbach, G.: PenLight: combining a mobile projector and a digital pen for dynamic visual overlay. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI 2009), pp. 143–152. ACM, New York (2009)
53. Liao, C., Guimbretière, F., Anderson, R., Linnell, N., Prince, C., Razmov, V.: PaperCP: Exploring the Integration of Physical and Digital Affordances for Active Learning. In: Baranauskas, C., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007, Part II. LNCS, vol. 4663, pp. 15–28. Springer, Heidelberg (2007)