Supporting Active Reading on Pen and Touch-Operated Tabletops

Fabrice Matulic, Moira C. Norrie
ETH Zurich, Institute of Information Systems
8092 Zurich, Switzerland
{matulic|norrie}@inf.ethz.ch

ABSTRACT

With the proliferation and sophistication of digital reading devices, new means to support the task of active reading (AR) have emerged. In this paper, we investigate the use of pen-and-touch-operated tabletops for performing essential processes of AR such as annotating, smooth navigation and rapid searching. We present an application to support these processes and then report on a user study designed to compare the suitability of our setup for three typical tasks against the use of paper media and Adobe Acrobat on a regular desktop PC. From this evaluation, we found out that pen and touch tabletops can successfully combine the advantages of paper and digital devices without their disadvantages. We however also learn from observations and participant feedback that there are still a number of hardware and software limitations that impede the user experience and hence need to be addressed in future systems.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - Input devices and strategies (e.g., mouse, touchscreen)

General Terms

Design, Human Factors

Keywords

Active reading, bimanual pen and touch interaction, digital tabletops

1. INTRODUCTION

Active reading (AR) is the act of interactively perusing a document by employing effective strategies to better comprehend and remember its content. When performed on paper, this usually involves annotating, underlining or highlighting elements with pens or markers and searching through the document by flipping through pages [17]. With technological advances and the need to digitise information for more efficient integration and recollection, researchers and engineers have attempted to support those processes in various ways, taking care, when possible, to preserve the naturalness and intuitiveness with which people engage with physical documents by more or less mimicking the properties of paper interaction. Thus, researchers came to consider a range of media and devices for prototype development, such as pen-operated tablets [15, 20], multitouch devices [23] and augmented paper [16]. This shift was paired with great efforts on the software front as well to provide more effective interaction techniques and integration of the written content with digital workflows [12].

In this paper we explore a new promising environment for AR that is digital tabletops with simultaneous bimanual pen and (multi) touch input capabilities [4, 10]. We envisage this medium as a digital office desk on which knowledge workers perform a wide range of document tasks, one of which is AR. We view this platform as an additional enabler in the landscape and not a replacement of tablets and other media, where each has its own advantages and disadvantages. We believe, for instance, that the additional work space provided by tabletops allows for increased efficiency and comfort compared to tablets, PDAs, smartphones etc., which have to make sacrifices for the sake of mobility. Moreover, the vast majority of current direct input devices are either limited to one interaction type (mostly touch) or allow only alternative use of stylus and touch input without differentiating between the two. In other words, none of them implement asymmetric bimanual input of commands and gestures with a conscious design choice of making a synergistic use of both modalities.

In the office context we set ourselves, we investigate how a pen and touch tabletop system compares to media commonly used in the workplace, i.e. regular paper and desktop computers. This choice of experimental conditions favouring typical settings over best possible competitors follows O'Hara and Sellen's rationale that selecting optimal configurations for a task is a subjective decision [19]. Furthermore, testing a system against conventional tools ensures that users can react and provide feedback based on familiar work environments. This approach yields valuable insights into the benefits and weaknesses of the considered platform that can subsequently help inform the design of future systems.

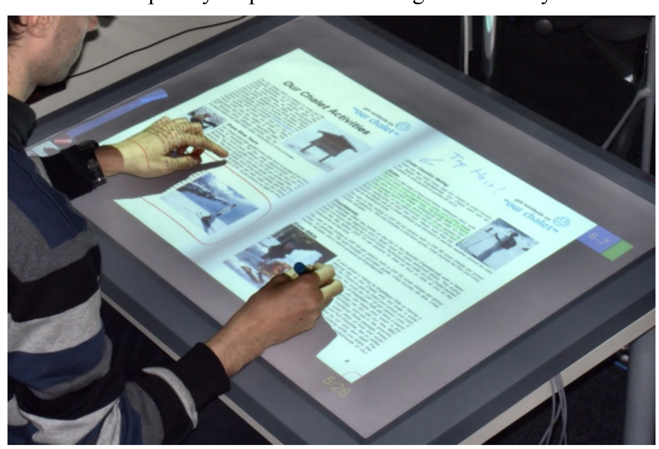

Figure 1. The main user interface of our system for a right-handed user (sidebar on the left)

Our contributions in this paper are threefold: we present a complete system designed for pen and touch tabletops to support active-reading tasks with several navigation functions and in-

AVI '12, May 21-25, 2012, Capri Island, Italy.

document searching capabilities. Our application includes a set of traditional, adapted and novel functions and gestures, which are packaged in a coherent whole to serve the execution of the task at hand. We then show the results of a comprehensive evaluation that not only validates our hypothesis that our system is able to effectively combine the advantages of paper and a typical digital platform such as Adobe Acrobat without their disadvantages, but also details what types of gestures seem to work well for such document tasks and which do not. Finally, from our experiments and user feedback we derive specific requirements detailing important hardware and software issues that need to be addressed for pen and touch surfaces to be able to become more suitable platforms for productive document tasks.

2. RELATED WORK

2.1 Active Reading

There is a considerable body of work tackling active reading and its challenges in the literature. In their seminal study, O'Hara and Sellen compared people's reading experience on paper vs. on a computer screen [19]. They observed that affordances of paper were compelling and derived design implications for digital systems to better support the task, in particular the need to provide adequate tools for annotations, quick navigation and control over spatial layout of content. 10 years later, Morris et al. conducted a study based on similar experimental conditions, in which they highlighted the progresses made by computers [17]. They reported, for instance, that tablets were able to perform on par with paper in their AR tasks. A more recent investigation by Tashman and Edwards reinforces the impression of computers' growing dominance over paper for AR [22]. The authors also notice that as digital tools have become more pervasive, people's habits have also changed so that AR tasks are now more readily executed on a digital device rather than on paper.

