Digital manuscripts and electronic publishing

Peter Robinson
University of Cambridge Computer Laboratory
March 1998

ABSTRACT

The publication of Vannevar Bush's seminal paper, *As we may think*, in 1945 was closely followed by the invention of the electronic digital computer in 1948. In the following 50 years much of Bush's vision has been realised by computer scientists. We can store vast amounts of data including speech and video in computers, and we can transmit them over networks. We can annotate the material with editorial notes and electronic cross references. However, our interaction with the material requires the use of a keyboard and mouse rather than the more natural pencil and paper. This paper describes how this problem is being addressed and will discuss its implications for scholars in the future.

INTRODUCTION

Scientists played a crucial role in the 1939-1945 World War. Most obviously we think of the atomic bomb that finally brought the war with Japan to an end, but there were other important technological contributions in areas such as radar or the mechanical and electronic calculation of ballistic trajectories and code breaking. These results spilled out into civilian academic research at the end of the war.

Vannevar Bush had witnessed many of these developments while serving as Director of the Office of Scientific Research and Development for the American government in the war. Moreover, he had also experienced the difficulties of managing the information arising from the scientists' research efforts. In July 1945 he published a vision of the future in the popular magazine, *Atlantic Monthly* [Bush 1945]. In it he identified the management of information as the great challenge for the future and posed three problems:

- How could an enduring record of ideas be maintained?

- How could the evolution of information through extraction and manipulation be supported?

- How could communication between individuals be assisted?

He went on to propose a technological solution, a machine which he called the *Memex*, that embodied many of the advances that had emerged or were foreseen during the war. In many ways it could be considered as a precursor of the computer workstation of today. He envisaged a desk-sized piece of furniture with a couple of large display screens, a keyboard and various knobs and levers to control the electromechanical workings inside. It used microfilm to store information together with annotations recorded using "dry photography" (today's xerography). The screens could be used for reading and also for scanning paper documents that were to be added to the repository.

However, he envisaged more radical ways of manipulating the data. In particular, he saw the need to link different pieces of data in separate places through "trails" (today's hyperlinks), and also saw that these trails would be valuable pieces of information in their own right. He also looked to voice recognition as a means of controlling the Memex and wireless communications as a means of accessing it remotely.

Clearly Bush was aware of the work on electronic calculators and their application to code-breaking that had been undertaken at the British Government's Cipher School at Bletchley Park during the war, although it was sufficiently highly classified that it could not be mentioned explicitly. However, the seeds had been sown and work progressed on the electronic computers both in America and at Manchester [Williams 1948] and Cambridge in England. In May 1949 the EDSAC [Wilkes 1948], the world's first practical electronic stored-program computer, successfully ran a program to print a table of squares and information age was born.

Twenty-five years later, in the 1970s, the world's first personal computer, the Alto, was demonstrated at the Xerox Corporation's Palo Alto Research Center [Thacker 1982]. Although earlier computers had been limited to a single user, this was the first that was designed to be a personal workstation. It included an A4-sized black-on-white raster display and featured a user interface that made extensive use of a mouse to control a window system and for other interaction. Parallel developments included the ethernet local area network and the laser printer.

Fuelled by decreasing costs of hardware, the personal computer with a graphical display and interaction through a mouse and windows has become the model for computing in the last decade of this century. In many ways, this fulfils Bush's vision. The personal computer provides an enduring record of ideas, word processing, assists the evolution of information, hypertext builds trails through the different pieces of data and networks allow communication between individuals. The desktop metaphor even models the physical design of Bush's Memex.

Is this the answer or should we aspire to better things?

At one extreme there are proponents of *artificial reality* who would immerse us in a synthetic world projected into a special headset. The *Origami* project at the University of Cambridge Computer Laboratory has taken the gentler line of *augmented reality* in which computational properties are added to everyday objects.

There is no desktop metaphor for computing on a screen because the desktop itself becomes part of the computer. We have investigated and implemented new methods for human-computer interaction, and applied and evaluated their use in the specific discipline of electronic publishing.

ANIMATED PAPER DOCUMENTS

Electronic, multi-media publishing is becoming established as an alternative to conventional publishing on paper. CD-ROM and on-line versions of reference books and fiction can augment their conventional counterparts in a number of ways:

- They offer elaborate indexing, glossaries and cross-referencing.