There have also been a number of actual systems developed to specifically support AR. Among the more prominent examples, one can cite Xlibris, a digital tablet with a paper-like interface [20] and PapierCraft [16], a gesture-based command system to manipulate digital documents via an optical pen and interactive paper. More recently, LiquidText shows an example of a multi-touch AR system that breaks with the paper metaphor for the document layout, yet to a large extent relies on a set of gestures such as collapsing and selection that are inspired by paper interaction [23]. The authors also reveal that they would have opted for a pen and touch interface had they managed to find adequate hardware.

2.2 Pen and Touch

Recent years have seen efforts to provide simultaneous pen and tactile interaction modalities simultaneously to tabletop systems that are typically limited to (multi-)touch input. Those efforts have consisted mainly in experimental setups combining separate touch-sensitive hardware with technology to capture pen data [4, 14, 25] although some vision-based systems have been able to use the visual signature of the pen (e.g. a recognisable IR LED on its tip [8, 10]) to differentiate its contact point on the table from fingertip touches.

The motivation to build such experimental systems stems from the benefits that simultaneous pen and touch input can bring compared to strictly unimodal systems [4]. Hinckley et al., for instance, underline that pen and touch not only have unique characteristics individually, but that the combination of the two modalities creates even richer interaction possibilities in terms of gestures that they separately cannot provide [11]. The advantages of mixed pen and touch input have been exemplified in a few applications such as Hands-On Math [26] and a diagram editor [8].

3. DESCRIPTION OF THE SYSTEM

Considering our target environment and input paradigm, which attempt to remain close to people's natural experience with physical documents, we opted to follow the path of the paper metaphor, but with a number of enhancements that our platform permitted and that we detail hereafter. Our decision was further motivated by the fact that PDF documents, arguably the most common format for documents intended for reading, use fixed, page-based layouts and hence lend themselves well to a close-to-paper setting.

We consider the main processes of AR outlined by O'Hara and Sellen, i.e. annotating, navigating and laying out content [19], where in a first iteration we limit ourselves to single-document situations. Given those requirements, we set about to support the following functions in our application:

- Annotating: adding, deleting and recalling of annotations
- Navigation: Sequential and random movement in the document. The user has to be able to rapidly browse through the pages of the document and easily locate relevant information.
- Layout: overview of the whole document to gain a quick panorama of its content at a glance.

Within our design philosophy to remain in a familiar paper-like environment yet also provide efficient tools from the digital world, we were particularly interested in coming up with suitable gestures to provide searching capabilities. We see searching as one of the most significant ability that paper lacks and that computers bring to the table. Following this approach, we decided to include keyword-querying functionality where the pen plays a direct role, as opposed to resorting to a touch-only virtual keyboard. Thus we adopted a technique similar to that found in InkSeine [12], where the keywords are either selected in the document by the pen or directly handwritten by the user.

3.1 Implementation

The application was developed on a platform composed of a DiamondTouch table [7] and an overlay with an Anoto pattern printed on it to be used in conjunction with an optical pen [1]. This combination of technologies is a popular choice when building an experimental pen and touch tabletop system [4, 14]. The advantage is that the input resolution for the pen is very high. Anoto claims a theoretical precision of 0.03 mm for the pen position [13] and while the actual value is perhaps lower, the accuracy of the digital pen is still much greater than that afforded by current sensing technologies used in multitouch tabletops. Another convenience of using separate technologies for the two input modalities is that pen and touch handling are uncoupled, which allows for a clear separation of the two types of input events at the software level, facilitating development. The disadvantage of using Anoto technology in such a way is the slight optical distortions created by the dot pattern on the sheet of paper, which impacts the readability of the text projected on it. This and other impediments are expounded further below in the paper. Those inconveniences, however, do not detract from the overall smoothness and operativeness of the interface and the usability of the system.

Other than the SDKs provided by the vendors, the pen-management engine relies on a customised version of the iPaper framework [18] designed initially for the creation of interactive paper but adapted to support streaming data for an Anoto pen to be used as a digital stylus on a drawing surface. A commercial handwriting recognition package is also utilised to handle handwritten input of text.

3.2 User Interface

For the user interface, we took inspiration from 3Book [5], which adopts a 3D interactive codex book as its representational model, supported by a number of digital library features such as text search and highlighting, annotating and dynamic bookmarks. We revisited the concepts developed in this PC application and adapted them for our pen and touch tabletop setting.

The UI thus follows a virtual desktop paradigm, in which the working space of a physical desk is modelled and represented on the tabletop surface. Figure 1 shows the system, with an open document. A toolbar containing selectable colour pens as well as a number of icons which are used to execute commands is also available. The toolbar docks on the side of the interface that is opposite to the user's handedness, since it is meant to be operated using the hand not holding the pen, i.e. the non-dominant hand.

To support the aforementioned spatial layout aspect of AR in our single-document configuration, we included an overview mode similar to Space-Filling Thumbnails [6], i.e. where all pages of the document appear as smaller icons on a single screen, mimicking a situation where all pages of a paper document are scattered and laid out on a surface. We describe the overview mode and its associated triggering gesture further below.

3.3 Hand Roles

Since Guiard's Kinematic Chain model [9] describing the roles of each hand in asymmetric bimanual tasks involving joint motor actions by the dominant and non-dominant hands, many HCI researchers have sought to apply those principles to their work [4, 10]. In this framework, the non-dominant hand sets a general spatial and temporal frame of reference in which the dominant hand performs fine-precision operations. In pen and touch contexts, a broad interpretation of this division of labour can imply having the dominant hand use the pen for inking (fine-precision action) while the non-dominant hand is committed to positioning the document as well as control actions triggered by touch (coarser actions). This assignment of input modalities to specific and distinct roles is generally a sensible guideline for interface designers, but as noted by Hinckley et al. [11] a strict separation of the two hand functions would be overly restrictive and needlessly limit the richness of gestures achievable by combined actions of both hands. This relaxed stance was also adopted in the design of the gestures for our system, where the pen is mostly used for inking on the document, but sometimes also to issue commands (see following subsection).

3.4 Gesture Set

The implemented gestures are a combination of classic unimodal gestures to manipulate digital objects on an interactive surface (i.e. pan, zoom, rotate etc.) and a number of bimanual gestures exploiting the two modalities separately or simultaneously. As stated above, the dominant hand holding the pen is mostly used for inking, while the other hand performs coarser manipulations to position the document on the workspace and activate functions. However, it is also easy to perform two-handed multitouch actions with a simple movement of the pen-holding hand to let a finger enter into contact with the surface, without releasing the pen. Hence all gestures are executable without burdening the user with time-consuming and effort-requiring context switches.

The following list presents some of the main gestures used in our system in two categories: "Multitouch" and "Touch + Pen", depending on the input types involved. The "Touch + Pen" gestures are executed by combining a simple touch action and issuing a command with the pen. Because writing or drawing with the pen mobilises most of the user's focus and concentration, we think it makes sense not to require complex coordinate motions of the other hand as well, in order to not sacrifice speed and smoothness of the motor actions as well as cause cognitive strain. Thus, we designed the "Touch + Pen" gestures to adopt a two-step "activate and pen" model, where a finger of the non-dominant hand simply touches a specific interactive or hotspot area of the interface and is maintained pressed to select a function, whose parameters are then input by the pen. Depending on the nature of the function, the execution of the corresponding command is then triggered by lifting either the finger or the pen from the tabletop surface.

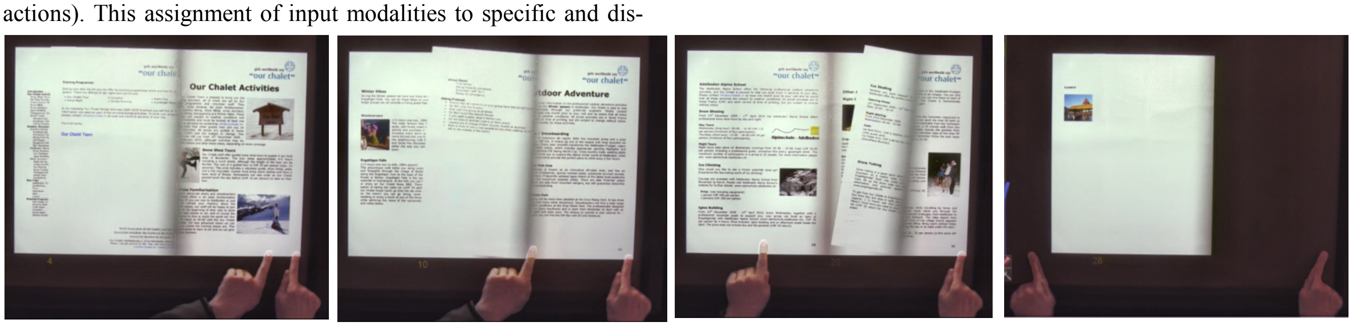

Figure 2. Forward multiple page flip: the right index activates the right corner and the left index moves laterally to flick through the pages

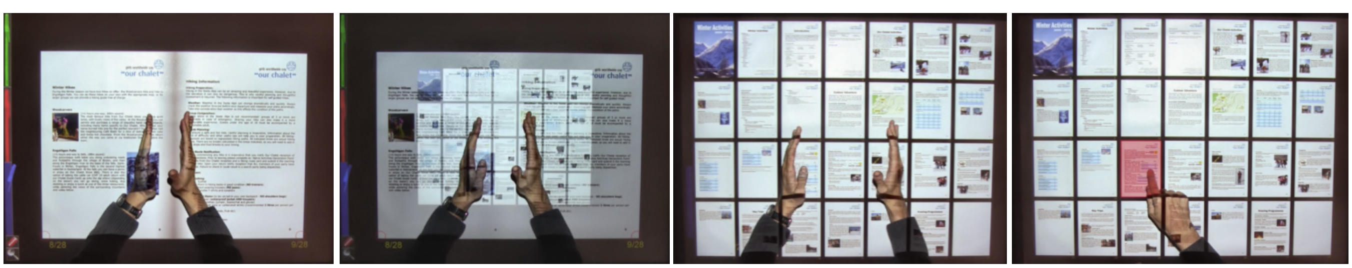

Figure 3. Chop and spread gesture to trigger the overview screen (1-3) in which a particular page to display can then be selected (4)

We believe that this ability to set the context of the pen input using appropriate touch activations enables more rapid command entry than unimodal solutions requiring explicit context switches performed by the same tool (pen or touch). Repeated successions of individual actions such as searching with pen-written keywords, annotating, navigating in the document etc. are therefore presumably more efficiently performed using such an input model. Furthermore, requiring that a finger must remain on the trigger area of a function to execute it ensures that the user is always consciously activating that function. We felt that tap-selections might be more disruptive as the user would have to constantly switch between pens and commands by tapping the respective icons, thereby also maybe forgetting which function is currently activated.