- They allow non-linear progression through the text.

- Sound and moving images can be added.

- Sections can be copied into new documents.

However, screen-based documents have a number of disadvantages:

- People find screens harder to read than paper.

- Electronic bookmarks are less convenient than bits of paper or flicking through a book.

- Adding personal notes to electronic documents is more complicated than jotting in the margin of a book.

- Writing, editing and proof-reading a non-linear, multi-media document is still a specialised and difficult task.

Our solution is to publish material as an ordinary, printed document that can be read in the normal way, enjoying the usual benefits of readability, accessibility and portability. However, when observed by a camera connected to a computer, the material acquires the properties of an electronic document, blurring the distinction between the two modes of operation [Robinson 1997a].

This builds on the construction at the Computer Laboratory in Cambridge of the *DigitalDesk* – an ordinary desk augmented with a computer display using projection television and a video camera to monitor inputs. A number of prototype systems had been implemented to demonstrate its feasibility [Robinson 1995, Stafford-Fraser 1996, Wellner 1993]. We have now investigated the technology further, addressed issues of performance and explored and evaluated its application to electronic publishing. In particular, we have developed a strategy for combining electronic and printed documents to give a richer presentation than that afforded by either separate medium. This also related to earlier work in Cambridge on educational computing materials, especially *Computer Illustrated Texts* [Harding 1993].
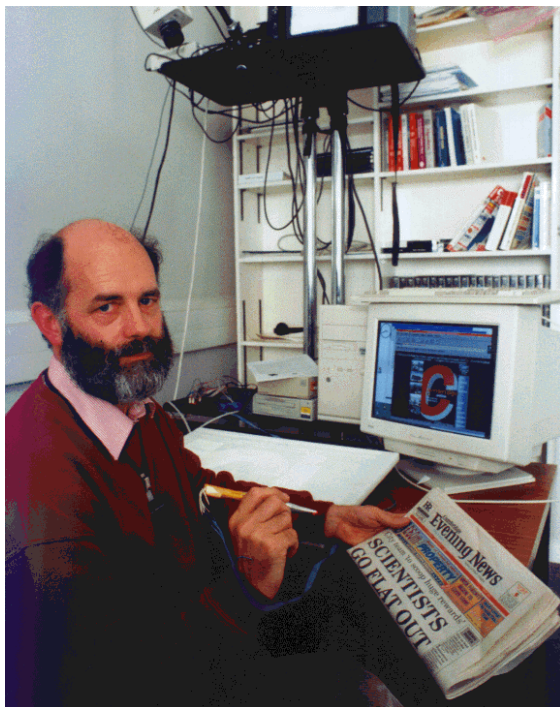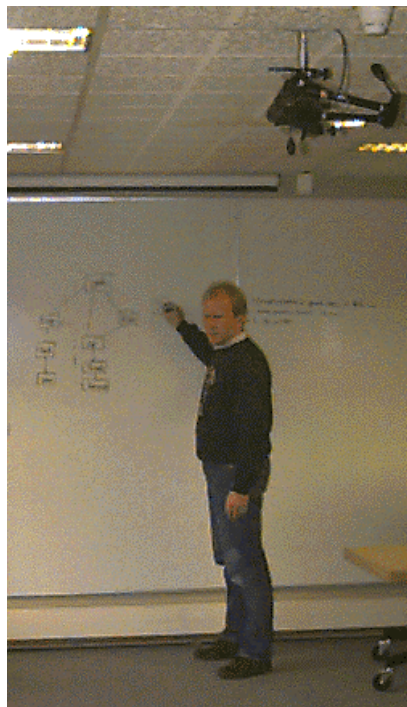
*DigitalDesk*   *BrightBoard*

As a result we have proved the concept of combined documents and demonstrated their enhanced functionality, paving the way for subsequent development of mixed-media publications through:

- the design and implementation of an architecture for mixed-media documents,

- the construction of adaptors to create, manipulate, present, import and export such documents,

- demonstration applications for Web browsing, electronic mail and the teaching of maths and grammar,

- underlying technology to locate and identify printed documents, to extract text and images from them and to allow direct interaction with them.