Multitouch

- Multiple page flip (Figure 2): In addition to turning pages individually by tapping their corners, users can execute this gesture to leaf through several pages of the document to casually browse through its content as if flipping the pages of a real book. It is initiated by pressing a finger in a page corner (typically the index) and sliding the fingertips of the other hand between the two lateral edges of the document to flick through its remaining pages. Depending on whether the left or right corner is activated, the pages before or after the currently opened page(s) are shown. The displayed page and the state of its folding animation are a function of the position of the non-static finger(s) where the edges of the document represent both extrema. The user can lift the finger and touch the surface again to initiate another page-flicking operation.
- Document overview (Figure 3): the document overview is summoned by executing a chop-and-spread gesture where the two hands are placed vertically on the surface, palms facing, and then spread apart, as if scattering all pages of the document around. The document overview screen gradually fades in and replaces the main interface as the hands separate, until a threshold distance is reached and the overview completely fills the screen. Some level of relaxation is permitted in that a fingertip can be used for one of the hands instead of a full chop posture which may be more convenient for the pen-holding hand. This gesture is somewhat similar to Wu et al.'s Pile-n-Browse gesture [24] although it fulfils a different purpose. If the document contains a large number of pages, it is possible to switch to a multi-level overview, where only a subset of the pages are first displayed (e.g. pages with chapters or main sections if the metadata is available) and the user can view the remaining pages contained in a virtual "pile" by tapping the top page.

Touch + Pen

As mentioned previously, touch actions in this mode only involve pressing and maintaining a finger of the non-writing hand on an active area of the UI and is hereafter simply referred to as "activating".

- Erase annotation: annotations can be removed by activating the delete icon in the command sidebar and striking them out with the pen. Several annotations can be erased using a single stroke that intersects at least one of the lines that make up said annotations. Thanks to this combination that does not rely on error-prone gesture or stroke recognition, annotations can be speedily and reliably erased.
- Go-to-page: it is possible to turn to a specific page of the document by activating one of the interactive page corners and writing the page number to turn to anywhere on the surface. When the finger is released, the handwritten number is recognised and the desired page is shown after a page-flipping animation.
- Searching with handwritten keywords (Figure 4): users can perform text searches in the document by activating the magnifying glass icon in the command sidebar and writing keywords anywhere on the surface. When the finger is released, the handwritten keywords are recognised and the results displayed. Matches in the currently open page(s) are highlighted, while hits located on non-visible pages appear with contextual snippets and the number of occurrences on the referenced page in tappable sticky note-style bookmarks at the side of the pages (Figure 5). If the number of bookmarks exceeds the space available at the side of the document page, navigation arrows appear at the bottom to cycle through the next or previous set of bookmarks. In the overview screen, terms matching search keywords are shown highlighted on all pages, although text might not actually be readable if the thumbnails are too small. Annotations incidentally benefit from a similar treatment, in that notes occurring on pages currently not visible are referenced by side bookmarks filled with the colours of the pen(s) used for the annotation(s) on the corresponding page.

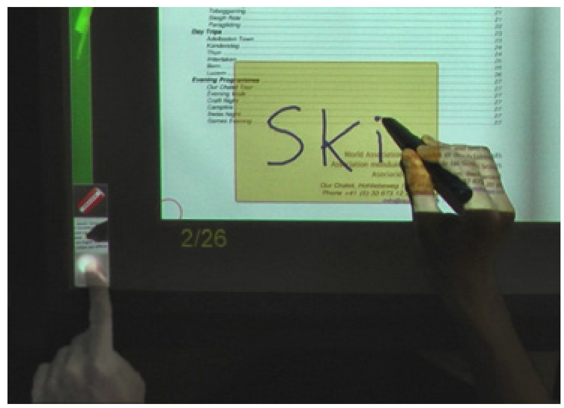

Figure 4. Handwriting a search keyword

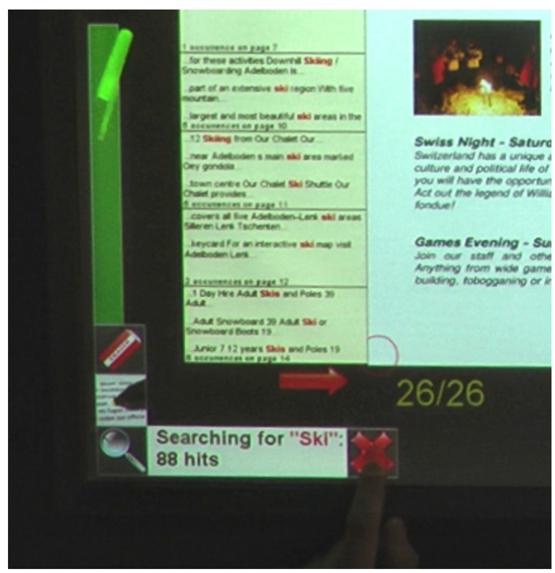

Figure 5. Dynamic bookmarks on the side of the page with context snippets for search results

- Searching with terms selected in the document: Keyword queries can also be executed by simply circling terms in the document. For that, the user activates the text marking function in the side bar and selects text in the document. The selection marquee appears as a dotted line to differentiate it from normal inking mode.

4. USER EVALUATION

When designing our study, we were particularly interested in gaining insights into the following questions: would people's individual interaction practices and behavioural habits for paper and computers successfully merge when using the pen and touch tabletop? Could they be more or at least as efficient on our system as on the medium they would otherwise naturally use to execute a given AR task? Would the provided tools and gestures be sufficiently accessible and easy to master so that users could become rapidly productive with our application? To address those interrogations, we created a set of experiments in which individual tasks would rely on different characteristics and inherent strengths of the considered media. For instance, we imagined that paper would prove more adequate than a regular computer for annotating, but that the latter would outshine the former in a task that involved much text searching. The question was then to see if our tabletop system including both pen interaction and search functions could combine those intrinsic advantages and prove to be efficient in all cases. Those considerations led us to lean towards a methodology with a series of tasks focusing on different aspects of AR separately rather than the summarisation exercise used by O'Hara and Sellen and Morris et al. [17, 19], hence the following protocol:

The study was comprised of three sets of data-searching and annotating tasks performed on three different media: paper, Adobe Acrobat on a regular desktop PC and the tabletop. The display for the desktop PC was a 19" monitor with a resolution of 1200 x 1024, whereas the projected image on the tabletop had a resolution of 1600 x 1200 (we estimated that a higher resolution was needed for the tabletop to compensate for the optical distortions introduced by the Anoto dot pattern).

We tried to design the tasks so as to reflect scenarios where information needs to be located in a particular document (in this case a PDF) following which comments can be added, for example for correction or summarising purposes. We timed participants when performing each assignment and observed the approach taken to solve it as well as the manner in which they interacted with the documents for later analysis. On the tabletop, we recorded the number of times users activated the different gestures and functions for each task. We applied counterbalancing measures by rotating the orders of documents and media between users in order to reduce bias and learning effects. Like Morris et al. [17] we chose a within-subject design so that testers could manipulate and directly compare the competing platforms in their feedback.

The study involved the participation of 20 people, 8 males and 12 females. The ages ranged from 21 to roughly 60 years old. All participants rated their ability to search for information on the Internet using keywords as good or very good. A majority also revealed they owned a touch-based device on which they read text content. As for their familiarity with PDF documents and Adobe software, most participants declared they used only the basic functions of the Reader to read and search within PDF files. A few testers had also tried the Text Edit tools in Acrobat, but none of them were proficient with them.

The first 2 persons were used as pilot testers, whose feedback was utilised to make a few adjustments to the study protocol that was then followed to perform the actual study with the remaining 18 people. Prior to the evaluation, participants were given adequate training on Acrobat and the tabletop with a focus on functions they were likely to need in order to fulfil the given tasks. During the execution of the tasks however, testers were left to decide which tools they preferred to employ and were not interrupted, except in cases where it was evident they had misunderstood the question.

In a nutshell, the evaluation consisted of three independent within-subject studies, each performed on all three media (i.e. three levels for the variable), that is, each participant had to execute 3x3=9 tasks in total.

4.1 Description of the Tasks

4.1.1 Visual Search

A set of 6 documents with roughly the same number of pages was laid out on a table and on top of each was placed a piece of paper with a question, whose answer had to be looked for inside the document. The questions referred to items easily identifiable with a glance of the eye, that is, photographs, illustrations, headlines, captions etc. so that participants were able to find the answers by simply flipping through the pages of the document without taking the pain to peruse the actual content (e.g. "Where are the objects that may not be taken on a plane?" in an airport guide). The idea was to determine whether people adopt different strategies when trying to casually glean information from paper documents compared to digital ones and whether gestures like the chop-spread for the document overview and the two-finger multiple-page flip would feel natural enough for the users to take advantage of. On a performance level, the average task completion time was expected to be comparable for the 3 media with perhaps a slight edge for the tabletop thanks to its overview functionality that could help spot relevant information more rapidly.

4.1.2 Text Search

A 20-page compilation of news articles was given to the participants with a list of 5 questions to answer in order. This time, people were told that the answers were located in the text of the articles and so they had to read or at least skim through the content to find the solutions. Because of the tediousness of such a task when performed on paper, however, the questions were formulated in such a way that the corresponding news article could be easily determined from them (e.g. "What was one of the most popular names given to male babies in Switzerland in the 1990s?"). On the digital platforms, of course, testers could avail themselves of the text search functions in order to directly locate the relevant passages and sometimes even the answer. It is evident that here, paper was at a significant disadvantage and so the goal was rather to see how the tabletop would fare compared to Acrobat, particularly with respect to the entry of handwritten vs. typed keywords.

4.1.3 Editing

Participants were handed the news articles again and given 7 annotation tasks to perform. Those included circling, crossing out and marking items using (regular or digital) pens of different colours (e.g. "Circle in red all dates on page 15"). The last 2 tasks required users to go back to annotations they had previously made, this in an attempt to simulate a short review of the modifications they had done in the document (e.g. "Draw a vertical line next to the text where you made the circles"). For the latter, it was expected that the "Comments list" feature of Acrobat [3] and the annotation bookmarks of our application would come in handy, but overall, that users would prefer paper and the tabletop to make pen annotations.

Through the execution of those tasks, participants were made to realise the advantages and disadvantages of the tested media as well as their suitability to perform those tasks. We asked users to report on their impressions and provide feedback in a questionnaire after they completed the tasks.

4.2 Results

We were pleased to see that our tabletop application proved efficient and pleasant to use among a large majority of the participants, despite people's initial unfamiliarity with the system. With only minimal training, users were able to learn most of the gestures and immediately apply them to solve the given tasks. The bimanual Touch + Pen gestures, in particular, were efficiently put into use, as users did not find it unnatural to use the pen on the same surface for inking as well as for command writing and to quickly switch between the two modes by activating a function with the other hand. One participant likened the operation to using keyboard shortcuts on the computer. Generally, users liked the larger surface of the tabletop, especially in the overview mode, where they could easily have a bird's eye view of the content of the document as well as a perception of its length. They appreciated the freedom to move the document around the surface and write on the pages just like paper on a desk. There were also a few frustrations reported, which we detail further below.

4.2.1 Platform Efficiency and Task Suitability

Table 1: Mean execution times in seconds with standard deviation for the three tasks Visual Search (VS), Text Search (TS) and Editing (E).