AN ARCHITECTURE FOR MIXED-MEDIA DOCUMENTS

We have investigated two document structures, one based on SGML and one based on active objects [Robinson 1997b]. Both provide links to other documents and to the programs implementing annotations. These can be held in a separate registry which allows links embedded in a paper document to be activated remotely. When used on the DigitalDesk, the system resembles a Web browser working from paper rather than a screen. Links can be activated by pointing with a stylus and forms can be completed simply by writing.

The overall architecture of the animated paper document system is shown below:
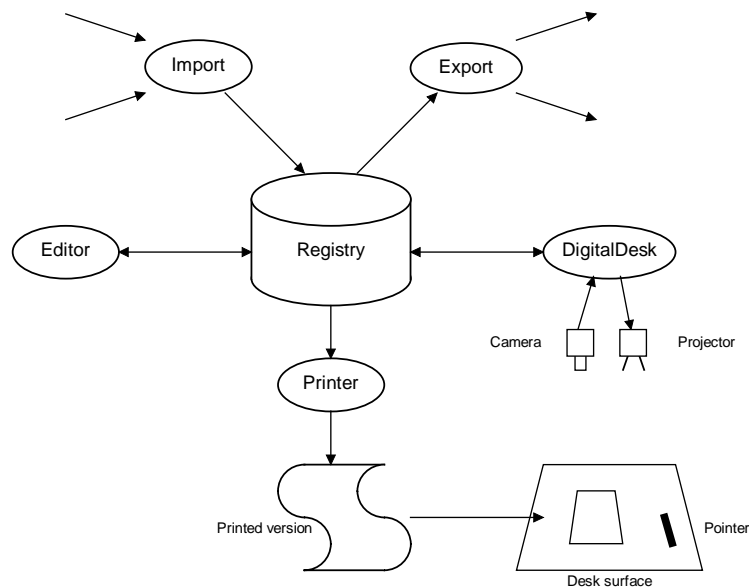
Animated paper document framework

At the core of the system is a *Registry* which maintains the association between electronic documents and their printed variants. It stores the image of each active document and the code of any interactions required for the document, together with cross references between these and further indexes to identify them. In the context of WWW documents, these correspond to links to other URLs, but the facilities allow much more general forms of interaction.

CREATION, MANIPULATION AND PRESENTATION

The central registry is surrounded by a suite of adaptors linking it to other information sources and to interaction with printed documents.

Import

Conventional hypertext can be absorbed into the animated paper document system; paper access to the World-Wide Web is possible through such an adaptor. Given a uniform resource locator (URL), the adaptor captures the information currently on the associated Web page in the registry. This includes the URLs of any links embedded within the page.

An HTML parser breaks the document into blocks of text (usually paragraphs, but at a finer grain where there are links) and images. These are then rendered as PostScript and the positions and content of the links recorded. All this information is kept in the registry. The page can then be printed simply from the PostScript, with further embellishment to assist subsequent page recognition.

Editing

Documents in the registry can be edited with a fairly conventional WYSIWYG editor. Text and diagrams are entered and amended in the usual way. However, it is also possible to mark areas of the document as hyperlinks and associating interactors with them. These are recorded as references to the associated code.

One version of the editor actually operates on the DigitalDesk, which means that text, diagrams and interactors from other printed documents can be copied into the new document. If the other printed documents are active documents known to the system, this copying is entirely digital, just as it would be in a conventional word processor. However, text and pictures can also be copied from conventional printed documents by using the overhead camera to capture an image and passing any text through an optical character recognition system.

Printing

Another adaptor prints out documents from the registry onto paper so that they can be used for direct interaction on the DigitalDesk. The printed documents are annotated with marks in their corners to facilitate recognition and location on the desk top, and are also have a unique identifier printed in an OCR fount.

Once the document has been printed, its contents are retained in the registry as an immutable copy of its structure for future interaction. This allows the paper to continue working in the same way even if its electronic original is edited. However, any URLs referred to in the electronic version will have been remembered and will be followed when the paper version is animated. The contents of the pages identified by such URLs can change in the usual way.

The paper version can be carried away, read, annotated, posted or even sent by fax and will still retain its hypertext attachments whenever it is placed on another DigitalDesk.