	Paper	Acrobat	Tabletop
VS	*M=200,SD=67*	*M=172,SD=53*	*M=174,SD=73*
TS	*M=394,SD=99*	*M=296,SD=121*	*M=294,SD=91*
E	*M=346,SD=93*	*M=454,SD=133*	*M=331,SD=81*

Measurements of the completion times for the three tasks are reported in Table 1. A one-way repeated measures ANOVA (with the sphericity condition verified by Mauchly's test) revealed that while for visual search users were not significantly faster on any of the three platforms ($F_{2,38} = 1.71$, $P = 0.19 > 0.05$), statistically significant disparities were observed in the other two tasks ($F_{2,38} = 31.91$, $P < 0.0005$ for text search and $F_{2,38} = 13.01$, $P < 0.0005$ for editing). Post-hoc pairwise comparisons confirmed what is already apparent in Table 1, i.e. that the last two tasks were executed at least as fast on the tabletop as on the next best medium and faster than on the third (all P values < 0.007 for significantly different mean execution times). For text search, users could take advantage of the appropriate functions on Acrobat and the tabletop in order to almost instantly locate the relevant contexts of the sought information, something that was of course not possible on paper. Furthermore, people were able to be almost as efficient handwriting keywords as typing them with a keyboard. The few occasional recognition errors were indeed compensated by the availability of term highlighting and bookmarks with context snippets to help find the required information.

For the editing tasks, participants found themselves in familiar territory with the digital pen, with the added benefit compared to its felt tip counterparts that mistakes could be easily deleted. Acrobat's commenting tools proved too cumbersome and understandably users felt more at ease with regular or digital pens to mark and annotate text. Those feelings were confirmed when study participants were asked to rate on a Likert scale their overall experience with the three media as well as for the three individual tasks, particularly with respect to the suitability of the medium to execute said task. Those results are reported in Figure 6.

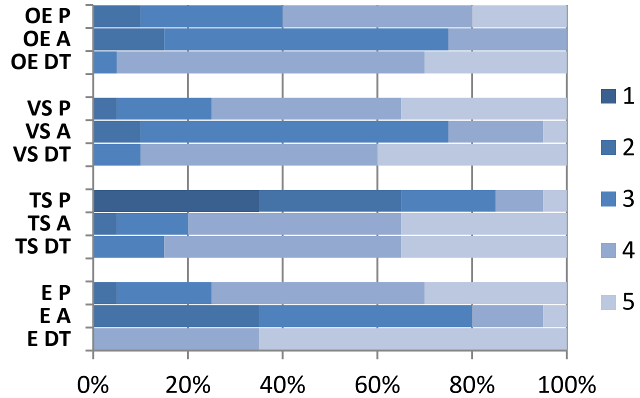

Figure 6: Participants' ratings of their overall experience (OE) on the three media and the latter's suitability to perform the given tasks from 1=Very Poor to 5=Very Good.

Friedman tests performed for all four comparisons exhibited statistical significance in each case so post-hoc Wilcoxon signed-rank tests with Bonferroni adjustments were made to determine pairwise significance. For overall experience, the tabletop was rated significantly higher than Acrobat ($Z = -3.493$, $P < 0.0005$) and even paper ($Z = -2.524$, $P = 0.012$). For the visual search task, participants found that the tabletop was significantly more suitable than Acrobat ($Z = -3.508$, $P < 0.0005$), but not with respect to paper ($Z = -1.076$, $P = 0.282$). In text search, the tabletop was rated equally suitable as Acrobat and both significantly higher than paper ($P < 0.0005$). For editing, our application was rated as more suitable than Acrobat ($Z = -3.831$, $P < 0.0005$) and even paper ($Z = -2.754$, $P = 0.006$), which was somewhat surprising. User feedback revealed that while they still felt more comfortable marking paper, they valued the additional tools provided by our application. Two users commented that they imagined in a real situation they would want to store their annotations digitally and so they welcomed the potential benefit of being able to do that on the tabletop.

4.2.2 Function Rating

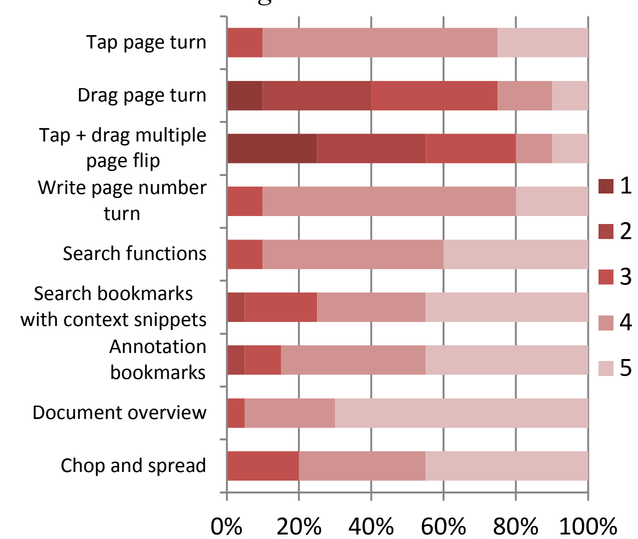

Figure 7: Users' ratings of the usefulness of the different gestures and functions ranging from 1=Not useful at all to 5= Extremely useful.

Although participants were instructed, prior to the task assignments, in the use of all gestures and functions of the tabletop without any particular bias for any of them, the latter did not all enjoy equal popularity. The post-task questionnaire requested testers to assess the general usefulness and practicality of those features, not necessarily in association with the given tasks (Figure 7). The answers they supplied very closely matched the frequency in which they used the different functions, as logged by the system. For instance, users found the overview feature and its associated gesture very helpful and hence used it extensively, whereas the bimanual multiple page flip, on the other hand, saw little appeal and so was almost never used. This is most likely due to the difference in complexity and precision required by the gestures, as the former essentially executes a transition from one state to another and can be triggered anywhere on the surface, whereas the latter maps a precise position of the finger to a particular visual state (the multiple pages being viewed and their folded appearance as they are being turned). However, further observation of selected test users, who spent longer time with the application, hinted that this discrepancy could also be put down to learnability and awareness issues. Some gestures simply take more time to master than others.