DigitalDesk

The DigitalDesk actually animates the paper documents. This involves recognising that a page printed by the system has appeared on the desk, determining its position, reading its unique identifier and locating any interactors. A transformation is then set up between the page representation stored in the registry and physical co-ordinates on the desk top. The printed document thus becomes part of the projected window system. In particular, any active links are highlighted by projecting a red background over them. For a document originating on the Web, these correspond to links in the original HTML.

A pen with a light-emitting diode in its tip is used for pointing. This is recognised by the camera system and converted to co-ordinates using a transformation calculated by occasional registration. It would be possible to use a conventional graphics tablet, but the light pen has the advantage that it works perfectly well over a stack of paper on the desk. The events are passed back through the window system and interpreted using information in the registry. For a URL, this involves opening a new projected window on the desktop and displaying the contents of the associated page in it.

Export

The interactions afforded by animated paper are considerably richer than straightforward HTML but if a document is sufficiently simple, it can also be exported as HTML.

This involves scanning the image of the page from top to bottom, left to right and emitting text or images as appropriate. When a page is published in this way, a series of HTTP PUT commands are sent to the WWW server which is going to hold the page. One of these is for the HTML of the page itself, the others are for the richer features of DigitalDesk documents that can not be translated into conventional HTML. These can be recovered either through another DigitalDesk or by a suitably extended WWW browser.
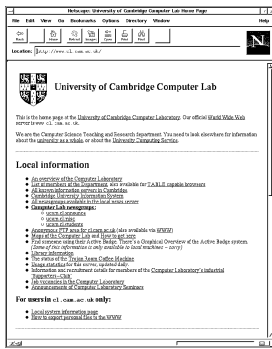
APPLICATIONS

Several demonstration applications have been written to exercise this framework.

Web browsing

The system can be used to interact with printed versions of Web pages [Robinson 1997c]. The figure below shows a conventional WWW page at (a). This is imported into the system and reprinted with additional coding to assist recognition (b). When this is placed on the DigitalDesk it is recognised and active areas of the document illuminated by projected highlights. When these are selected, links are followed or programs executed and the results projected into a further window on the work surface (c).
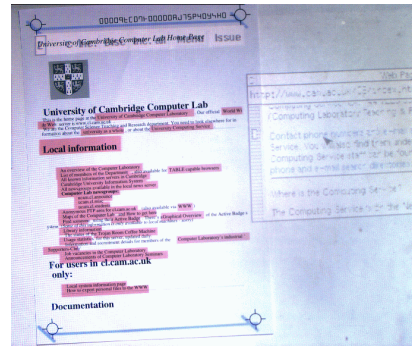
Moreover, fragments can be copied from the paper document into new electronic documents also projected onto the desk (d). The new document can be printed to give a new paper document (e) which can be animated on the desk in just the same way (f).
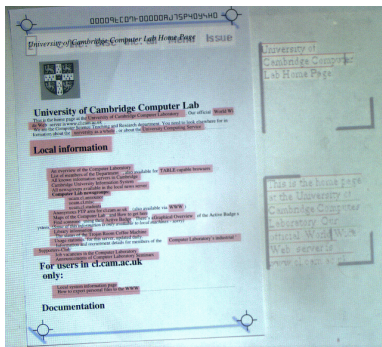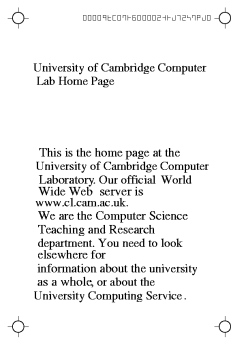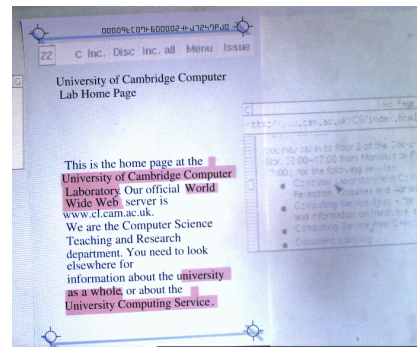
(a) Original Web page.
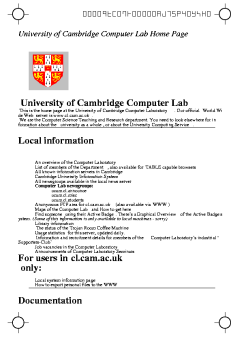
(b) Printed version.

(c) Animated on the DigitalDesk.

(d) Deriving a new document.


(e) Printed version.

(f) Animating the new document.

Paper access to the World-Wide Web

Educational texts

Two further applications explore the use of this technology for educational material. The first is a course book for teaching mathematics [Harding 1996]. The software which accompanies the course book is automatically launched when the book is first placed on the desk. The picture below shows a section on curve-sketching for polynomials. The generic equation of a quadratic polynomial is given with spaces for the values of the coefficients and an empty box underneath for plotting the graph. The software projects default values and draws the graph into the box. However, it also projects controls alongside the coefficients to allow the reader to change these values while observing the corresponding change in the graph.

Further down the page of the course book there is an assessment exercise. This time the polynomials are fixed and the student must draw the curve into the box (the active pen also has a real nib for writing). Clicking a projected button asks the computer to assess the sketch. The image is captured and analysed for features such as maxima, minima and axis crossings, and marked accordingly.
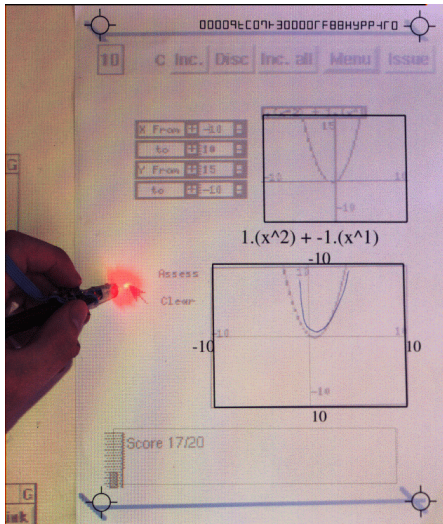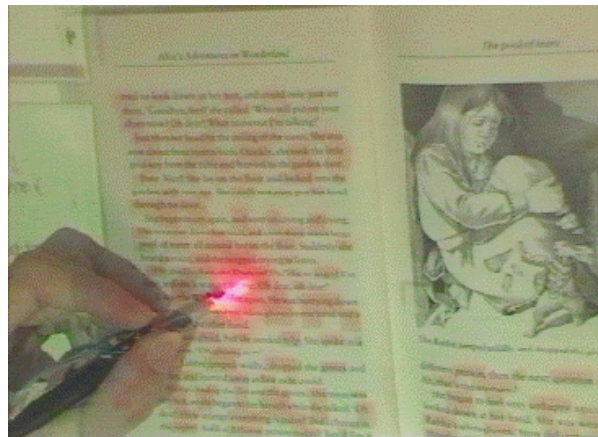
A maths tutor   *A grammar tutor*

A second educational application teaches elementary grammar by animating a standard printed book [Brown 1997, Brown 1998]. This uses additional information from an SGML edition of the book distributed as part of the Text Encoding Initiative.

Paper mail

A final application integrates paper notes with electronic mail. A note written on a piece of paper can be placed on the DigitalDesk and a recipient selected from a projected menu. An image of the note is captured and send as an e-mail attachment to the recipient.

UNDERLYING TECHNOLOGY

A substantial effort has been put into the vision system that underpins interaction on the DigitalDesk [Sheppard 1997]. This has involved work on lighting compensation, registration of projected and captured images, a driver for variable resolution image capture, page recognition and optical character recognition to yield text from images.

The registry and adaptors run as a distributed computing system using the unique identifier on each page to locate the home registry for the page and then loading attachments over the Internet. The system is extremely modular and the various components can easily be rearranged or supplemented to provide new facilities.

The user interacts with the system directly through the paper which is modelled as a colleague or instructor, watching the interactions and offering advice when appropriate. This involves locating a piece of paper on desk, identifying the page through its unique identifier and contacting the home registry, transforming co-ordinates, accepting pointing and written input from the pen, capturing images and possible optical character recognition, projection of annotations back onto the page and dealing with printed or projected menus and controls.

CONCLUSIONS

The Origami project has refined the underlying technology of the DigitalDesk and applied it to electronic publishing of educational material, resulting in a system for combining electronic and printed documents to give a richer presentation than that afforded by either separate medium. It fulfils much of Bush's vision. There is an enduring record of ideas, albeit using digital storage rather than microfilm. Trails are provided by hypertext and there natural interaction with the system to extract and manipulate information. Communication between individuals is possible either by publishing on CD-ROM or through digital communications networks.