4.2.3 System Limitations and Discussion

Due to a number of hardware and software constraints as well as design choices, the application suffered from a few limitations that more or less interfered with the execution flow of the user's work. Study participants were asked to report on these impediments and how far they were affected by them. The results are presented in Figure 8.

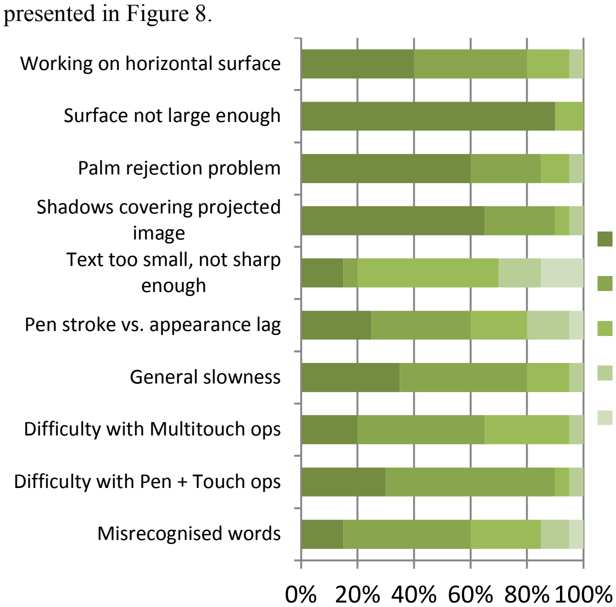

Figure 8: Hardware and software limitations of the system and how seriously users felt their experience was affected by them, ranging from 1=Posed no problem at all to 5=Extremely problematic.

Despite the higher resolution of the projected image on the tabletop compared to the PC monitor, the most serious hindrance was the insufficient sharpness and crispness of the text preventing pleasant viewing and reading of the document pages, especially in the overview screen. As hinted at in the system description section, this lack of adequate rendering quality is mostly due to the dot pattern on the overlay but also to some extent to the display resolution as well. We think this issue of resolution depth, both for display and input capturing, is critical and needs to be addressed by tabletop manufacturers if they are to support a wider range of applications requiring high-quality rendering of object details such as AR and generally document-centric applications. A possible yet more expensive solution for the viewing problem could be to utilise high-definition projectors or multi-projector displays [2]. As for sensing, Anoto technology, which was initially designed for asynchronous recording of pen inking on paper, has been adapted to support streaming position data, but because it relies on a printed dot pattern it is not ideal to integrate with touchscreen systems. Moreover, optical pens are still fairly bulky and need to be regularly recharged. The problem of high-resolution and high-accuracy pen input sensing for tabletop systems therefore remains.

Another reported problem was the misrecognition of handwritten keywords that sometimes caused a little frustration to some users. However, during the experiments, the misrecognition problem only occurred for a small number of participants, who had trouble adapting to the writing style required for the handwriting recognition engine to accurately determine the input text. With more training and better support by the interface, this handicap could most likely be eliminated or at least alleviated. In terms of speed and efficiency, however, it would be interesting to compare it with a soft keyboard solution.

Among the factors that posed only minor disturbances was the signal interference of the resting palm or arm on the tabletop while writing. Participants in the study were told to pay attention not to come into contact with the touch surface when using the pen, but many of them did so naturally and so were not troubled by that constraint. The evaluation tasks did not involve much inking, which is also a reason why users might not have felt too inconvenienced. One can imagine, however, that more pen-intensive tasks requiring a great deal of manual text entry or sketching could cause wrist fatigue in the long run and so we believe this problem merits consideration.

Finally, an often pointed out drawback of front-projected systems, namely the occlusions caused by shadows cast by arms reaching over the tabletop, turned out not to be an issue for the vast majority of the study participants. Shadows are natural occurrences that most people can cope with, especially if they can be controlled [21]. One interesting matter that was brought up by a left-handed user, however, was that shadows did get in the way when writing with a curled left hand (i.e. almost from above), since in that case the pen-holding arm and wrist cover the letters that have just been written as the arm moves from left to right. But all things considered, the occlusion problem of front-projected displays was found to be of minimal discomfort to the users and so the use of such systems should not necessarily be discouraged, at least in those kinds of usage scenarios.

5. DISCUSSION

The application we developed and the experiments we conducted were a first attempt to sound out the potential of interactive pen and touch tabletops. We opted to remain fairly close to the paper metaphor without strictly adhering to it when we felt practical concerns dictated that the interface should adopt a more straightforward approach to support useful functions (e.g. searching, overview mode). Yet, the user study revealed that we may sometimes have gone too far in trying to model paper-like behaviour. The multiple page flip gesture as it was implemented is an example where users found it added little to the program's functionality. The lesson one can draw from that when designing systems with natural user interfaces (NUI) is that there needs to be a balance between the desire to offer a close-to-real interaction paradigm to appeal to users' sense of familiarity and the necessity for the system to provide handy and effective tools to execute the target tasks.

The participants in our study also made suggestions on how they perceived we could improve and enhance the application. We are aware that our system is not feature-complete and that further work is needed to support other AR requirements such as content extraction and multiple documents as well as collating information with external sources. We believe however, that we have prepared the ground for a more extensive and feature-rich platform based on pen and touch interaction. We have already started extending our search functions so that the handwritten or selected keywords can also be used to retrieve external reference material, such as dictionaries, Wikipedia articles, translations etc. that could be of help for AR.