We now want to consider the problems of presenting this technology to a wider audience. This will involve achieving the same interaction through paper without video cameras, using simpler feedback than projection and

exploiting communication networks for publishing the electronic part of mixed-media documents. A proposal is being prepared to investigate the use of mixed-media documents as a vehicle for interacting with network computers and to demonstrate further applications in electronic publishing for education.

ACKNOWLEDGEMENTS

REFERENCES

1. H Brown, RD Harding, SW Lay, P Robinson, DP Sheppard & RR Watts: *Active paper for active learning*, Proceedings 4th annual conference Association for Learning Technology, Telford, September 1997, reprinted in Association for Learning Technology Journal, 1998.

2. H Brown, RD Harding, SW Lay, P Robinson, DP Sheppard & RR Watts: *Active Alice - using real paper to interact with electronic text*, Proceedings 7th International Conference on Electronic Publishing, Saint Malo, April 1998, ISBN 3 540 64298 6, pp 407-419.

3. V Bush: *As we may think*, Atlantic Monthly, July 1945.

4. RD Harding: *Computer Interactive Texts*, Proceedings of the Fourth ICTCM, Addison-Wesley, 1993.

5. RD Harding, SW Lay, P Robinson, DP Sheppard & RR Watts: *New technology for interactive CAL - the Origami project*, Proceedings 3rd annual conference Association for Learning Technology, September 1996, reprinted in Association for Learning Technology Journal 5(1), 1997, ISSN 0968 7769, pp 6-12.

6. P Robinson: *Virtual offices*, Proceedings of Royal Society discussion meeting on Virtual Reality, July 1995, British Telecom Publication number SRD/R5/1.

7. P Robinson, DP Sheppard, RR Watts, RD Harding & SW Lay: *Animated paper documents*, Proceedings HCI '97, San Fransisco, August 1997, reprinted in Design of computing systems: social and ergonomic considerations 21B, Elsevier 1997, ISBN 0 444 82183 X, pp 655-658.

8. P Robinson, DP Sheppard, RR Watts, RD Harding & SW Lay: *A framework for interacting with paper,* Proceedings Eurographics '97, Computer Graphics Forum 16(3), September 1997, ISSN 0167 7055, pp 339-324.

9. P Robinson, DP Sheppard, RR Watts, RD Harding & SW Lay: *Paper interfaces to the World-wide,* Procedings WebNet '97. Toronto, November 1997, ISBN 1 880094 27 4, pp 426-431.

10. P Robinson: *Digital manuscripts and electronic publishing*, invited presentation, International Congress on Production and Context, Constantijn Huygens Institute, The Hague, March 1998.

11. DP Sheppard & RR Watts: *Some DigitalDesk algorithms*, working paper, June 1997, to be revised for publication.

12. Q Stafford-Fraser & P Robinson: *BrightBoard - a video augmented environment,* Proceedings ACM Conference on Computer-Human Interaction, April 1996, ISBN 0 201 94687 4, pp 134-141.

13. CP Thacker, EM McCreight, BW Lampson, RF Sproull & DR Boggs: *Alto - a personal computer*, in D Sieworek, DG Bell & A Newell: *Computer structure - principles and examples*, Mc Graw-Hill, 1982.

14. PD Wellner: *Interacting with paper on the DigitalDesk*, Communications of the ACM 36(7), July 1993, pp 87-96.

15. MV Wilkes: *The design of a practical high-speed computing machine - the EDSAC*, Proceedings of the Royal Society 195(265), 1948.

16. FC Williams & T Kilburn: *Electronic digital computers*, Nature 162(487), 1948.

THE AUTHOR

Peter Robinson is a Lecturer in the Computer Laboratory at the University of Cambridge in England, where he is part of the Rainbow Group working on computer graphics and interaction. He is also a Fellow and Director of Studies in Computer Science at Gonville & Caius College. His research interests are human-computer interaction, where he works on the use of video and paper as part of the user interface, and electronic CAD for self-timed circuits. He studied Mathematics at the University of Cambridge and continued with a PhD in Computer Science. He is a Fellow of the British Computer Society and a Chartered Engineer.