6. CONCLUSION

In this paper, we considered the application of active reading to digital tabletops with pen and touch interaction capabilities. Inspired by prior work and intuition, we designed an application prototype addressing the basic needs of AR. The system and the experiments conducted with it showed that pen and touch interaction paradigms bring several advantages that we think can be extended to other document-centric scenarios. There is however still much work to be done to further the cause of digital tabletops as virtual work desks. On the software side, it remains to be seen how document functions people commonly use and expect to find in a productivity system can be efficiently integrated in a large, keyboardless surface setup. On the hardware side, interactive tabletops and surface devices in general need to meet certain resolution and reactivity standards in order to guarantee a smooth and (mostly) pains-free experience. Future systems will have to address those requirements if they are to support a broader scope of professional applications.

7. ACKNOWLEDGMENTS

We would like to thank Matthias Geel and Maximilian Speicher for their work adapting the iPaper and Anoto frameworks to suit the requirements of the application and Michael Haller for his valuable help and feedback when editing this paper.

8. REFERENCES

[1] Anoto AB. http://www.anoto.com

[2] Ashdown, M., Tuddenham, P. and Robinson, P. High-resolution interactive displays. In *Tabletops - horizontal interactive displays*, Springer Verlag (2010), 71–100.

[3] Baker, D.L. Communicating with Comments in Adobe Acrobat 8. In *Adobe Acrobat 8 in the Office*. Adobe Press (2006).

[4] Brandl, P., Forlines, C., Wigdor, D., Haller, M. and Shen, C. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *Proc. AVI 2008*, 154– 161.

[5] Card, S.K., Hong, L., Mackinlay, J.D. and Chi, E.H. 3Book: A 3D Electronic Smart Book. In *Proc. AVI 2004*, 31, 303–307.

[6] Cockburn, A., Gutwin, C. and Alexander, J. Faster Document Navigation with Space-Filling Thumbnails. In *Proc. CHI 2006*, 1–10.

[7] Dietz, P. and Leigh, D. DiamondTouch: a multi-user touch technology. In *Proc UIST 2001*, 219–226.

[8] Frisch, M., Heydekorn, J. and Dachselt, R. Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. In *Proc. Tabletop 2009*, 167–174.

[9] Guiard, Y. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. In *Journal of Motor Behavior* (1987). 19(4), 486–517.

[10] Hinckley, K., Rodenhouse, J., Yatani, K., Wilson, A., Pahud, M., Benko, H., Coddington, N., Buxton, B. Manual deskterity: an exploration of simultaneous pen+ touch direct input. In *Proc. CHI 2010*, 2793–2802.

[11] Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H. and Buxton, B. Pen + Touch = New Tools. In *Proc. UIST 2010*, 27–36.

[12] Hinckley, K., Zhao, S., Sarin, R., Baudisch, P., Cutrell, E., Shilman, M., Tan, D. InkSeine: In Situ Search for Active Note Taking. In *Proc. CHI 2007*, 251-260.

[13] Kauranen, B. The ANOTO pen - Why light scattering matters. http://www.acreo.se/Global/Meet-us-at/OE2000/00-KAURANEN.pdf.

[14] Leitner, J., Powell, J., Brandl, P., Seifried, T., Haller, M., Dorray, B., To, P. Flux - a tilting multi-touch and pen based surface. In *Proc. CHI 2009*, 3211–3216.

[15] Levine, S.R. and S.F. Ehrlich. The Freestyle System: A Design Perspective, in Human-Machine Interactive Systems, A. Klinger, Editor. 1991. 3–21.

[16] Liao, C., Guimbretière, F., Hinckley, K. and Hollan, J. Papiercraft: A Gesture-Based Command System for Interactive Paper. In *ACM Transactions on Computuer-Human. Interaction*, 14, 4 (Jan. 2008), 1–27.

[17] Morris, M.R., Brush, A.J.B. and Meyers, B. Reading Revisited: Evaluating the Usability of Digital Display Surfaces for Active Reading Tasks. In *Proc. of Tabletop. 2007*, 79-86.

[18] Norrie, M. C., Signer, B. and N. Weibel. General Framework for the Rapid Development of Interactive Paper Applications. In *Proc. CoPADD 2006*, 9–12.

[19] O'Hara, K. and Sellen, A. A Comparison of Reading Paper and On-Line Documents. In *Proc. of CHI 1997*, 335-342.

[20] Schilit, B.N., G. Golovchinsky, and M.N. Price. Beyond paper: supporting active reading with free form digital ink annotations. In *Proc. CHI 1998*, 249–256.

[21] Summet, J., Abowd, G. D., Corso, G. M. and Rehg, J. M. Virtual rear projection: Do shadows matter? In *Proc. CHI 2005*. 1997–2000.

[22] Tashman, C.S. and Edwards, W.K. Active reading and its discontents: the situations, problems and ideas of readers. In *Proc. of CHI 2011*, 2927-2936.

[23] Tashman, C.S. and Edwards, W.K. LiquidText: a flexible, multitouch environment to support active reading. In *Proc of CHI 2011*, 3285-3294.

[24] Wu, M., Shen, C., Ryall, K., Forlines, C., and Balakrishnan, R. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. In *Tabletop 2006*, 183–190.

[25] Yee, K. Two-handed interaction on a tablet display. In *Proc. CHI 2004*, 1493–1496.

[26] Zeleznik, R., A. Bragdon, F. Adeputra, and H.-S. Ko. Hands-on math: a page-based multi-touch and pen desktop for technical work and problem solving. In Proc. *UIST 2010*, 17–26.